Information as Foundation Principle for Quantum Mechanics

Bachelor Thesis
for the Degree of
Bachelor of Science (BSc)
at the
University of Vienna

submitted by

Christoph Regner

supervised by

Ao. Univ. Prof. Dr. Reinhold A. Bertlmann

Vienna, March 2, 2015

Table of Contents

List of Figures

1 Introduction

1.1 The Present Situation in Quantum Mechanics

At the beginning of the previous century, scientists had to face the problem that all the physical theories already known at that time could not explain some distinct phenomena observed at the atomic level, such as the photoelectric effect. Furthermore these theories even predicted outcomes that apparently contradicted experimental results, like electrons moving on "classical" orbits and crashing into the atomic nucleus due to loss of energy or the so-called ultraviolet catastrophe associated with black-body radiation.

The solution to all these problems was found in the 1920s and lead to the invention of a completely new theory of physics – *quantum mechanics*. Since that time scientists were able to explain numerous phenomena in nature and along with the theory of relativity quantum mechanics has become the most successful physical theory in the 20th century.

Despite its great success over the last few decades and the progress that has been made in the understanding of natural processes, qantum mechanics differs significantly from other physical theories as pointed out by Anton Zeilinger in [1]. Zeilinger argues that the crucial difference concerns primarily the conceptual foundations of quantum mechanics, that is, in contrast to most physical theories, like the theory of relativity or classical mechanics based on Newton's laws of motion, quantum mechanics has not yet been grounded on consistent fundamental principles.

At this point it is certainly necessary to clarify that the term "foundations of a physical theory" in the sense it is basically used in the present work does not refer to the underlying mathematical formalism. In the specific case of quantum mechanics the axiomatic mathematical formulation based on Hermitian operators and their corresponding eigenvectors and eigenvalues has been known ever since the invention of the theory.[1] What is actually meant is a conceptual foundation of quantum mechanics in terms of simple axioms and meaningful principles on which the entire theory can be built on.

Until now many generations of scientists have been debating on the fundamental principles of quantum mechanics but physicists still do not agree on the foundations of the theory. Thus, there currently exists a great variety of in part very different approaches. It is exactly this disagreement about the conceptual interpretation of the theory itself that characterises the present situation in quantum mechanics.

1.2 Quantum Theory from an Information-Theoretic Point of View

Due to its rather simple mathematical formalism and the fact that it predicts some very counterintuitive events that cannot be compared with any phenomena observed in our everyday life, quantum mechanics is regarded as one of the most abstract theories in

[1] The mathematical framework of quantum mechanics was mainly developed by Paul Dirac and John von Neumann between 1930 and 1932 (compare [2] and [3]).

physics. Therefore scientists are eager to find simple principles that are able to provide a physically meaningful basis for the mathematical axioms, as Hardy states in [4]:

> “A set of reasonable axioms provides us with a deeper conceptual understanding of a theory and is therefore more likely to suggest ways in which we could extend the domain of the theory or modify the axioms in the hope of going beyond quantum theory (for example, to develop quantum gravity).”[2]

Since some of the most important questions regarding current research concern the information content of quantum systems and the extent to which knowledge can be gained by performing measurements, many very promising approaches are based on information-theoretic constraints.

The idea of starting from an information-theoretic point of view first came up in the 1970s, when physicists began to realise that their search for fundamental principles in quantum mechanics has much in common with problems encountered in information theory (see [5]). One example in this respect is the question whether superluminal information transfer is possible. In 1982 Nick Herbert suggested a way in which quantum entanglement could be used to enable superluminal communication.[3] In the same year Wootters and Zurek proved the so-called *no-cloning theorem*[4] and showed that faster-than-light communication could only be performed, if cloning of an arbitrary quantum state was possible.

Due to the fact that approaches based on such constraints are very successful in deriving features which are regarded as being solely quantum mechanical, we will limit the discussion about the foundations of quantum theory to an information-theoretic point of view.

To begin with, we will look at Zeilinger’s fundamental principle [1], which states that quantum systems can only carry a finite amount of information. Starting from this constraint, Zeilinger and Brukner were able to show in [8] and [9] that some characteristic properties of quantum mechanics, such as entanglement and randomness, are natural consequences of this restriction. A quite different approach was developed by Clifton, Bub and Halvorson [10], who try to deduce quantum theory from three fundamental information-theoretic constraints. This concept is discussed in chapter 3.

The last sections deal with Hardy’s work [4], which is currently one of the most promising approaches able to derive the entire nature of quantum theory. In his paper Hardy sets out from five axioms and manages to show that quantum theory is basically very similar to classical probability theories. Referring to Hardy’s work, Spekkens’ toy theory [12] based on the so-called *knowledge balance principle* will also be presented. Even though Spekkens is not able to derive quantum theory, his toy theory suggests that quantum states are primarily states of imagination and incomplete knowledge than states of reality.

[2] L.Hardy, 2008, p. 27.

[3] compare [6]

[4] compare [7]

2 Anton Zeilinger - The Fundamental Principle

Since Anton Zeilinger's information-theoretic approach in [1] is quite intuitive and less mathematical than most of the concepts presented in the following chapters, we will start the discussion on the foundations of quantum mechanics with Zeilinger's idea.

According to Zeilinger, the description of the world, i.e. the information one possesses about physical objects, is expressed in terms of propositions in connection with their corresponding truth values. Any system, regardless of whether it is classical or quantum mechanical, can therefore be seen as consisting of a set of various propositions. In the classical case a physical object can carry an infinite number of truth values, whereas in the quantum case one has to deal with a non-commuting algebra of observables and thus cannot assign the answer to all possible propositions to a system simultaneously. One of the most prominent examples of this feature, known as quantum complementarity, are the position $\vec{x}$ and momentum $\vec{p}$ of an object which cannot be determined precisely at the same moment.

2.1 Finiteness of Information and the Fundamental Principle

In order to reproduce the characteristics of quantum complementarity, Zeilinger and Brukner consider it quite natural to suppose that the description of quantum objects in terms of propositions is subject to a limited amount of information, as they state in [8]:

"The information content of a quantum system is finite."[5]

According to this basic assumption quantum objects are related to an irreducible randomness, meaning that it is not possible to know exactly the outcome for all conceivable measurement settings.

Starting from this point of view the question arises how much information a quantum system can carry. To answer this question we have to decompose the system into its constituents until the most elementary individual system carries the truth value of a single proposition. Since the truth value is either "true" or "false", it can be expressed by one bit of information and this eventually leads to the statement which Zeilinger and Brukner refer to as their *fundamental principle* [1]:

"An elementary system carries one bit of information."[6]

This foundational assumption provides a reasonable explanation for the complementarity and randomness in quantum measurements. This can be illustrated if we consider the example of a spin-1/2 particle with spin up along the z-axis, i.e. $|\psi\rangle = |+z\rangle$ [9]. The particle is able to carry one bit of information and this corresponds to the truth value

[5] C. Brukner and A. Zeilinger, 2002, p. 2.

[6] A. Zeilinger, 1999, p. 635.

of a single proposition. Thus, it is only possible to predict the outcome of a particular measurement and all other observations contain a certain randomness. Otherwise, the particle carries more than one bit of information which contradicts the fundamental principle.

The quantum randomness signifies a crucial difference between quantum and classical mechanics. From a classical point of view the result of a measurement is predetermined, meaning that the properties of a physical object are independent of observation and are assigned to the system even before the measurement is performed. In quantum mechanics, on the contrary, the outcome is random except when the quantum state is an eigenstate of the operator describing the measurement [1].

In the next section 2.2 we are going to consider more complex objects consisting of more than one elementary system. Once again we are interested in the information content of such systems and will see that if the information is contained in the correlations of the constituents, i.e. the joint properties, this corresponds to the case of quantum entanglement.

2.2 Entanglement as Consequence of the Fundamental Principle

Since there exists no classical counterpart, entanglement is regarded as one of the most characteristic features of quantum objects that differ classical mechanics from quantum theory. In order to investigate entanglement in terms of an information-theoretic point of view, setting out from the fundamental principle, we need to know how much information is contained within N elementary systems and, more importantly, how the information is distributed over the individual constituents.
Following our basic assumptions, each of the N systems carries one bit of information and hence Zeilinger and Brukner consider it suggestive to generalize the fundamental principle [8]:

"N elementary systems carry N bits."[7]

If these systems interact with each other, the information content can be distributed within the composite system in many different ways.[8] Either the information is carried by the N elementary systems separately or it is in any way contained in the correlations between the constituents [1].

One of the simplest composite systems we can consider is a two-qubit product state, e.g. $|\psi\rangle = |{+z}\rangle\,|{+z}\rangle$. As we know from the generalization of the fundamental principle this system can carry a total number of $N = 2$ bits of information. A possible list of

[7] C. Brukner and A. Zeilinger, 2002, p. 3.

[8] At this point it is crucial to mention that we do not consider any exchange with the environment and therefore the total information of the composite system is conserved.

propositions could be two individual statements about the spin of particle 1 and particle 2, respectively, along the z-axis.
Another way of representing the system can be achieved if we replace one of these statements with a proposition concerning joint properties, like for example: “The two spins are equal along the z-axis”. If this proposition expressed our entire knowledge, i.e. the total information content, the quantum state would be incompletely described, since we are left with two possibilities:

$$|\psi\rangle = |{+z}\rangle\,|{+z}\rangle \quad or \quad |\psi\rangle = |{-z}\rangle\,|{-z}\rangle\,. \tag{2.1}$$

As we can see, the correlations can be represented only by a single proposition and thus the second bit of information must be carried by one of the systems individually. The crucial point here is that a two-qubit product state is characterized by at most one proposition which describes the joint properties. The information content of the correlations is therefore[9] [9]:

$$I_{\text{corr}}^{\text{product}} = 1. \tag{2.2}$$

Instead of a pure product state we will now consider one of the four maximally entangled Bell states:

$$|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}\left(|{+z}\rangle\,|{-z}\rangle \pm |{-z}\rangle\,|{+z}\rangle\right) \tag{2.3}$$

$$|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}\left(|{+z}\rangle\,|{+z}\rangle \pm |{-z}\rangle\,|{-z}\rangle\right). \tag{2.4}$$

In all of these cases the information content represented by joint properties of the individual systems is not restricted to one bit as opposed to product states. This is another consequence of the finiteness of information and the fundamental principle as we will show in the following example.
If we take one of the Bell states, e.g. $|\phi^{-}\rangle$, we see that this quantum state can be described in different bases [1] [9]:

$$\begin{aligned} |\phi^{-}\rangle &= \frac{1}{\sqrt{2}}\left(|{+z}\rangle\,|{+z}\rangle - |{-z}\rangle\,|{-z}\rangle\right) && (2.5)\\ &= \frac{1}{\sqrt{2}}\left(|{+x}\rangle\,|{-x}\rangle + |{-x}\rangle\,|{+x}\rangle\right) && (2.6)\\ &= \frac{1}{\sqrt{2}}\left(|{+y}\rangle\,|{+y}\rangle + |{-y}\rangle\,|{-y}\rangle\right). && (2.7) \end{aligned}$$

