

Standard Quantum Theory Derived from First Physical Principles

Mehran Shaghghi

University of Illinois-Chicago, Chicago 60612, IL, USA

The mathematical formalism of quantum theory has been known for almost a century, but its physical foundation has remained elusive. In recent decades, many physicists have noted connections between quantum theory and information theory. In this study, we present a physical account of the derivation of quantum theory's mathematical formalism based on information considerations in physical systems. We postulate that quantum systems are physical systems with only one independent adjustable variable. Using this physical postulate along with the conservation of the total probability, we derive the standard Hilbert space formalism of quantum theory, including the Born probability rule. Our complete derivation of quantum theory provides a clear and concise physical foundation for the mathematical formalism of quantum mechanics.

1 Introduction

Quantum theory has been a successful mathematical framework for describing the behavior of quantum systems. It was developed almost a century ago mainly by Dirac and von Neumann and is based on Hermitian operators and their eigenvectors and eigenvalues [1, 2]. Despite its success, a fundamental physical foundation for quantum theory is still lacking. Since its development, numerous attempts have been made to derive the formalism of quantum theory from physical axioms or first principles [3-20]. However, these attempts have either been incomplete or based on abstract mathematical assumptions that lack a clear physical basis.

In the last few decades, there has been a growing interest in using an information-theoretic approach to quantum theory. This is partly due to the advocacy of John Wheeler for the relevance of information theory for understanding quantum physics [21-23], as well as the developments in the field of quantum information. Some efforts involve analyzing the internal structure and logic of the theory to identify its cornerstones. Others have looked to gain new insights into the characteristics that shape quantum theory by examining the relationships between the various features of quantum systems. For example, Clifton, Bub, and Halvorson [17] used a C*-algebraic formalism to demonstrate how certain information-theoretic constraints on physical systems, namely, no superluminal information transfer, no perfect broadcasting, and no bit commitment, can lead to other features of quantum theory such as kinematic independence, noncommutativity, and nonlocality. Another approach, initiated by Hardy [16], focuses on the general properties of probability theories and discusses the criteria that distinguish quantum theory from classical probability theories.

In this report, we present a derivation of the standard formalism of quantum theory based on a clear physical foundation. The fundamental concept we employ is the limited information capacity of quantum systems: quantum systems are those physical systems that can hold only a single physical message (a piece of information) at a time, as formalized in the complementarity principle. This contrasts with classical systems, which can hold multiple messages simultaneously. The limits on the information capacity of a physical system can be understood as each independently adjustable variable of the system can physically represent one independent piece of information at a time. For physical systems with only one adjustable variable, this constraint prevents them from carrying more than one piece of information (one message) at a time.

A range of factors can restrict the adjustable variables of physical systems. For example, the adjustable variables can be confined by the screenings in the experimental setup, such as in the double-slit experiment where the use of a coherent beam of photons determines their energy and direction, leaving their polarization as the only adjustable variable. Alternatively, the adjustable variables of physical systems can be constrained under extreme high pressures, low temperatures, or intense electromagnetic fields, such as in laser-cooled trapped ions or superconductivity. And in some cases, the other independent adjustable variables of a system may be abstracted away, as the specified property of the system (e.g., the electron spin) can be considered isolated and entirely separate from the rest. In this report, we demonstrate that the standard formalism of quantum mechanics (the Hilbert space and the Born probability rule) can be explicitly and systematically derived as the theory of the physical systems with a single adjustable variable.

2 Single Variable Systems: Information-Theoretic Considerations

Several information-theoretic definitions of a quantum system have been proposed in the literature as potential foundations for the derivation of quantum theory. For instance, Rovelli proposed the axiom that “there is a maximum amount of relevant information that can be extracted from a [quantum] system” [12]. Similarly, Zeilinger suggested “a foundational principle for quantum mechanics” that states a quantum system is “an elementary system [that] carries 1 bit of information” [14]. While these axioms offer some interesting explanations for certain quantum phenomena, such as randomness and entanglement [24-28], or a framework for reconstructing the formalism [12], they were not successful in deriving the full formalism of quantum theory.