Thus, we can state two different propositions representing joint properties, meaning that the total number of two bits of information is completely contained in the correlations of the individual systems [9]:

$$I_{\text{corr}}^{\text{entangled}} = 2. \tag{2.8}$$

[9] An adequate measure of information will be introduced in the next section 2.3

A possible list of propositions comprises the statements “The spins along the z-axis are equal” and “The spins are not equal along the x-axis”. Due to the generalized fundamental principle, the Bell states can carry only two bits of information and consequently the third proposition concerning the spins of the particles along the y-axis must already be determined by the other two statements.[10]

The examples discussed above suggest that quantum entanglement is just another natural implication of the finiteness of information and provide an information-theoretic entanglement criterion for two-qubit systems [9]:

$$I_{\text{corr}}^{\text{entangled}} > 1. \tag{2.9}$$

In other words, in the case of a two-qubit entangled state the two available bits are required to represent joint properties of the composite system and the individual particles do not carry any information on their own.
Zeilinger and Brukner’s entanglement criterion is deeply connected with Erwin Schrödinger’s paper *“The present situation in quantum mechanics”* from 1935 (compare [13]). In this work, Schrödinger first came up with the idea that information is bound and can primarily reside in correlations between the individual systems [13]:

> “Maximal knowledge of a total system does not necessarily include total knowledge of all its parts, not even when these are fully separated from each other and at the moment are not influencing each other at all.”[11]

The general concept of the entanglement criterion does not only apply to two-qubit systems but is also valid for more complex objects (see Zeilinger, Brukner and Zukowski, 2001). In this paper it is also shown that a general Bell inequality can be derived which corresponds to our entanglement criterion and therefore agrees with the fundamental principle.

2.3 Measure of Information

So far we talked a lot about the information content of quantum systems and discussed some consequences of the fundamental principle but we did not specify an appropriate measure of information or, respectively, the information gain in quantum experiments.

For that purpose we first consider a probabilistic experiment with only two possible outcomes ”yes” and ”no” and associated probabilities p_1 and p_2 [8]. Such experiments obey the binomial distribution with standard deviation

$$\sigma = \sqrt{p_1 p_2 N} = \sqrt{p_1(1-p_1)N}, \tag{2.10}$$

[10] This is obviously true, since the relation $\sigma_y^1\sigma_y^2 = -\left(\sigma_x^1\sigma_x^2\right)\left(\sigma_z^1\sigma_z^2\right)$ implies that an eigenstate of $\sigma_y^1\sigma_y^2$ is a joint eigenstate of $\sigma_x^1\sigma_x^2$ and of $\sigma_z^1\sigma_z^2$ [9].

[11] E. Schrödinger, translated by J.Trimmer, 1980, p.331.

where N denotes the number of experimental trials. Since one cannot exactly predict the sequence of these N outcomes, the result of each trial is subject to an element of uncertainty [8]. According to Zeilinger and Brukner it is reasonable to assume that the experimenter's uncertainty U is proportional to σ^2. Thus, after each trial one's uncertainty is reduced by [8]:

$$U = \frac{\sigma^2}{N} = p_1(1 - p_1). \tag{2.11}$$

If we now generalize our example and consider an experiment with n possible outcomes and probabilities p_1, p_2, ..., p_n, the uncertainty can be expressed in the following way [8]:

$$U(\vec{p}) = \sum_{i=1}^{n} U(p_i) = \sum_{i=1}^{n} p_i(1 - p_i) = 1 - \sum_{i=1}^{n} p_i^2. \tag{2.12}$$

This implies that $U(\vec{p})$ ($\vec{p}$ defines the probability vector $\vec{p} = (p_1, p_2, ..., p_n)$) represents the total lack of information about all conceivable measurement results.
In order to obtain a proper expression for the measure of information $I(\vec{p})$, Zeilinger and Brukner argue that the information gain in a single trial is the complement of the uncertainty $U(\vec{p})$ [8]:

$$I(\vec{p}) = 1 - U(\vec{p}) = \sum_{i=1}^{n} p_i^2. \tag{2.13}$$

As we can see, if one of the probabilities is equal to one, $I(\vec{p})$ reaches its maximum whereas in the case of equal probabilities the uncertainty is maximal. In the second situation $p_i = 1/n$ ($i = 1, 2, ..., n$) and the experimenter is not able to gain any information regarding subsequent measurement outcomes. Hence, it is useful to introduce an adequate normalization [8]:

$$I(\vec{p}) = \mathcal{N} \sum_{i=1}^{n} \left(p_i - \frac{1}{n} \right)^2. \tag{2.14}$$

$\mathcal{N}$ denotes the normalization constant which, in the case of k bits of information and consequently $n = 2^k$ possible measurement outcomes, is given by $\mathcal{N} = 2^k k/(2^k - 1)$.[12]
In particular, in an experiment with only two outcomes it is easy to prove that the measure of information is:

$$I(\vec{p}) = 2\left(p_1 - \frac{1}{2}\right)^2 + 2\left(p_2 - \frac{1}{2}\right)^2 = (p_1 - p_2)^2. \tag{2.15}$$

Equation (2.15) specifies the appropriate expression for the measure of information needed to describe experiments with two-dimensional quantum system, e.g. spin-1/2

[12] The derivation of this result is shown in Brukner and Zeilinger, 1999 [14].

particles, as discussed in section 2.2.

Let us recall the Bell state $|\phi^-\rangle$ and the two propositions concerning the spins along the x- and the z-axis. In the case of the first statement "The spins are not equal along the x-axis" we can define the probabilities p^+_{xx} and p^-_{xx}. p^+_{xx} quantifies the probability that the spins of both particles are equal along the x-axis and therefore p^-_{xx} describes the probability that the spins are not equal. According to this definition, we know that $p^+_{\mathrm{xx}} = 0$ and, consequently, $p^-_{\mathrm{xx}} = 1$. Thus, the measure of information I_{xx} regarding a measurement in x-direction is given by:

$$I_{\mathrm{xx}} = \left(p^+_{\mathrm{xx}} - p^-_{\mathrm{xx}}\right)^2 = 1. \tag{2.16}$$

The measure of information I_{zz} with respect to the z-axis is determined analogously:

$$I_{\mathrm{zz}} = \left(p^+_{\mathrm{zz}} - p^-_{\mathrm{zz}}\right)^2 = 1. \tag{2.17}$$

We see that in the case of entangled states the total information resides in the correlations between the particles as pointed out in section 2.2.

Therefore, the measure of information we deduced here for quantum experiments agrees with the predictions of the fundamental principle and this fact indicates that there exists a certain link between quantum mechanics and a limited amount of information.

3 Clifton, Bub and Halvorson - Quantum Theory in Terms of Three Information-Theoretic Constraints

In this section we are going to discuss another interesting approach, developed by Clifton, Bub and Halvorson (see [10]), that is quite different from Zeilinger's fundamental principle but also manages to derive some of the most significant traits of quantum mechanics in a mathematically abstract way. Starting point here is the assumption that the physical world is subject to the following three fundamental information-theoretic constraints [10]:

- Superluminal information transfer between separated physical systems is not possible.
- Information contained in unknown physical states cannot be broadcasted perfectly.
- Unconditionally secure bit commitment is not possible.

In order to derive quantum theory, these fundamental constraints have to be connected with some of the very characteristics of quantum systems [10]:

- The algebras of observables related to distinct physical system need to commute. This property is referred to as kinematic independence.
- The algebra of observables of an individual system has to be noncommutative.

- The physical world needs to be nonlocal, meaning that separated systems are at least sometimes represented by entangled states.

To prove that the information-theoretic constraints suffice to deduce the features of quantum mechanics listed above, we will use the concept of C^*-algebras. Since C^*-algebras are algebras of linear operators acting on a Hilbert space, they are an important tool for the description of the mathematical formalism of quantum mechanics. Thus, in the next section we will start with an introduction to the basics of our mathematical framework.

3.1 Introduction to C^*-Algebras

In the most general case a C^*-algebra is defined as a Banach *-algebra[13] over the field of complex numbers $\mathbb{C}$. In this context, the *-algebra (also referred to as involutive algebra) represents in a certain way the "adjoint" operation that satisfies the following properties for all elements A and B of the corresponding Banach space and complex numbers α and β [15]:

$$\begin{aligned}
&(A^*)^* = A \\
&(AB)^* = B^* A^* \\
&(\alpha A + \beta B)^* = \alpha^* A^* + \beta^* B^* \\
&\|A^* A\| = \|A\|^2.
\end{aligned}$$

The last property, $\|A^*A\| = \|A\|^2$, characterizes C^*-algebras and the $\|\cdot\|$-operation denotes the standard operator norm. Any algebra containing the identity operator is called a unital C^*-algebra. A very important example, the one that will be used throughout this chapter, is the algebra $\mathfrak{B}(\mathscr{H})$ of all bounded operators on a Hilbert space $\mathscr{H}$.

In connection with algebras, we also need to define the concept of representations. A representation of a C^*-algebra $\mathfrak{A}$ is a mapping π, symbolized by $\pi : \mathfrak{A} \to \mathfrak{B}(\mathscr{H})$, preserving the linear and *-structure of $\mathfrak{A}$ [10]. A representation is said to be irreducible, if $\mathscr{H}$ and the null space are the only closed subspaces of $\mathscr{H}$ that are invariant under the mapping π.

An important subalgebra of $\mathfrak{B}(\mathscr{H})$ are von Neumann algebras $\mathscr{R}$. These algebras are characterized by the property $\mathscr{R} = \mathscr{R}''$, where $\mathscr{R}''$ denotes the double commutant of $\mathscr{R}$.[14] The center of a von Neumann algebra is defined by $\mathscr{R} \cap \mathscr{R}'$ and referred to as factor, if it includes only multiples of the identity operator [10].
A quantum mechanical state is any linear functional $\rho : \mathfrak{A} \to \mathbb{C}$ on a C^*-algebra $\mathfrak{A}$ that

[13] An detailed introduction to Banach algebras and spaces goes beyond the scope of this work. We limit our discussion to the very basics of C^*-algebras.

[14] The commutant $\mathscr{R}$' describes the set of all operators on the Hilbert space $\mathscr{H}$ that commute with the operators of $\mathscr{R}$.

is positive and normalized, i.e. [15]:

$$\begin{aligned} &\rho(\alpha A+\beta B)=\alpha\rho(A)+\beta\rho(B), && \forall A,B\in\mathfrak{A} \text{ and } \forall\alpha,\beta\in\mathbb{C}\\ &\rho(A^*A)\geq 0, && \forall A\in\mathfrak{A}\\ &\rho(\mathbb{1})=1. && \end{aligned}$$

The standard form for the description of a quantum state defined on the algebra $\mathfrak{B}(\mathscr{H})$ can be achieved by using a trace-one density operator D [10]:

$$\rho(A)=\mathrm{Tr}(AD), \quad \forall A\in\mathfrak{B}(\mathscr{H}).$$

In the case of a pure state, the density operator satisfies $D^2=D$, meaning that the only possible eigenvalues of D are 0 and 1.

At this point we have to introduce a theorem that is very significant for the abstract characterization of C^*-algebras. The Gelfand-Naimark-Segal theorem (GNS theorem) claims that for all states ρ there exists a representation $(\pi_\rho,\mathscr{H}_\rho)$ of $\mathfrak{A}$ and a vector Ω_ρ defined on the Hilbert space $\mathscr{H}_\rho$ in such a way that [10]:

- $\rho(A)=<\Omega_\rho|\pi_\rho(A)\Omega_\rho> \ \ \forall A\in\mathfrak{A}$, and
- the set $\{\pi_\rho(A)\Omega_\rho : A\in\mathfrak{A}\}$ is dense in $\mathscr{H}_\rho$.

Just in the case of pure states ρ, the triple $(\pi_\rho,\mathscr{H}_\rho,\Omega_\rho)$, the so called GNS representation of $\mathfrak{A}$, defines an irreducible representation.
Referring to the GNS theorem, the Gelfand-Naimark theorem states that for an abstract C^*-algebra one can obtain a faithful[15] representation as a norm-closed *-subalgebra of $\mathfrak{B}(\mathscr{H})$, for a Hilbert space $\mathscr{H}$ [10]. According to this theorem, C^*-algebras can be regarded as general algebras of operators acting on Hilbert spaces.

In the following section 3.2 we will discuss how classical as well as quantum theories fit into the theory of C^*-algebras we established here.

3.2 Classical and Quantum Theories in the C^*-algebraic framework

In order to investigate classical theories in terms of our mathematical framework we need to introduce another important theorem concerning the so called function representation of a C^*-algebra $\mathfrak{A}$ [10]. This kind of representation implies that an abelian C^*-algebra is isomorphic to the set $C(X)$ of all complex-valued, continuous functions defined on a compact Hausdorff space X.
In this case each element A of $\mathfrak{A}$ is mapped to a function $\hat{A}$, referred to as Gelfand transformation of A, whose value at a particular state ρ of the pure state space $\mathscr{P}(\mathfrak{A})$ corresponds to the value that ρ assigns to A [10]. Consequently, we see that every phase

[15] A representation π of $\mathfrak{A}$ is called faithful, if $\pi(A)=0$ implies $A=0$ $(A\in\mathfrak{A})$, see [10].

space representation of a classical theory agrees with a C^*-algebra and, furthermore, we can also conclude that every abstract abelian C^*-algebra represents a classical phase space theory [10].

A similar condition can be derived for quantum theories, but in contrast to the classical case, such theories are represented by non-abelian C^*-algebras. This property, however, does not suffice to deduce the entire nature of quantum mechanics as Clifton, Bub and Halvorson suggest in [10]. According to these physicists, the most characteristic feature that fundamentally sets quantum mechanics apart from its classical counterpart is the existence of entangled states.
The significance of entanglement for quantum systems was first described by Erwin Schrödinger in his famous paper from 1935 (see [13]). Schrödinger did not believe that nonlocal entangled states can indeed exist, since he considered physical systems, no matter whether they are classical or quantum mechanical, to be characterized by local states only. But, in contrast to Schrödinger's notion of quantum mechanics, the experimental violation of Bell's inequality has actually proved the existence of nonlocal entanglement.

In the subsequent sections we are going to show that our three information-theoretic constraints suffice to deduce the characteristic features of quantum mechanics stated at the beginning of chapter 3. Clifton, Bub and Halvorson even prove the converse derivation in [10], but that goes beyond the scope of this work and, therefore, we limit our discussion to the deduction of the fundamentals of quantum theory starting from an information-theoretic point of view.

3.3 Kinematic Independence

First, we set out from the point that superluminal information transfer is not possible and demonstrate that this single restriction already implies the kinematic independence of distinct physical systems [10].
For this purpose let us consider a composite system, consisting of two subsystems $\mathcal{A}$ and $\mathcal{B}$, where local measurements in $\mathcal{A}$ do not have any impact on $\mathcal{B}$. Due to this constraint we assume all measurements in $\mathcal{A}$ to be represented by nonselective (measurement) operators T [10]:

$$T(A) = \sum_{i=1}^{n} E_i^{\frac{1}{2}} A E_i^{\frac{1}{2}}, \quad A \in \mathfrak{A} \vee \mathfrak{B} \tag{3.1}$$

where $\mathfrak{A} \vee \mathfrak{B}$ denotes the C^*-algebra of the composite system $\mathcal{A} + \mathcal{B}$ and E_i describe positive operators in $\mathfrak{A}$ with the defining property $\sum_{i=1}^{n} E_i = 1$. The restriction to nonselective operators T endorses the following definition regarding the information content of system $\mathcal{B}$ after a measurement in system $\mathcal{A}$ has been performed [10]:

> **Definition 1:** *"An operation T conveys no information to $\mathcal{B}$ just in case $(T^*\rho)|_{\mathfrak{B}} = \rho|_{\mathfrak{B}}$ for all states ρ of $\mathfrak{B}$."* (Halvorson, 2003, p. 1574)

At this point it is important to mention that we will consider only such C^*-algebras, where $\rho(A) = \rho(B)$ implies $A = B$ for all states ρ [10]. Hence, $(T^*\rho)|_{\mathfrak{B}} = \rho|_{\mathfrak{B}}$ if and only if $\rho(T(B)) = \rho(B)$ for all $B \in \mathfrak{B}$ and states ρ of $\mathfrak{A} \vee \mathfrak{B}$. According to our information-theoretic constraint all states of $\mathfrak{B}$ are also states of $\mathfrak{A} \vee \mathfrak{B}$ and therefore $(T^*\rho)|_{\mathfrak{B}} = \rho|_{\mathfrak{B}}$ if only if $T(B) = B$ for all $B \in \mathfrak{B}$ [10].

In order to show the kinematic independence of $\mathfrak{A}$ and $\mathfrak{B}$ one has to prove that local measurements in $\mathcal{A}$ do not suffice to convey any information to system $\mathcal{B}$. Clifton, Bub and Halvorson deem it reasonable to consider nonselective measurement operators T of the following form [10]:

$$T(A) = PAP + (\mathbb{1} - P)A(\mathbb{1} - P), \quad A \in \mathfrak{A} \vee \mathfrak{B} \tag{3.2}$$

with P representing a projection in $\mathfrak{A}$. We already showed that for any $B \in \mathfrak{B}$ [10]

$$\begin{aligned} B = T(B) &= PBP + (\mathbb{1} - P)B(\mathbb{1} - P) && (3.3) \\ \Rightarrow 2PBP - PB - BP &= 0. && (3.4) \end{aligned}$$

If we now multiply equation (3.4) on the left and, respectively, on the right with P, we obtain [10]:

$$\begin{aligned} PBP - PB &= 0 && (3.5) \\ PBP - BP &= 0. && (3.6) \end{aligned}$$

Note that here we used the characteristic property of projections, $P^2 = P$. Subtracting these two equations yields [10]:

$$[P, B] = 0. \tag{3.7}$$

We know that $\mathfrak{A}$ is spanned by all its projection operators P, whereas $\mathfrak{B}$ is spanned by its self-adjoint operators and therefore we can conclude that $\mathfrak{A}$ and $\mathfrak{B}$ are kinematically independent [10].

3.4 Noncommutativity

In this section we are going to show that the *"No Broadcasting"*-constraint entails the noncommutativity concerning the algebra of observables of a distinct quantum system [10]. Starting point here is a general theorem, which states that the broadcasting of an arbitrary pair of pure states can be performed if and only if the corresponding density matrices are mutually commuting.[16] Hence, we assume classical mechanics to be characterized by the property that all states can be broadcast, whereas in the quantum case this is generally not possible.

In the following, we will need a number of definitions and theorems, which will not be entirely proved in this work.[17] First of all, we begin with a very important lemma [10]:

[16] This theorem was proved by Barnum et al., see [11].

[17] All complete proofs are given in [10].

Lemma 1: Let us consider two physical systems $\mathcal{A}$ and $\mathcal{B}$ with associated C^*-algebras $\mathfrak{A}$ and $\mathfrak{B}$. If $\mathfrak{A}$ and $\mathfrak{B}$ are kinematically independent, there is at most one state σ of $\mathfrak{A} \vee \mathfrak{B}$ for any state ω of $\mathfrak{A}$ and any state ρ of $\mathfrak{B}$ that satisfies $\sigma(AB) = \omega(A)\rho(B)$ for all $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$.

In addition to this lemma we give the definition [10]:

Definition 2: *"Suppose two kinematically independent, isomorphic C^*-algebras $\mathfrak{A}$ and $\mathfrak{B}$. A pair of states $\{\rho_0, \rho_1\}$ of $\mathfrak{A}$ can be broadcast only if there exists a standard state σ of $\mathfrak{B}$ and a dynamical evolution represented by an operation T of $\mathfrak{A} \vee \mathfrak{B}$such that $T^*(\rho_i \otimes \sigma)|_{\mathfrak{A}} = T^*(\rho_i \otimes \sigma)|_{\mathfrak{B}} = \rho_i \;\; (i = 0, 1)$. The pair of states $\{\rho_0, \rho_1\}$ of $\mathfrak{A}$ can be cloned just in the case that $T^*(\rho_i \otimes \sigma) = \rho_i \otimes \rho_i \;\; (i = 0, 1)$".* (Halvorson, 2003, p. 1578)

The subsequent theorem proves that in the case of Abelian algebras of observables any pair of pure states can be cloned and broadcast by a certain broadcasting map [10]:

Theorem 1: For two isomorphic Abelian C^*-algebras $\mathfrak{A}$ and $\mathfrak{B}$ there exists an operation T on $\mathfrak{A} \vee \mathfrak{B}$ that is able to broadcast all states of $\mathfrak{A}$.

For the proof of this theorem we use the fact that due to the isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ both Abelian algebras are also isomorphic to the set $C(X)$ defined on a compact Hausdorff space X[18] and, hence, $\mathfrak{A} \otimes \mathfrak{B} \cong C(X) \otimes C(X) \cong C(X \times X)$[19] [10].