Here, we present a mathematical theory that describes the behavior of systems with a single adjustable variable under measurements. Several information-theoretic considerations affect the properties of such systems. Since these systems have no more than one adjustable variable, they can contain only one message (i.e., one piece of information) at a time. We refer to these systems as Single Message (SM) systems. Since an SM system contains only a single piece of information, its state can be defined by the outcome of the last measurement performed on it. Note that performing a measurement always produces an outcome, even if the outcome is a zero reading.

The unique property of SM systems is that when the information content of an SM system is determined via a measurement, performing any subsequent independent measurement (which measures an independent proposition) must yield a result with zero informational gain, that is, a random outcome. The equivalence of randomness and zero information gain is well established through Shannon's information measure [29]. This means that in SM systems, performing independent measurements leads to random transformations of the system into new unpredictable states. In other words, performing measurements on SM systems generally involves an element of randomness since the system cannot hold more than one piece of information at any time.

In short, due to the single messaging capacity of the SM systems, performing independent measurements does not yield deterministic results, but rather involves elements of unpredictability and randomness, as in measuring the X component of an electron spin that is in the Z+ state. This constraint also implies that performing independent measurements on an SM system results in completely random changes in the system's state, making it impossible to predict the outcomes of measurements with certainty. However, if measurements are not entirely independent, the future states of the SM system can be predicted probabilistically based on the outcome of the last measurement performed on the system. Therefore, the dynamics of the SM system can be described probabilistically.

3 Construction of the Formalism

In analyzing the possible outcomes of an SM system in a given measurement scenario, it is crucial to examine the relationship between two measurements: the last measurement performed on the system (which has defined the state of the system) and the measurement for which we want to calculate the probabilities of its outcomes. These two measurements can either be dependent or independent. In the case of independent measurements, the outcome of the first measurement does not influence the outcome of the second, while for dependent measurements, certain outcomes of the second measurement can happen less or more likely based on the outcome of the first. A dependent measurement, for example, would be measuring the spin of an electron in the direction that is tilted 20 degrees from the z-axis in the zx-plane when it is in \tilde{S}_z^z state.

Consider an SM system in a certain state. The aim is to determine the probabilities of the system to end up in each of the outcomes of a measurement, \hat{M} , to be performed on it. Without loss of generality, we consider measurements with N distinguishable outcomes (generalization to the infinite case is straightforward). In what follows this notation is used: a measurement type K is represented as \hat{M}^K , with possible outcomes $\tilde{M}_1^K, \tilde{M}_2^K, \dots, \tilde{M}_N^K$ that are independent members of the set $S(\hat{M}^K)$ defined as:

$$S(\hat{M}^K) = \{\tilde{M}_1^K, \dots, \tilde{M}_N^K\} \quad (1)$$

The probabilities of the SM system for the outcomes of this measurement can be represented as:

$$P(\widehat{M}^K) = \{P_1^K, \dots, P_N^K\} \quad (2)$$

Since performing the measurement eventually results in an outcome, it follows that:

$$\sum_{j=1}^N P_j^K = 1 \quad (3)$$

In certain cases, the probabilities, P_j^K , are easily determinable. For example, if the SM system has just undergone the measurement \widehat{M}^Q and the outcome \widetilde{M}_i^Q is resulted, then repeating the same measurement \widehat{M}^Q will not change the state of the system. Therefore, in the case that the next measurement is \widehat{M}^Q , the probabilities of the system for the measurement outcomes are:

$$P_j^Q = \delta_{j,i} \quad (4)$$

On the other hand, in the case that the next measurement is an independent measurement \widehat{M}^R , then the outcomes $\{\widetilde{M}_1^R, \dots, \widetilde{M}_N^R\}$ are all equally likely: $P_1^R = P_2^R = \dots = P_N^R$. Hence, the probabilities of the system for the outcomes of the measurement can be written as:

$$P_j^R = \frac{1}{N}, \forall j \in \{1, \dots, N\} \quad (5)$$

where N is the total number of the outcomes.