We can now define a linear mapping T on $C(X \times X)$ that satisfies $T(\mathbb{1}) = \mathbb{1}$, i.e. T represents a nonselective operation on $C(X) \otimes C(X)$. The operation T has the following properties for any functions f and g on some Hausdorff space X and for any states ρ, σ of $C(X)$ [10]:

$$T(f \otimes g) = fg \otimes \mathbb{1} \tag{3.8}$$

$$T(\mathbb{1} \otimes f) = f \otimes \mathbb{1} \tag{3.9}$$

$$\Rightarrow T^*(\rho \otimes \sigma)(\mathbb{1} \otimes f) = (\rho \otimes \sigma)(f \otimes \mathbb{1}) = \rho(f). \tag{3.10}$$

As we can see, $T^*(\rho \otimes \sigma)|_{\mathfrak{B}} = \rho$. Since $T(f \otimes \mathbb{1}) = f \otimes \mathbb{1}$, we can conclude that [10]:

$$T^*(\rho \otimes \sigma)(f \otimes \mathbb{1}) = (\rho \otimes \sigma)(f \otimes \mathbb{1}) = \rho(f) \tag{3.11}$$

$$\Rightarrow T^*(\rho \otimes \sigma)|_{\mathfrak{A}} = \rho. \tag{3.12}$$

Note that the nonselective operator T introduced above represents a many-to-one mapping of the underlying pure state space and therefore T is irreversible.

According to the previous theorem, all states in classical systems can be broadcast,

[18] See section 3.2.

[19] Note that $\mathfrak{A} \vee \mathfrak{B}$ is isomorphic to $\mathfrak{A} \otimes \mathfrak{B}$, see [10].

since such theories are characterized by Abelian algebras of observables. To prove that this property significantly differs classical from quantum mechanics, we need to show that if any pair of states can be broadcast, the corresponding system has an Abelian algebra of observables [10]. For this purpose, Clifton, Bub and Halvorson suggest to demonstrate that if the broadcasting of an arbitrary pair of pure states is possible, then these states can be cloned.

We will proceed with our proof by introducing the concept of unitary operators in the C^*-algebraic framework. Consider a state ρ of $\mathfrak{A}$ and let U be a unitary operator in $\mathfrak{A}$. We can define another state ρ_U[20] by [10]:

$$\rho_U(A) = \rho(U^*AU) \quad (A \in \mathfrak{A}). \tag{3.13}$$

Additionally, if U is a unitary operator in $\mathfrak{A}$ and V is another unitary operator in $\mathfrak{B}$, then for all $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ we obtain [10]:

$$(\omega \otimes \rho)_{U\otimes V}(A \otimes B) = (\omega \otimes \rho)(U^*AU \otimes V^*BV) = (\omega_U \otimes \rho_V)(A \otimes B). \tag{3.14}$$

The uniqueness of product states then implies that $(\omega \otimes \rho)_{U\otimes V} = \omega_U \otimes \rho_V$.

In order to prove that the possibility to broadcast arbitrary pairs of pure states entails that the underlying algebra of observables is Abelian, we will need a number of different lemmas[21] and theorems [10].

> **Lemma 2:** Consider two unitarily equivalent pure states ρ_0 and ρ_1 of $\mathfrak{A}$ and another state σ of $\mathfrak{B}$, then [10]
>
> $$p(\rho_0 \otimes \sigma, \rho_1 \otimes \sigma) = p(\rho_0, \rho_1) \tag{3.15}$$
> $$p(\rho_0 \otimes \rho_0, \rho_1 \otimes \rho_1) = p(\rho_0, \rho_1)^2, \tag{3.16}$$
>
> where $p(\rho_0, \rho_1)$ denotes the transition probability between ρ_0 and ρ_1.
>
> **Lemma 3:** Let $\mathfrak{A}$ and $\mathfrak{B}$ be kinematically independent C^*-algebras. If we define a state ρ of $\mathfrak{A} \vee \mathfrak{B}$ in such a way that either $\rho|_{\mathfrak{A}}$ or $\rho|_{\mathfrak{B}}$ is pure, then ρ is a product state.
>
> **Theorem 2:** If there exists a (nonselective) operation T on $\mathfrak{A} \vee \mathfrak{B}$ that broadcasts each pair $\{\rho_0, \rho_1\}$ of $\mathfrak{A}$, the algebra $\mathfrak{A}$ is Abelian.

For the proof of the previous theorem let us consider a nonabelian C^*-algebra $\mathfrak{A}$. If there indeed exists an operation T that is able to broadcast each pair of states $\{\rho_0, \rho_1\}$, then there are certain pure states ρ_0 and ρ_1 of $\mathfrak{A}$ that satisfy $0 < ||\rho_0 - \rho_1|| < 2$[22].

[20] ρ and ρ_U are called unitarily equivalent.

[21] The proofs of the following two lemmas will not be given in this work. For further details see [10].

[22] Here we use the fact that two pure states ρ_0 and ρ_1 are orthogonal, if $||\rho_0 - \rho_1|| = 2$, compare [10].

We now assume that a state σ of $\mathfrak{B}$ is given, such that we obtain [12]:

$$T^*(\rho_0 \otimes \sigma)|_{\mathfrak{A}} = T^*(\rho_0 \otimes \sigma)|\mathfrak{B} = \rho_0, \tag{3.17}$$

$$T^*(\rho_1 \otimes \sigma)|_{\mathfrak{A}} = T^*(\rho_1 \otimes \sigma)|\mathfrak{B} = \rho_1. \tag{3.18}$$

According to *Lemma 3* we can conclude that $T^*(\rho_0 \otimes \sigma) = \rho_0 \otimes \rho_0$, as well as $T^*(\rho_1 \otimes \sigma) = \rho_1 \otimes \rho_1$. For the transition probability p it follows that [10]:

$$\begin{aligned} p(\rho_0, \rho_1) &= p(\rho_0 \otimes \sigma, \rho_1 \otimes \sigma) && (3.19)\\ &\leq p(T^*(\rho_0 \otimes \sigma), T^*(\rho_1 \otimes \sigma)) && (3.20)\\ &= p(\rho_0 \otimes \rho_0, \rho_1 \otimes \rho_1) && (3.21)\\ &= p(\rho_0, \rho_1)^2. && (3.22) \end{aligned}$$

Note that the nonselective operation T cannot decrease the transition probability. Since the inequality we deduced above, $p(\rho_0, \rho_1) \leq p(\rho_0, \rho_1)^2$, contradicts our requirement that $0 < p(\rho_0, \rho_1) < 1$, we can conclude that $\mathfrak{A}$ has to be Abelian [10].

As we can see, a pair of states of a system can be broadcast, if and only if, the corresponding algebra of observables is Abelian, i.e. only classical theories can broadcast arbitrary states. But since we already know that in this case the density matrices are commuting (see Barnum et al., [11]), we can conclude that individual quantum systems are characterized by a noncommutative algebra of observables.

3.5 Nonlocality

Finally, we want to demonstrate that our last remaining information-theoretic constraint, the *"No Bit Commitment"*-theorem, implies that separated physical systems are at least sometimes represented by entangled states[23] [10].
We start with a lemma concerning the non-uniqueness of the decomposition of quantum mechanical mixed states[24] [10]:

> **Lemma 4:** Consider a C^*-algebra $\mathfrak{A}$. $\mathfrak{A}$ is nonabelian if and only if there exist specific pure states $\omega_{1,2}$ and $\omega_\pm$ of $\mathfrak{A}$ that satisfy the relation:
>
> $$\frac{1}{2}\left(\omega_1 + \omega_2\right) = \frac{1}{2}\left(\omega_+ + \omega_-\right). \tag{3.23}$$

For the subsequent proof we will have a close look at a bit commitment protocol between two distinct physical systems A and B. In the case that the mixed states in both systems

[23] This property is generally referred to as nonlocality, compare section 3.

[24] The proof of the lemma will not be given in this work. For further details see [10].

can be decomposed in a non-unique way, we can define a pair of correlated states $\{\rho_0, \rho_1\}$ in $A + B$, which has identical marginals relative to A and B [10]:

$$\rho_0 = \frac{1}{2}\left(\omega_1 \otimes \omega_1 + \omega_2 \otimes \omega_2\right) \tag{3.24}$$

$$\rho_1 = \frac{1}{2}\left(\omega_+ \otimes \omega_+ + \omega_- \otimes \omega_-\right). \tag{3.25}$$

Provided that the algebras of observables pertaining to A and B do not commute but are kinematically independent, let us consider a party, called Alice, that wants to supply an encoded bit to another party, called Bob [10].
A commitment scheme usually consists of two phases, the *commitment stage* and the *revelation stage*. In the commitment stage we suppose that Alice is able to decide whether the state ρ_0 or ρ_1 is prepared [10]. Her choice either relates to the commitment 0 or to the commitment 1, respectively. If Alice committed to 0, Bob is commanded to do a measurement that discriminates between ω_1 and ω_2 in the revelation stage. If, on the other hand, Alice committed to 1, Bob is instructed to perform a measurement able to distinguish between ω_+ and ω_-.

We proceed with an important theorem[25] that proves the security of the bit commitment protocol, if Alice and Bob are able to make use of only classically correlated states[26] [10]. In addition, we prove that due to the impossibility of superluminal information transfer Alice can convert product states only to other product states. Thus, it is ensured that Alice cannot cheat by preparing an arbitrary state σ and transforming it later, at some other stage, into ρ_0 and ρ_1 [10].

> **Theorem 3:** Let $\mathfrak{A}$ and $\mathfrak{B}$ be two nonabelian C^*-algebras. There exists a pair of states $\{\rho_0, \rho_1\}$ of $\mathfrak{A} \vee \mathfrak{B}$ that satisfies:
>
> - $\rho_0|_{\mathfrak{B}} = \rho_1|_{\mathfrak{B}}$.
> - There exists no classically correlated state σ of $\mathfrak{A}$ and $\mathfrak{B}$ and nonselective measurement operations T_0 and T_1 that fulfill:
>
> $$T_0^* \sigma = \rho_0 \tag{3.26}$$
>
> $$T_1^* \sigma = \rho_1. \tag{3.27}$$

We can conclude that if the composite system $A + B$ has a pair of classically correlated states $\{\rho_0, \rho_1\}$ with identical marginals relative to A and B, then A and B can have an entangled state that can be converted to ρ_0 or ρ_1 by a local operation [10]. Hence, the impossibility of unconditionally secure bit commitment implies the nonlocality of physical systems.

As we can see, the three information-theoretic constraints stated at the beginning of

[25] Once again, the proof of the lemma will not be given in this work. For further details see [10].

[26] Classically correlated states relate to convex combinations of product states [10].

section 3 indeed suffice to deduce some of the very characteristic features of quantum mechanics. This indicates that the foundations of quantum physics are at least to some extent subject to a small number of fundamental information-theoretic axioms.