Besides these two cases, an extensible framework is required to evaluate the probabilities of the SM system for the outcomes of a general type of measurement, i.e., for measurements that are neither the same as the previous one nor fully independent of it. Consider two measurements, \widehat{M}^K & \widehat{M}^L , which are not fully independent, meaning certain outcomes of the second measurement can happen less or more likely based on the outcome of the first. The interdependence between the outcomes of the measurements can be defined as:

$$T_{j,i}^{L,K} = P(\widetilde{M}_j^L | \widetilde{M}_i^K) \quad (6)$$

which represent the probability of obtaining the j^{th} outcome in measurement \widehat{M}^L , given the i^{th} outcome of measurement \widehat{M}^K . These interdependences can be framed in an $N \times N$ "interdependence matrix" of the two measurements.

The interdependence matrix has several properties. For example, for any fixed i

$$\sum_j T_{j,i}^{L,K} = 1 \quad (7)$$

as the \widehat{M}^L measurement ultimately produces an outcome, and the probabilities should add up to 1. Representing the probabilities of the SM system for the measurement \widehat{M}^K as P_i^K , the probabilities of the system for the measurement \widehat{M}^L is determined as

$$P_n^L = \sum_i P(\widetilde{M}_n^L | \widetilde{M}_i^K) P_i^K = \sum_i T_{n,i}^{L,K} P_i^K \quad (8)$$

This is the sum over all possible ways that the system can result in \widetilde{M}_n^L . The interdependence matrices need to conserve the total probability, i.e., $\sum_n P_n^L = 1$. This is ensured by:

$$\sum_n P_n^L = \sum_n \sum_i T_{n,i}^{L,K} P_i^K = \sum_i P_i^K \sum_n T_{n,i}^{L,K} = 1 \quad (9)$$

based on (3) & (7).

Alternatively, \widehat{M}^L can be taken as the first measurement, in which the system has the probabilities of P_i^L . Then the probabilities of the system for the second dependent measurement \widehat{M}^K , can be evaluated as above using the interdependence matrix of:

$$S_{j,i}^{K,L} = P(\widetilde{M}_j^K | \widetilde{M}_i^L). \quad (10)$$

The interdependence matrices represent the probabilistic correlations between the outcomes of two measurements. Functionally, they map the probabilities of the system for one measurement to that of the other as shown in (8). In general, these mapping matrices should have certain properties. Firstly, the components represent probabilities, therefore they are positive numbers not greater than 1:

$$\begin{aligned} 0 \leq T_{j,i} &\leq 1 \\ 0 \leq S_{j,i} &\leq 1 \end{aligned} \quad (11)$$

Additionally, they must obey:

$$\begin{aligned} \sum_j T_{j,i} &= 1 \\ \sum_j S_{j,i} &= 1 \end{aligned} \quad (12)$$

for any fixed i , as measurements necessarily produce an outcome (cf. (7)).

The interdependence matrices map the probabilities of the SM system from one measurement-space to another according to (8):

$$\begin{aligned}
P_j^L &= \sum_i T_{j,i} P_i^K \\
P_j^K &= \sum_i S_{j,i} P_i^L
\end{aligned} \tag{13}$$

Therefore, the consecutive application of these reciprocal maps should take any initial state back to itself, which means the following identity must hold for the interdependence matrices:

$$TS = ST = \mathbf{I} \tag{14}$$

in which \mathbf{I} is the identity matrix. However, the current construction of the interdependence matrices does not satisfy this property in general since the components are positive probability value that do not sum up to zero for the non-diagonal components of the TS and ST matrices. And the only instance where $S = T = \mathbf{I}$ is the trivial scenario of identical measurements.

To accommodate this issue, a different *probability measure* is needed for determining the probabilities in the mappings between various measurement types. A probability measure that is a continuous function, accepting non-positive inputs and yielding values in the interval $[0,1]$, with $P(0) = 0$ and $P(1) = 1$. These attributes lead to the probability measures in the form of:

$$P(\tilde{M}_j^K | \tilde{M}_i^L) = |\rho_{j,i}^{K,L}|^a, a \in \mathbb{R}^+ \tag{15}$$

defined based on the probability-intensities, $\rho_{j,i}$, of events. The probability-intensities in these probability measures can be negative or complex, as long as their resulting probabilities fall within the unit interval $[0,1]$. Among these probability measures, imposing the conservation of the total probability between the mappings leads to the probability measure with $a = 2$ (see Appendix A for proof). Therefore, in the final analysis, the only probability measure that results in consistent mappings between different measurements probability spaces is:

$$P(\tilde{M}_j^K | \tilde{M}_i^L) = |\rho_{j,i}^{K,L}|^2 \tag{16}$$

which defines probabilities based on probability-amplitudes.

Following the same procedure as before and conforming to the above probability measure, the revised interdependence matrices

$$\begin{aligned}
\Gamma_{j,i}^{L,K} &= \rho_{j,i}^{L,K} \\
\Delta_{j,i}^{K,L} &= \rho_{j,i}^{K,L}
\end{aligned} \tag{17}$$

have the following properties. Parallel to (12), for any fixed i the probabilities should add up to 1:

$$\sum_j |\Gamma_{j,i}|^2 = 1 \tag{18}$$

$$\sum_j |\Delta_{j,i}|^2 = 1$$

and parallel to (11), the components are confined according to:

$$\begin{aligned} 0 \leq |\Gamma_{j,i}|^2 &\leq 1 \\ 0 \leq |\Delta_{j,i}|^2 &\leq 1 \end{aligned} \tag{19}$$

which grants that the components can have negative or complex values. Analogous to the earlier construct, these interdependence matrices can be used to pursue the SM system's probabilities. However, instead of acting on the SM system probabilities, P_j^I , these mappings act on the SM system probability-amplitudes, σ_j^I , defined as:

$$P_j^I = |\sigma_j^I|^2 \tag{20}$$

with the following property:

$$\sum_j |\sigma_j^I|^2 = \sum_j P_j^I = 1 \tag{21}$$

With these adjustments, the adapted interdependence matrices can be used as before to determine the probabilities of the system for a second measurement. The interdependence matrices $\Gamma^{L,K}$ and $\Delta^{K,L}$ transform the SM system's probability-amplitudes between the two measurements \hat{L} and \hat{K} . $\Gamma^{L,K}$ maps the probability-amplitudes from measurement \hat{K} to measurement \hat{L} , and $\Delta^{K,L}$ maps them from measurement \hat{M}^L to measurement \hat{L} , according to:

$$\begin{aligned} \sigma_j^L &= \sum_i \Gamma_{j,i} \sigma_i^K \\ \sigma_j^K &= \sum_i \Delta_{j,i} \sigma_i^L \end{aligned} \tag{22}$$

The mappings should conserve the total probability, i.e.,

$$\sum_j |\sigma_j^L|^2 = \sum_j |\sigma_j^K|^2 = 1 \tag{23}$$

which indicates that Γ and Δ matrices are unitary. (The antiunitary case is not considered since these matrices describe continuous transformations between the probability-amplitudes; see also (18)).

The second property of these interdependence matrices stems from the fact that they are reciprocal mappings between the two measurement probability spaces, and their consecutive actions should map any state back to itself (cf. (14)):

$$\Gamma\Delta = \Delta\Gamma = I \quad (24)$$

This indicates that these unitary matrices are conjugate transposes of each other:

$$\Delta = \Gamma^{-1} = \Gamma^* \quad (25)$$

$$\Gamma = \Delta^{-1} = \Delta^*$$

In other words, the interdependence matrices between two measurements are conjugate transposes of each other, which means that for SM systems, the probability-amplitudes between any pair of measurement I and II outcomes are related as follows:

$$\rho(\tilde{M}_b^{II}|\tilde{M}_a^I) = \rho^*(\tilde{M}_a^I|\tilde{M}_b^{II}). \quad (26)$$