4 Lucien Hardy - Quantum Theory from five Reasonable Axioms

Another promising approach able to derive many characteristics of quantum mechanics is represented by Hardy's work [4]. In contrast to the rather abstract standard formulation of quantum theory based on Hermitian operators acting on complex Hilbert spaces, Hardy proposes a set of five intuitive axioms starting from an information-theoretic point of view. As we will see in this chapter, these axioms have much in common with classical probability theories and it turns out that in fact only a single axiom distinguishes quantum theory from its classical counterparts.
We will start our discussion with an introduction to Hardy's reasonable axioms [4]:

> **Axiom 1, *Probabilities*:** *"Relative frequencies (measured by taking the proportion of times a particular outcome is observed) tend to the same value (which we call the probability) for any case where a given measurement is performed on an ensemble of n systems prepared by some given preparation in the limit as n becomes infinite."* (L. Hardy, 2008, p. 2)

This first axiom is fundamental to all probability theories, in that the measurement probabilities depend only on the preparation and not on the particular ensemble being used. Thus, we can associate with each preparation a distinct state of the system which contains all the information required to calculate the probability of a specific measurement outcome [4].

Before we can proceed with the remaining axioms, we need to define two integers that will be very important for the description of the physical systems being investigated [4]:

- The integer K represents the number of degrees of freedom. In other words, K denotes the minimum number of measurements necessary to specify the state.
- The integer N represents the dimension of the system, i.e. N defines the maximum number of states distinguishable in a single measurement.

> **Axiom 2, *Simplicity*:** *"K is determined by a function of N (i.e. $K = K(N)$) where $N = 1, 2, ...$ and where, for each given N, K takes the minimum value consistent with the axioms."* (L. Hardy, 2008, p. 2)

The second axiom implies that there exists no natural restriction on the number of dimensions a physical system can have, but choosing the smallest possible number of

K for a given N leads to the simplest feasible theory in agreement with the axioms [4]. As we will see in the following sections, $K = N^r$ with r being a positive integer. In particular, it will be proved that $K = N$ in the case of classical probability theories, whereas in the case of quantum theory we have $K = N^2$.

> **Axiom 3, *Subspaces*:** *"A system whose state is constrained to belong to an M dimensional subspace (i.e. have support on only M of a set of N possible distinguishable states) behaves like a system of dimension M."*
> (L. Hardy, 2008, p. 2)

The *subspaces*-axiom emphasizes the notion that all ensembles with the same number of distinguishable states carry an equal amount of information [4].

> **Axiom 4, *Composite systems*:** *"A composite system consisting of two subsystems A and B having dimension N_A and N_B respectively, and number of degrees of freedom K_A and K_B respectively, has dimension $N = N_AN_B$ and number of degrees of freedom $K = K_AK_B$."* (L. Hardy, 2008, p. 2)

Hardy considers it reasonable to assume that the composite system $A + B$ comprises at least N_AN_B distinguishable states[27] [4]. Provided that $N = N_AN_B$, Hardy manages to show in his work that the number of degrees of freedom is given by $K = K_AK_B$[28].

> **Axiom 5, *Continuity*:** *"There exists a continuous reversible transformation on a system between any two pure states of the system."*
> (L. Hardy, 2008, p. 2)

The last axiom 5 is the one that significantly differs quantum theory from classical probability theories [4]. It is quite astonishing, but one even just has to remove the word *"continuous"* from the previous axiom to recover classical theories. To see this, we will give a description of the basic concepts of classical as well as quantum probability theories in the following sections.

4.1 The Structure of Classical Probability Theories

First, we will have a look at classical probability theories. Let us consider a classical system that can be located in one of N distinguishable states, the so-called basis states, which span the corresponding state space. It is now possible to assign each of the N basis states a probability p_n $(n = 1, 2, ..., N)$ of finding the system in a specific state after a measurement has been performed [4]. We can therefore represent the state of

[27] Basically, the composite system can have a larger number of distinguishable states but we will not consider this case here, compare [4].

[28] The proof will not be performed in this work. For further details see [4].

the system by a single vector whose entries are related to the distinct measurement probabilities [4]:

$$\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{pmatrix}. \tag{4.1}$$

This vector is completely characterized by a total number of N measurement probabilities and, thus, classical probability theories satisfy $K = N$[29]. Since a general state of the state space is given as a convex sum of the basis states, the basis states can be written in the following form [4]:

$$\mathbf{p}_{\text{null}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathbf{p}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathbf{p}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \cdots, \ \mathbf{p}_{\text{N}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{4.2}$$

The null state $\mathbf{p}_{\text{null}}$ corresponds to the case in which the system does not enter the measurement device.

In the following we will associate the basis states, except the null state, with the pure states of the classical system and refer to the minimum number of K probability measurements needed to specify the state as *fiducial* measurements [4]. In order to prove whether the system $\mathbf{p}$ is in one of the pure states, we relate all these measurements to the vectors $\mathbf{r}_n$ $(n = 1, 2, ..., K = N)$ [4]:

$$\mathbf{r}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathbf{r}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \cdots, \ \mathbf{r}_K = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{4.3}$$

Hence, in the case of pure states the measurement probability p_{meas} is given by the dot product [4]:

$$p_{\text{meas}} = \mathbf{r}_n \cdot \mathbf{p}, \qquad n = 1, 2, ..., K. \tag{4.4}$$

This simple formula does not only apply to pure state measurements but is also valid for general types of measurements, as we can see, if we consider the situation where we decide to measure an arbitrary $\mathbf{r}_n$ with probability λ and $\mathbf{r}_k$ with probability $1 - \lambda$. The corresponding measurement is thus represented by $\mathbf{r} = \lambda \mathbf{r}_n + (1 - \lambda)\mathbf{r}_k$ and from

[29] Note that we do not care about normalization. In the case of normalized states one obtains $K = N - 1$ [4].

linearity of the dot product we conclude that the measurement probability can generally be written in the form [4]:

$$p_{\text{meas}} = \mathbf{r} \cdot \mathbf{p}. \tag{4.5}$$

If one just wants to check whether the system is present or not, the associated measurement vector is defined by [4]:

$$\mathbf{r}^{\mathbb{1}} = \sum_{n} \mathbf{r}_n. \tag{4.6}$$

Consequently, the measurement probability is confined by [4]:

$$0 \leq \mathbf{r}^{\mathbb{1}} \cdot \mathbf{p} \leq 1. \tag{4.7}$$

Note that normalized states $\mathbf{p}_{\text{norm}}$ satisfy the upper bound, i.e. $\mathbf{r}^{\mathbb{1}} \cdot \mathbf{p}_{\text{norm}} = 1$. Additionally, equation (4.6) can be redefined if we take into account that the measurement device is able to distinguish L different outcomes, each related to a vector $\mathbf{r}_l$ $(l = 1, ..., L)$ [4]:

$$\mathbf{r}^{\mathbb{1}} = \sum_{l=1}^{L} \mathbf{r}_l. \tag{4.8}$$

So far, classical probability theories are obviously consistent with our first four axioms, but axiom 5, however, is not satisfied. According to Hardy's *continuity*-axiom one can always find a continuous reversible transformation between any two pure states of the system [4]. It is easy to prove that reversible transformations applied to a pure state will yield another pure state. To see this we will consider the contrary case for a moment. Let us assume there exists a reversible transformation Z and a pure state $\mathbf{p}$ such that

$$Z\mathbf{p} = \lambda \mathbf{p}_n + (1 - \lambda)\mathbf{p}_k, \tag{4.9}$$

with $\mathbf{p}_{n,k}$ being two different states [4]. Since Z is reversible, we conclude

$$\mathbf{p} = \lambda Z^{-1}\mathbf{p}_n + (1 - \lambda)Z^{-1}\mathbf{p}_k, \tag{4.10}$$

which implies that $\mathbf{p}$ is a mixture and therefore our initial assumption is apparently contradicted [4]. Hence, axiom 5 demands the existence of a continuous transformation that takes a pure state to another one without leaving the pure states' subspace. This, however, is not possible, since the system has only a finite number of N pure states [4]. Thus, classical probability theories do not fulfill axiom 5 as indicated at the beginning of chapter 4.

In contrast to the classical case, we will see in the next section 4.2 that quantum theory does indeed satisfy Hardy's postulated axioms.

4.2 The Structure of Quantum Theory

According to Hardy, quantum theories can be characterized in a way very similar to classical probability theories, but before we will deal with Hardy's description we give an overview of the standard formulation of quantum mechanics [4]:

- A quantum mechanical state can be expressed in terms of a positive, Hermitian operator $\hat{\rho}$ that fulfills the condition $\mathrm{Tr}(\hat{\rho}) = 1$[30].
- Each outcome of a probability measurement corresponds to a positive operator $\hat{A}_l$ $(l = 1, 2, ..., L)$ [4]:

$$\sum_{l=1}^{L} \hat{A}_l = \hat{\mathbb{1}}, \tag{4.11}$$

  where $\hat{\mathbb{1}}$ denotes the identity operator.

- When a measurement $\hat{A}$ is performed on a system represented by a state $\hat{\rho}$, the measurement probability is given by [4]:

$$p_{\text{meas}} = \mathrm{Tr}(\hat{A}\hat{\rho}). \tag{4.12}$$

- The evolution of a states is generally described by a completely positive, linear superoperator \$ [4]:

$$\hat{\rho} \rightarrow \$(\hat{\rho}). \tag{4.13}$$

In order to represent quantum theory by means of probability vectors, we first need to relate the number of degrees of freedom K with the number of dimensions N a quantum system has. For that purpose let us consider the matrix representation of Hermitian operators acting on an N dimensional Hilbert space [4]. Such matrices are specified by N real parameters along the diagonal and $\frac{1}{2}N(N-1)$ complex numbers above the diagonal resulting in N^2 real numbers. Consequently, Hermitian operators are made up of N^2 linearly independent projection operators $\hat{P}_k$ $(k = 1, ..., K = N^2)$ and we therefore have $K = N^2$ in the quantum case.
Using the vector

$$\hat{\mathbf{P}} = \begin{pmatrix} \hat{P}_1 \\ \hat{P}_2 \\ \vdots \\ \hat{P}_{K=N^2} \end{pmatrix}, \tag{4.14}$$

[30] This condition refers to normalized state vectors. If we do not care about normalization this relation generalizes to $0 \leq \mathrm{Tr}(\hat{\rho}) \leq 1$ [4].