In summary, in the above analysis by investigating the constraints on transforming probabilities of SM systems between measurements, the characteristics of probability mappings in SM systems were derived. The standard probability measure used in classical physics was found to be inadequate for determining probabilities in SM system mappings. To address this, consistent probability measures were sought that could accommodate negative or complex probability-intensities while conserving total probability, leading to the probability measure based on probability-amplitudes and consequently the unitarity of the interdependence matrices. Throughout the analysis, the conservation of total probability in mappings was the primary constraint, and the transformation properties adhered to this consistency requirement.

State Vectors, Operator Algebra, Hilbert Space Representation, and the Probability Rule

The Hilbert-space formalism of SM systems theory is easily recognizable in the above construct, with clear denotations of its elements. The probability-amplitude of the SM system, denoted as σ^I , can be regarded as a vector of length 1 in the N-dimensional space defined by the N independent outcomes of the measurement \tilde{M}^I , namely $\tilde{M}_1^I, \tilde{M}_2^I, \dots, \tilde{M}_N^I$ (as in (1)). Employing the conventional bra-ket notation, the state-vector of the SM system can be expressed as $|\sigma^I\rangle = \sum_i \sigma_i^I |\tilde{M}_i^I\rangle$ with the following normalization from (2121):

$$\langle\sigma^I|\sigma^I\rangle = 1. \quad (27)$$

With this representation of the SM states as unit vectors in the complex vector space of the probability-amplitudes, the algebraic structure of SM systems is apparent. Once a measurement basis is chosen to represent the SM state, e.g., $|\sigma^I\rangle$, the probability-amplitude of the system for another measurement can be expressed as a linear combination of those bases, using the interdependence of the two measurements, according to (22) as:

$$|\sigma^{II}\rangle = \rho^{II,I}|\sigma^I\rangle \quad (28)$$

in which $\rho^{II,I}$ is the unitary transformation portrayed in (26). In other words, mappings of the state of SM systems between measurements are carried out by the interdependence matrices of the measurements. The logic is simple: the dependence of the states can be determined by the dependence of the measurements that would produce those states since SM states are defined by the outcome of measurements.

The Born rule for calculating probabilities is also clear in this representation. The probability-amplitude of a measurement outcome, \tilde{M}_b^{II} , given the initial state of the SM system, $\sigma^I = \tilde{M}_a^I$, determined from (28) is

$$\rho_{b,a}^{II,I} = \langle \sigma_a^I | \sigma_b^{II} \rangle = \langle \tilde{M}_a^I | \tilde{M}_b^{II} \rangle \quad (29)$$

and applying (16) leads to the probability of the described event as:

$$P_{b,a}^{II,I} = |\rho_{b,a}^{II,I}|^2 = |\langle \sigma_a^I | \sigma_b^{II} \rangle|^2 = |\langle \tilde{M}_a^I | \tilde{M}_b^{II} \rangle|^2 \quad (30)$$

This result is the Born probability rule; it is a built-in part of the theory rooted in the conservation of total probability in transformations.

In summary, the Hilbert space formalism of SM systems theory allows us to represent the state of the system as a unit vector in the complex vector space of probability-amplitudes. The state-vector can be expressed in terms of a chosen measurement basis and can be transformed between measurements probability spaces using the unitary interdependence matrices. The Born rule for calculating probabilities is naturally derived from the probability-amplitudes, and total probability is conserved in all transformations. Overall, this formalism provides a clear and powerful framework for understanding the behavior of SM systems.

Superposition of Possibilities and the Interference Effect

In the above formulation, the transformation of the system probability-amplitudes under a series of measurements can be described by consecutively applying the interdependence matrices of those measurements, as expressed in:

$$\begin{aligned} |\sigma^{III}\rangle &= \rho^{III,II} |\sigma^{II}\rangle = \rho^{III,II} \rho^{II,I} |\sigma^I\rangle \\ &= \rho^{III,I} |\sigma^I\rangle \end{aligned} \quad (31)$$

where

$$\rho_{f,i}^{III,I} = \sum_n \rho_{f,n}^{III,II} \rho_{n,i}^{II,I} \quad (32)$$

describes the relationship between the interdependence matrices of the measurement.