any Hermitian operators is given by the dot product $\mathbf{a}\cdot\hat{\mathbf{P}}$, where $\mathbf{a}$ denotes a real vector. Following the concept of the trace formula (see equation (4.12)), Hardy suggests to alternatively characterize the states of the system by the probability vector [4]:

$$\mathbf{p}_S = \mathrm{Tr}(\hat{\mathbf{P}}\hat{\rho}). \tag{4.15}$$

Each component of this vector corresponds to the measurement probability when the respective projector $\hat{P}_n$ is applied to $\hat{\rho}$. Similar to the case of classical probability theories we can also associate the measurement operators $\hat{A}_l$ $(l = 1, 2, ..., L)$ with certain vectors $\mathbf{r}_M$ [4]:

$$\hat{A} = \mathbf{r}_M \cdot \hat{\mathbf{P}}. \tag{4.16}$$

Using this equation and equation (4.15) we can rewrite the trace formula in the form [4]:

$$p_{\text{meas}} = \mathbf{r}_M \cdot \mathbf{p}_S. \tag{4.17}$$

Moreover, a measurement $\hat{A}$ can be characterized in an analogous manner to the state probability vector by [4]

$$\mathbf{p}_M = \mathrm{Tr}(\hat{A}\hat{\mathbf{P}}), \tag{4.18}$$

whereas it is possible to define the state of a system by

$$\hat{\rho} = \hat{\mathbf{P}} \cdot \mathbf{r}_S. \tag{4.19}$$

Thus, the measurement probability is given by [4]:

$$p_{\text{meas}} = \mathbf{p}_M \cdot \mathbf{r}_S = \mathrm{Tr}(\hat{A}\hat{\rho}) = \mathrm{Tr}\left(\mathbf{r}_M^T \hat{\mathbf{P}} \hat{\mathbf{P}}^T \mathbf{r}_S\right) = \mathbf{r}_M^T D \mathbf{r}_S, \tag{4.20}$$

where D is a $K \times K$-matrix and the corresponding matrix elements are determined by [4]:

$$D_{ij} = \mathrm{Tr}\left(\hat{P}_i \hat{P}_j\right). \tag{4.21}$$

Note that the identity operator $\hat{\mathbb{1}}$ can also be represented in terms of projection operators and the identity measurement $\mathbf{r}^{\mathbb{1}}$ [4]:

$$\hat{\mathbb{1}} = \mathbf{r}^{\mathbb{1}} \cdot \hat{\mathbf{P}}. \tag{4.22}$$

In Hardy's formulation of quantum theory the superoperator \$ describing the evolution of a quantum system relates to a linear transformation of the probability state vector $\mathbf{p}_S$. This can easily be seen, if we substitute the transformation (4.13) into equation (4.15) [4]:

$$\begin{aligned}
\mathbf{p}_S = \mathrm{Tr}(\hat{\mathbf{P}}\hat{\rho}) &\rightarrow \mathrm{Tr}\left(\hat{\mathbf{P}}\$(\hat{\rho})\right) && (4.23)\\
&= \mathrm{Tr}\left(\hat{\mathbf{P}}\$(\hat{\mathbf{P}}^T D^{-1}\mathbf{p}_S)\right) && (4.24)\\
&= Z\mathbf{p}_S. && (4.25)
\end{aligned}$$

The matrix Z characterizing the linear transformation is defined by [4]

$$Z = \mathrm{Tr}\left(\hat{\mathbf{P}}\$(\hat{\mathbf{P}}^T)\right) D^{-1}. \tag{4.26}$$

In the previous derivation we made use of the relation $\mathbf{p}_S = D\mathbf{r}_S$, which follows from equation (4.17) and (4.20) [4].

4.3 Hardy's Axioms as Foundation of Quantum Theory

In this section we will forget about the characteristics of quantum and classical probability theories discussed before and show how quantum theory can be derived starting from Hardy's five axioms[31].

First, we will deal with the number of degrees of freedom K and the number of dimensions N a quantum system has and proof that in general $K(N) = N^r$, with r being a positive integer [4]. For that purpose Hardy suggests to consider an $N + 1$ dimensional system whose state is either restricted to an N dimensional subspace W or belongs to the one dimensional complement. In the first case, the *subspaces*-axiom implies that the system has $K(N)$ degrees of freedom, whereas in the second case, the system needs to have at least one degree of freedom, since K can never be less than one [4]. Generally, the state of the system will be a combination of W and its complement, each occurring with probability λ and $\lambda - 1$ respectively ($0 \leq \lambda \leq 1$). Such systems are at least characterized by $K(N) + 1$ degrees of freedom and, thus, we obtain that $K(N)$ is a strictly increasing function [4]:

$$K(N + 1) \geq K(N) + 1. \tag{4.27}$$

Additionally, from axiom 4 it follows that $K(N)$ is a completely multiplicative function, i.e. $K(N)$ fulfills $K(N_A N_B) = K(N_A)K(N_B)$. In his work Hardy manages to prove that any strictly increasing, completely multiplicative function can be written in the form $K(N) = N^\alpha$ [4]. Since K was initially defined as an integer, we conclude [4]:

$$K(N) = N^r \qquad (r = 1, 2, 3, \ldots). \tag{4.28}$$

Hardy proceeds with his proof by introducing K fiducial measurements which can be related to a basis set of measurement vectors $\mathbf{r}_n$ ($n = 1, 2, ..., K$). In an analogous manner Hardy also defines K linearly independent probability vectors $\mathbf{p}_k$ ($k = 1, 2, ..., K$) needed to uniquely specify the state of the system [4]. In the previous sections we have already seen that it is possible to represent both measurements and states either in terms of $\mathbf{r}$-type or $\mathbf{p}$-type vectors which are related to each other by a linear transformation [4]:

$$\mathbf{p}_S = D\mathbf{r}_S \tag{4.29}$$

$$\mathbf{p}_M = D^T\mathbf{r}_M. \tag{4.30}$$

[31] Note that we will only give a sketch of Hardy's proof in this work. The complete derivation is performed in [4].

The subscripts S and M refer to states and measurements, respectively. Using this linear transformation the measurement probability p_{meas} is given by [4]:

$$p_{\text{meas}} = \mathbf{r}_M^T D \mathbf{r}_S. \tag{4.31}$$

In the following we will eliminate the case $K = N$ and show that quantum theory satisfies the condition $K(N) = N^2$. If we consider a system with $K = N$ degrees of freedom, the K fiducial measurements can be chosen to be identical with the N basis vectors spanning the state space [4]. The equality of fiducial and basis vectors indicates that D correlates with the identity operator, i.e. $D = \mathbb{1}$, and therefore pure states fulfill [4]:

$$\mathbf{r}^T D \mathbf{r} = 1. \tag{4.32}$$

The relation $D = \mathbb{1}$ implies that [4]

$$\sum_{k=1}^{N} (p_k)^2 = 1, \tag{4.33}$$

with p_k being the kth component of the vector $\mathbf{p}$. Moreover, $0 \leq p_k \leq 1$, and since we now require $\mathbf{p}$ to be normalized, we obtain [4]:

$$\sum_{k=1}^{N} p_k = 1. \tag{4.34}$$

As we can see, only those vectors $\mathbf{p}$ satisfy the conditions stated above that have one p_k equal to one, while all the other entries are equal to zero. Hence, the pure states are identical with the basis states which is exactly the case in classical probability theories [4]. Since these N basis vectors form a discrete set, the *continuity*-axiom is not fulfilled and, thus, we can rule out any theory with $K = N$.
According to the *simplicity*-axiom, K takes the minimum value in agreement with our axioms, leading to the case of $K(N) = N^2$ and by using the *subspaces*-axiom we can easily generalize to arbitrary N [4].

In order to reproduce the formalism of quantum theory we derived in section 4.2, we primarily need to recover the expressions for states $\hat{\rho}$ and measurements $\hat{A}$ given by equations (4.16) and (4.19). This can be achieved by calculating the vector $\hat{\mathbf{P}}$ – whose components are N^2 linearly independent projection operators $\hat{P}_k$ (see equation (4.14)) – from D and correlating it with the vector representing the state and the measurement, respectively [4]. Thus, we obtain the measurement probability p_{meas} described by the well-known trace formula [4]:

$$p_{\text{meas}} = \text{Tr}(\hat{A}\hat{\rho}). \tag{4.35}$$

Hardy is finally able to show that the most general evolution of the state $\hat{\rho}$ after a measurement is equivalent to the evolution predicted by quantum theory [4]. If we consider the case of a system exiting the measurement device after a measurement has been performed, the state of this system will be changed into one of l ($l = 1, 2, ..., L$) distinct configurations corresponding to the possible measurement outcomes. Therefore it is possible to relate each outcome to a linear transformation Z_l that alters the normalization factor of the state such that it agrees with the probability of obtaining the specific outcome [4]. This results in the following condition [4]:

$$\mathbf{r}^{\mathbb{1}} \cdot Z_l \mathbf{p} = \mathbf{r}_l \cdot \mathbf{p} \qquad \forall \mathbf{p}. \tag{4.36}$$

Furthermore we can deduce that [4]

$$\left(\sum_{l=1}^{L} Z_l \right)^T \mathbf{r}^{\mathbb{1}} = \mathbf{r}^{\mathbb{1}}, \tag{4.37}$$

meaning that the set of all feasible transformations does not change the normalization factor of the state. The relation above correlates with the constraint [4]

$$\mathrm{Tr} \sum_l \$(\hat{\rho}) = \mathrm{Tr}(\hat{\rho}) \qquad \forall \hat{\rho}, \tag{4.38}$$

and since the evolution of the state given by the completely positive superoperator \$ can be described by $\$(\hat{\rho}) = \sum_l \hat{M}_l \hat{\rho} \hat{M}_l^\dagger$, it can be proved that this equation is identical with the usual constraint on superoperators [4]:

$$\sum_l \hat{M}_l^\dagger \hat{M}_l = \hat{\mathbb{1}}. \tag{4.39}$$

Since the two expressions (4.36) and (4.37) limiting the number of possible transformations are also valid for classical probability theories, Hardy believes this might lead to another approach to the measurement problem encountered in quantum theory [4]. Thus, we see that Hardy's set of reasonable axioms provides new insights into some of the most characteristic issues of quantum mechanics.

5 Robert W. Spekkens - A Toy Theory

In this last chapter we will present Robert Spekkens' *toy theory* [12] which, in contrast to the other approaches discussed before, is not able to derive quantum theory but still reproduces a great number of quantum phenomena. Spekkens' model is based on a single principle, which states that the amount of information one possesses about the state of a system must always be quantitatively equal to the lack of information in a state of maximal knowledge [12]. This assumption is deeply connected with the question whether quantum states are considered to be ontic states, i.e. states of reality, or epistemic states, that is states of (incomplete) knowledge[32].

[32] The simplest example of an ontic state in classical mechanics is a state in the phase space. Epistemic states, on the other hand, are often described as a convex sum of pure ontic states. Such states are usually encountered in statistical mechanics [12].

Physicists have been debating on this issue since the very beginning of quantum physics and today the most widespread concept is the ontic point of view [12]. Spekkens, however, argues that the epistemic view of states grants a better conceptual understanding of various quantum features, such as interference, noncommutativity and entanglement, whereas the ontic viewpoint is less intuitive.

In the following, we will develop Spekkens' model starting from our fundamental principle, the so-called *knowledge balance principle*, which will be discussed in more detail in the next section 5.1.
But before we proceed, it is important to emphasize once again that Spekkens' toy theory cannot be considered as an alternative concept to the standard formulation of quantum mechanics. Spekkens' model is built on local and noncontextual hidden variables, but since the violation of Bell's inequality and the Kochen-Specker-theorem we know that it is impossible to understand quantum theory in this way [12].