This chain rule allows the determination of the system's probability-amplitude for a measurement based on the interdependencies of measurements. Importantly, this sum accounts for the interference effects: the probability of events, calculated according to:

$$P(\tilde{M}_f^{III}|\tilde{M}_i^I) = P_{f,i}^{III,I} = |\rho_{f,i}^{III,I}|^2 = \left| \sum_n \rho_{f,n}^{III,II} \rho_{n,i}^{II,I} \right|^2 \quad (33)$$

can include extra terms, “the interference terms,” that lead to results different from what the classical method of calculating probabilities predicts. For example, using the above relation, the interference in the double-slit experiment follows as the photons have two options (slit1 and slit2) to go from the source (O) to the screen (S):

$$\begin{aligned} P(\tilde{M}_S^{\text{Screen}}|\tilde{M}_O^{\text{Source}}) &= |\rho_{S,O}^{\text{Screen,Source}}|^2 \\ &= |\rho_{S,\text{slit1}}^{\text{Screen,Slit}} \rho_{\text{slit1},O}^{\text{Slit,Source}} + \rho_{S,\text{slit2}}^{\text{Screen,Slit}} \rho_{\text{slit2},O}^{\text{Slit,Source}}|^2 \end{aligned} \quad (34)$$

This differs from the classical method of calculating probabilities, which predicts:

$$\begin{aligned} P(\tilde{M}_S^{\text{Screen}}|\tilde{M}_O^{\text{Source}}) &= P(S|\text{slit1})P(\text{slit1}|O) + P(S|\text{slit2})P(\text{slit2}|O) \\ &\equiv |\rho_{S,\text{slit1}}^{\text{Screen,Slit}} \rho_{\text{slit1},O}^{\text{Slit,Source}}|^2 + |\rho_{S,\text{slit2}}^{\text{Screen,Slit}} \rho_{\text{slit2},O}^{\text{Slit,Source}}|^2 \end{aligned} \quad (35)$$

The transformation of the probability-amplitudes between the measurements, as described by the chain rule in (31), highlights interference as a fundamental characteristic of SM systems. It can be seen that it is the connection between the probability-amplitudes of the outcomes of different measurements, rather than their probabilities, that modifies the result from classical probability. The above derivation provides a clear understanding of the basis of interference and demonstrates the essential role of the presence of the intermediate measurement. In the physics of SM systems, interference results from the superposition of possible outcomes of the intermediate measurement, indicating that the presence of intermediate non-performed measurements (the “interaction-free” measurements) cannot be neglected. It is also evident that interference results from the mathematical record-keeping of SM systems’ probability-amplitudes for possible outcomes of intermediate measurements rather than the physical occurrence of those outcomes.

Time Evolution and Derivation of the Schrödinger Equation

The analysis of the mappings between the probability spaces of different measurements for SM systems led to the algebraic structure of their state-space, specifically, the Hilbert space and the Born probability rule. Incorporating time evolution in this framework is a straightforward process, as discussed in various mathematical physics textbooks (see for example [30] Sec. 3.3).

A measurement can be labeled by the time variable t , denoting the time at which it is performed. Since the relationship between measurements does not depend on time, their interdependence matrices must maintain the same structure at different times; hence, there must exist a unitary transformation $\mathbb{T}(t_2 - t_1)$ such that:

$$\rho(t_2) = \mathbb{T}^{-1}(t_2 - t_1) \rho(t_1) \mathbb{T}(t_2 - t_1) \quad (36)$$