5.1 The Knowledge Balance Principle

In line with Zeilinger's approach (see chapter 2), Spekkens' toy theory is based on the assumption that quantum states are primarily states of incomplete knowledge, which indicates that the description of quantum systems is subject to a limited amount of information. Starting point is Spekkens' *knowledge balance principle* from which the entire toy theory can be derived [12]:

> *"If one has maximal knowledge, then for every system, at every time, the amount of knowledge one possesses about the ontic state of the system at that time must equal the amount of knowledge one lacks."*[33]

In order to apply this principle, we need to define an adequate measure of information to quantify the knowledge one posses about the state of a system. For that purpose we introduce a minimal number of propositions with corresponding truth values, "yes" or "no", required to completely characterize the ontic state [12].
To illustrate the implications of this set of propositions, which we refer to as *canonical* set, let us consider the case of a system that can be located in one of four distinct ontic states. A possible list of propositions able to determine the actual state of the system could be four statements concerning the yes/no-questions in which of the ontic states the system can be found [12]. As one might conjecture, this list of propositions does not represent the canonical set, since the number of questions can be reduced, if we always group two of the four ontic states to form a subset. Thus, two questions yielding four answers suffice to uniquely specify the state of the system, e.g. [12]:

- *"Is the system in the set* $\{1,2\}$*, or not?"*
- *"Is the system in the set* $\{1,3\}$*, or not?"*

[33] R. Spekkens, 2008, p. 3.

Using a canonical set of propositions to specify the state of a system, we can now determine an appropriate expression of the measure of information. Spekkens suggests to define it as the maximum number of propositions with known truth values in a variation over all possible canonical sets [12].

In the following, we will only consider systems that satisfy the knowledge balance principle stated above. In order to obtain an equality between the amount of information one possesses and the lack of information in a state of maximal knowledge, there always needs to be an even number of propositions in the canonical set. Hence, the simplest system, which we call the elementary system, is obviously characterized by a total of two questions and can therefore be located in one of four possible ontic states [12].
To describe more complex systems within the toy theory we suppose all systems to be composed of these elementary systems. A general system consisting of N elementary systems will then have $2N$ propositions in the canonical set and 2^{2N} possible ontic states [12].

5.2 The Description of Elementary Systems

In the last section 5.1 we have already ascertained that the most elementary system in accordance with Spekkens' knowledge balance principle is determined by a minimum number of two propositions in the canonical set. The actual ontic state, however, is yet unspecified, since the fundamental principle allows only one of these two questions to be answered [12]. We are therefore left with a system that still can be located in two possible ontic states. If we denote the four distinct states with '1', '2', '3' and '4', the epistemic states of the system are defined as disjunctions of the ontic states [12]:

$$\begin{gathered} 1 \vee 2 \\ 1 \vee 3 \\ 1 \vee 4 \\ 2 \vee 3 \\ 2 \vee 4 \\ 3 \vee 4. \end{gathered}$$

In contrast to these six states, which represent the states of maximal knowledge, there exists only a single state of nonmaximal knowledge for elementary systems that is associated with the epistemic state [12]

$$1 \vee 2 \vee 3 \vee 4.$$

This state corresponds to the case that the truth values of both propositions in the canonical set are unknown [12].
The elementary systems in Spekkens' toy theory can be characterized in a way very similar to qubits in quantum information theory [12]. The six epistemic states of maximal

knowledge correlate with the pure qubit states[34] [12]

$$\begin{aligned} 1 \vee 2 &\Leftrightarrow |0\rangle \\ 3 \vee 4 &\Leftrightarrow |1\rangle \\ 1 \vee 3 &\Leftrightarrow |+\rangle \\ 2 \vee 4 &\Leftrightarrow |-\rangle \\ 2 \vee 3 &\Leftrightarrow |+i\rangle \\ 1 \vee 4 &\Leftrightarrow |-i\rangle, \end{aligned} \tag{5.1}$$

whereas the state of nonmaximal knowledge can be associated with the completely mixed state [12]:

$$1 \vee 2 \vee 3 \vee 4 \Leftrightarrow \frac{\mathbb{1}}{2}.$$

The analogy between the elementary systems in the toy theory and qubits suggests a simple geometric representation of the epistemic states very similar to the Bloch sphere representation encountered in quantum information theory [12] (see figure 1). Orthogonal pure qubit states correspond to antipodal points on the surface of the Bloch sphere, while mixed states, i.e. convex combinations of pure states, are represented by the points inside the ball.
In order to reproduce a relation analogous to the orthogonality of quantum states, Spekkens deems it reasonable to introduce the notion of an ontic base of an epistemic state [12]. The ontic base is defined as the set of ontic states that is in agreement with the epistemic state. For example, if we consider the case of an elementary system represented by the epistemic state $1 \vee 2$, the corresponding ontic base is the set $\{1, 2\}$. In particular, two distinct epistemic states are called disjoint, if the intersection of their respective ontic bases is empty. Thus, we conclude that the disjointness of epistemic states in the toy theory resembles the feature of orthogonality in quantum mechanics [12].

We proceed with the introduction of the concept of pure and mixed states in the toy theory. As already mentioned, the six epistemic states of maximal knowledge are described in a manner analogous to pure qubit states in quantum information theory, whereas the single state of nonmaximal knowledge, $1 \vee 2 \vee 3 \vee 4$, correlates with the completely mixed state $\mathbb{1}/2$ [12]. Since a general mixed quantum state can be written as a convex sum of pure states, it is reasonable to introduce a similar operation in the toy theory. In order to define a meaningful convex combination, Spekkens argues that epistemic states need to fulfill the following two requirements [12]:

- First, the epistemic states must be disjoint.
- Secondly, the union of the ontic bases of the epistemic states has to correspond with the ontic base of the resulting epistemic state.

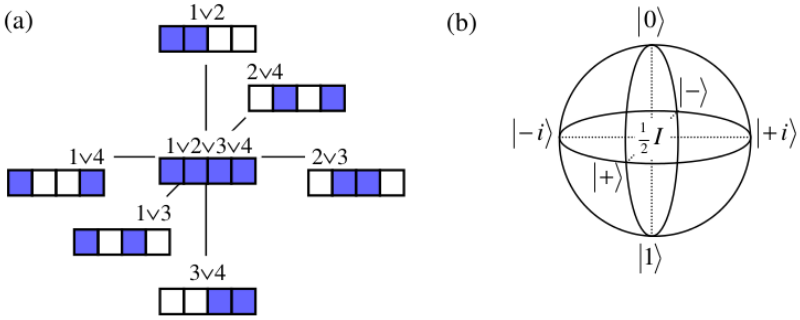

Figure 1: Analogy to the Bloch sphere. (a) A graphical representation of epistemic states. The states are specified by four squares representing the four possible ontic states. The colored squares characterize the set of ontic states in which the system can be located. (b) Bloch sphere representation of single qubits. Pure states correspond to the points on the surface of the sphere. (R. Spekkens, 2008, p. 6)

Using the operation symbol '$+_{cx}$' to denote the convex combination of two epistemic states, it is therefore possible to decompose the mixed state $1 \vee 2 \vee 3 \vee 4$ in various ways in accordance with the conditions listed above [12]:

$$\begin{aligned} 1 \vee 2 \vee \vee 3 \vee 4 =& (1 \vee 2) +_{cx} (3 \vee 4) \\ =& (1 \vee 3) +_{cx} (2 \vee 4) \\ =& (2 \vee 3) +_{cx} (1 \vee 4). \end{aligned} \tag{5.2}$$

This decomposition is analogous to the well-known convex decomposition of the completely mixed qubit state $\mathbb{1}/2$ [12]:

$$\begin{aligned} \frac{\mathbb{1}}{2} =& \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \\ =& \frac{1}{2}\,|+\rangle\,\langle +| + \frac{1}{2}|-\rangle\langle -| \\ =& \frac{1}{2}\,|+i\rangle\,\langle +i| + \frac{1}{2}\,|-i\rangle\,\langle -i|. \end{aligned} \tag{5.3}$$

As we can see, in quantum theory as well as in Spekkens' toy theory the convex decomposition of mixed states is not unique [12]. This fact indicates a certain connection between quantum states and epistemic states, i.e. states of incomplete knowledge, and therefore supports the epistemic point of view.

Another characteristic feature of quantum mechanics Spekkens' toy theory is able to reproduce concerns the impossibility of creating a universal state inverter [12]. Generally speaking, a universal state inverter maps an arbitrary quantum state $|\psi\rangle$ onto the

[34] Note that in quantum information theory $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |-\rangle)$ and $|\pm i\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|-\rangle)$ [12].

orthogonal state $|\bar{\psi}\rangle$, that is it transforms $|\psi\rangle$ to $|\bar{\psi}\rangle$ such that $\langle\psi|\bar{\psi}\rangle = 0$. However, such a device cannot be constructed, since mappings of this kind are not unitary [12]. An analogous transformation within the toy theory is represented by any mapping which takes a pure epistemic state of an elementary system to the state that is disjoint with it. We therefore have the following transformation [12]:

$$\begin{aligned} 1 \vee 2 &\leftrightarrow 3 \vee 4, \\ 1 \vee 3 &\leftrightarrow 2 \vee 4, \\ 2 \vee 3 &\leftrightarrow 1 \vee 4. \end{aligned} \tag{5.4}$$

This transformation is also impossible to achieve, since the first two conditions in equation (5.4) require the mapping 1 ↔ 4 and 2 ↔ 3, which apparently contradicts the third condition [12]. Thus, we conclude that there exists no universal state inverter in the toy theory.

As we have seen so far, the elementary systems in Spekkens' toy theory have much in common with qubits in quantum information theory and this fact supports the point of view that quantum states are primarily states of incomplete knowledge.
In the next section 5.3 we are going to extend the toy theory and consider more complex systems composed of pairs of elementary systems.

5.3 Pairs of Elementary Systems

Composite systems consisting of pairs of elementary systems are obviously specified by four propositions in the canonical set, since each elementary systems itself is represented by two propositions [12]. Consequently, the pair can be located in one of sixteen distinct ontic states, which can be written as conjunctions[35] of the ontic states of the individual elementary systems [12]:

$$\begin{array}{cccc} 1\cdot 1 & 1\cdot 2 & 1\cdot 3 & 1\cdot 4 \\ 2\cdot 1 & 2\cdot 2 & 2\cdot 3 & 2\cdot 4 \\ 3\cdot 1 & 3\cdot 2 & 3\cdot 3 & 3\cdot 4 \\ 4\cdot 1 & 4\cdot 2 & 4\cdot 3 & 4\cdot 4. \end{array}$$

Analogous to the description of single elementary systems it is also possible to represent pairs of elementary systems graphically by 4 × 4-arrays of cells (compare figure 2) [12]. If we consider a composite system consisting of two elementary subsystems A and B, the rows of the 4 × 4-array determine the ontic states of system A and the columns specify the possible ontic states of system B.

Even though the pair has four propositions in the canonical set, the knowledge balance principle allows only two of them to be answered. Thus, in a state of maximal

[35] Conjunctions will be denoted by the sysmbol '·' [12].

knowledge the composite system can be located in a set of four distinct ontic states and the possible pure epistemic states are given as disjunctions, e.g. [12]:

$$(1 \cdot 3) \vee (1 \cdot 4) \vee (2 \cdot 3) \vee (2 \cdot 4). \tag{5.5}$$

The valid set of pure epistemic states can be further reduced, if we apply the knowledge balance principle to each of the elementary systems, A and B, separately [12]. It follows that each system is represented at most by the truth value of a single proposition and Spekkens is finally able to show that two different kinds of 4×4-arrays suffice to completely characterize the possible epistemic states of maximal knowledge (see figure 2) [12].

Analogous to quantum states, it is also possible to distinguish between correlated and uncorrelated epistemic states [12]. Conjunctions of states of maximal knowledge lead to uncorrelated states, which can be generally described in the form [12]

$$(a \vee b) \cdot (c \vee d), \tag{5.6}$$

with $a, b, c, d \in \{1, 2, 3, 4\}$ and $a \neq b, c \neq d$. These states are graphically represented by permutations of the left array in figure 2 and correspond to product states in quantum theory, e.g. $|0\rangle|0\rangle$ (compare equation (5.1)).
Correlated epistemic states, on the other hand, can be written in the form [12]

$$(a \cdot e) \vee (b \cdot f) \vee (c \cdot g) \vee (d \cdot h), \tag{5.7}$$

with $a, b, c, d, e, f, g, h \in \{1, 2, 3, 4\}$ and $a \neq b \neq c \neq d, e \neq f \neq g \neq h$. In this case the propositions with known truth values in the canonical set concern the relations between the individual elementary systems and, thus, such epistemic states are analogous to maximally entangled states [12]. Correlated states are represented by any permutation of the second array in figure 2.
In contrast to the description of single elementary systems, there exist more than one epistemic state of nonmaximal knowledge for composite systems in the toy theory. In the case of a pair of elementary systems these states are characterized by either one or none answered questions in the canonical set [12]. Specifically, the completely mixed state

$$(1 \vee 2 \vee 3 \vee 4) \cdot (1 \vee 2 \vee 3 \vee 4), \tag{5.8}$$

which corresponds to none answered proposition, can be associated with the two-qubit state, $\mathbb{1}/2 \otimes \mathbb{1}/2$ [12].

In the following, we will deal with the well-known *no-cloning theorem* and demonstrate how this no-go theorem arises from an epistemic point of view. According to the theorem, which was first proved by Wootters and Zurek in 1982 (compare [7]), the cloning of an arbitrary quantum state is impossible, since there exists no unitary operator that is able to perform such a transformation[36].

[36] A transformation that copies an arbitrary initial quantum state onto another one must preserve inner products and therefore needs to be unitary [12].

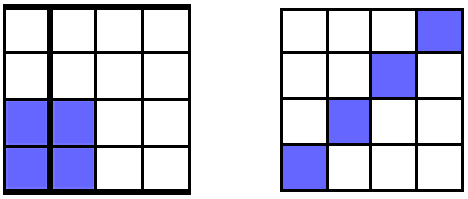

Figure 2: Graphical representation of valid epistemic states of maximal knowledge for a pair of elementary systems in agreement with the knowledge balance principle [12]. The possible ontic states of the composite system AB are specified by 4×4-arrays of squares. The colored squares characterize the set of ontic states in which the system is known to be. The entire set of valid epistemic states is given by permutations of the rows and columns of the arrays depicted above [12]. (R. Spekkens, 2008, p. 12)

In the context of an epistemic viewpoint, the cloning process is represented by a transformation that copies an unknown initial state, $(a \vee b)$, onto a fixed state, $(c \vee d)$ [12]:

$$(a \vee b) \cdot (c \vee d) \to (a \vee b) \cdot (a \vee b). \tag{5.9}$$

In order to illustrate the impossibility of the no-cloning theorem for nondisjoint epistemic states, Spekkens suggests to consider the cloning process for the set $\{1 \vee 3, 3 \vee 4\}$ [12]. If we choose the fixed state to be $(1 \vee 2)$, we have the transformation [12]:

$$\begin{aligned} (1 \vee 3) \cdot (1 \vee 2) &\to (1 \vee 3) \cdot (1 \vee 3) \\ (3 \vee 4) \cdot (1 \vee 2) &\to (3 \vee 4) \cdot (3 \vee 4). \end{aligned} \tag{5.10}$$

To prove that this transformation does not exist in the toy theory we need to overlap the uncorrelated epistemic states on both sides of equation (5.10). The resulting ontic states are $3 \cdot 1$ and $3 \cdot 2$ on the left-hand side and $3 \cdot 3$ on the right-hand side. Due to the fact that the number of ontic states is not preserved, we conclude that the cloning transformation is not possible [12]. By taking any permutation of the epistemic states, $1 \vee 3$ and $3 \vee 4$, we can generalize the proof to an arbitrary set of states.

As we can see, Spekkens' toy theory based on the simple knowledge balance principle is indeed able to reproduce a large variety of quantum features. The analogy between the epistemic states in the toy theory and quantum states certainly supports an epistemic view of quantum states, even though we must be aware of the fact that the toy theory is in no way equivalent to quantum theory.

6 Summary and Concluding Remarks

In this work we have considered four promising approaches that try to base quantum theory on conceptual foundations in terms of simple axioms and meaningful principles.

As already discussed at the beginning, we are not interested in the mathematical formalism of quantum mechanics concerning complex Hilbert spaces and Hermitian operators, instead, we want to find a foundation that is able to derive the very characteristics of quantum physics, which significantly distinguish quantum from classical mechanics.

Although all the approaches introduced here start from an information-theoretic point of view, the strategies to deduce quantum theory are in part very different. Zeilinger's fundamental principle is based on the assumption that the description of the world is represented in terms of propositions (see section 2.1). Consequently, quantization arises from the fact that quantum objects can carry only a limited amount of information [1].

In Clifton, Bub and Halvorson's approach, on the other hand, the physical world is subject to three fundamental information-theoretic constraints. As shown in section 3, these constraints are connected with some characteristic features of quantum mechanics, such as nonlocality and noncommutative algebras of observables.

Moreover, Hardy's work (see section 4) based on a set of five reasonable axioms represents another promising approach able to derive quantum theory. In contrast to the other approaches discussed before, Hardy suggests that quantum theory is more similar to classical probability theories as one might expect at first sight. Hardy is able to show that just the *continuity*-axiom – more precisely, just the single word "continuous" of this axiom – is not consistent with classical probability theories [4]. This fact provides completely new insights into the discussion on the conceptional foundations of quantum theory and raises the question why so many quantum phenomena, for example entanglement, do not have a classical counterpart.

Consequently, Hardy regards quantum theory as a new continuous kind of probability theory, as he states in [4]:

> "Quantum theory is, in some respects, both superior to and more natural than classical probability theory (and therefore classical theories in general) since it can describe evolution for finite systems in a continuous way. Since nature is quantum, not classical, it is to be expected that quantum theory is ultimately the more reasonable theory."[37]

In connection with the approaches developed by Zeilinger and Hardy, we also discussed Spekkens' toy theory in section 5. Even though not equivalent to quantum theory, the toy theory based on the knowledge balance principle supports an epistemic view of quantum states. If one accepts that both pure and mixed states are rather states of incomplete knowledge than states of reality, many quantum features can be understood in a very intuitive way.

On the other hand, the epistemic point of view raises questions concerning the reality of quantum objects. What does the knowledge one possesses about a state actually represent, if quantum states are not states of reality? This and other question are still

[37] L. Hardy, 2008, p. 27.

unanswered and therefore the knowledge balance principle does certainly not suffice to deduce quantum theory without an additional axiom regarding the reality of states. Nevertheless, Spekkens' toy theory supports Hardy's point of view and therefore we can conclude that quantum theory is primarily a theory of information transfer, as Clifton, Bub and Halvorson put it in [10]:

> "That is, we are suggesting that quantum theory be viewed, [...] as a theory about the possibilities and impossibilities of information transfer."[38]

[38] Clifton, Bub and Halvorson, 2003, p. 1563.

References

[1] A. Zeilinger, “Foundational principle for quantum mechanics”, Foundations of Physics, Vol . 29, No. 4, 1999.

[2] Paul Dirac, ”The Principles of Quantum Mechanics”, 4. Auflage, The International Series of Monographs on Physics 27, Oxford Science Publications, Oxford University Press 1988, ISBN 0198520115.

[3] J. Neumann, “Mathematical Foundations of Quantum Mechanics”, Beyer, R. T., trans., Princeton Univ. Press. 1996 edition: ISBN 0-691-02893-1.

[4] L. Hardy, “Quantum Theory From Five Reasonable Axioms”, quant-ph/0101012 (2001).

[5] Michael A. Nielsen and Isaac L. Chuang, “Quantum Computation and Quantum Information”, Cambridge: Cambridge UP, 2010.

[6] N. Herbert, “FLASH–A Superluminal Communicator Based upon a New Type of Quantum Measurement”, Foundations of Physics, Bd.12, 1982, S.1171.

[7] W. Wootters and W. Zurek, “A Single Quantum Cannot be Cloned”, 1982, Nature 299: 802–803.

[8] C. Brukner and A. Zeilinger , “Information and Fundamental Elements of the Structure of Quantum Theory” , in ”Time, Quantum, Information”, edited by L. Castell and O. Ischebek (Springer, 2003).

[9] C. Brukner, M. Zukowski, A. Zeilinger, “The essence of entanglement” quant-ph/0106119 (2001).

[10] R. Clifton, J. Bub, H. Halvorson, “Characterizing quantum theory in terms of information-theoretic constraints”, Foundations of Physics 33, 1561-1591 (2003).

[11] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, “Non-commuting mixed states cannot be broadcast”, Phys. Rev. Lett. 76, 2318 (1996).

[12] R. Spekkens, “In defense of the epistemic view of quantum states: a toy theory”, quant-ph/0401052 (2004).

[13] E. Schrödinger, “Die gegenwärtige Situation in der Quantenmechanik”, Naturwissenschaften 23, 807 ff. (1935). Translation published in Proc. Am. Phil. Soc., Vol. 124, No. 5 (Oct. 10, 1980), pp.323-338 edited by J. Trimmer.

[14] C. Brukner and A. Zeilinger, “Operationally Invariant Information in Quantum Measurements”, Phys. Rev. Lett. 83 (1999) 3354-3357, [arxiv] quant-ph/0005084.

[15] K. Davidson, “C^*-algebras by Example”, American Mathematical Soc., 1996. ISBN: 0821871897, 9780821871898.