Under common assumptions about time evolution, the transformation can be written as:

$$\mathbb{T}(t_2 - t_1) = e^{-iH(t_2-t_1)} \quad (37)$$

where H is a self-adjoint matrix that defines the Hamiltonian in the Hilbert space. Borrowing conventional quantum terminology, thus far, our discussion was in the “Heisenberg picture,” in which the states of isolated systems remain fixed, but the interdependence matrices that represent the observables change with time. Transposing to the “Schrödinger picture” shifts the focus to the time evolution of the state, and the Schrödinger equation is obtained as follows:

$$i \frac{d}{dt} |\sigma^I(t)\rangle = H |\sigma^I(t)\rangle \quad (38)$$

Here, we have shown how the analysis of the constraints imposed by the probabilistic nature of SM systems under measurements leads to the derivation of the standard Hilbert-space formalism of quantum mechanics, as well as the Born probability rule. The full framework of quantum theory emerged from studying the general properties of the mappings that transform the state of the SM system from one measurement probability-space to another. Using this perspective, the elements of the theory can be properly understood.

The state-vector, σ^I , contains information about the outcome of the last measurement performed on the system. The interdependence matrices, $\rho^{I,I'}$, are unitary matrices that transform the system state-vector, σ^I , between different measurements’ probability spaces. These matrices embody information on how the outcomes of distinct types of measurements correlate with one another probabilistically. Accordingly, the state-vector can be mapped into different measurements’ probability spaces via the interdependencies of the measurements. The transformed state-vectors represent the probability-amplitudes of the system for those measurements. The construct of the theory ensures that the total probability is conserved in these transformations. Finally, when a measurement is performed, the state of the system adjusts accordingly to reflect the result of the observation.

4 Discussion and Conclusion

In this report, we have presented a systematic derivation of the standard formalism of quantum theory from a physical foundation. The underlying physical idea in our derivation is the recognition of physical systems with a single adjustable variable and their inherently probabilistic nature due to their limited capacity to carry messages. This is whence the indeterministic nature of quantum mechanics arises. Based on this physical foundation, we have derived quantum theory in a transparent and intuitive manner. Similar postulates have been suggested in the past [12, 14], however, they did not succeed in deriving the full formalism of the theory.

This derivation shows that quantum mechanics describes the physics of systems that possess only a single adjustable variable. The theory describes the transformations of the probability-amplitudes of such systems for different measurements, with the algebraic structure of the theory rooted in the conservation of total probability in these transformations. The current derivation of the quantum formalism for finite-dimensional individual systems can easily be extended to the general case. It is worth noting that this derivation places the ensemble interpretation of quantum mechanics [31] as a secondary interpretation of the theory, rather than its primary one.

In addition, this derivation clarifies the fundamental concept of state in quantum theory. The state of a quantum system is a mathematical representation of the physical state of the system, determined by the outcome of the most recent measurement performed on it. The theory provides a way to calculate the system's probabilities for future measurement outcomes from its current state, taking into account the interdependencies of the measurements. This perspective resolves misconceptions about the state of a quantum system, namely that it stores all information about the system's representations in various measurements [32]. Rather, this information is contained in the interdependencies of measurements, not in the quantum system itself.

The interpretation of quantum mechanics becomes clear from our derivation. The foundation of quantum theory lies in the concept that a quantum system is a physical system with no more than one independent adjustable variable. Since such systems can only contain one proposition of objective reality, the theory is inherently probabilistic rather than deterministic. At its essence, quantum theory is a mathematical framework for calculating the probabilities of a measurement's outcomes. It describes how the probabilities of single-variable systems transform among different measurements probability spaces. Rather than describing the measurement process, the theory focuses on what can be known about the potential results of measurements.

Our work does not change the existing formalism of quantum theory. Instead, it provides a comprehensive framework for interpreting quantum phenomena and represents a significant step towards a deeper understanding of the theory. Quantum mechanics is the probability theory for physical systems that possess a single adjustable variable. The core mathematical structure of the theory is based on consistent record-keeping of probabilities between different measurements. Understanding the physical foundation of quantum theory allows us to revisit the phenomena described by conventional quantum mechanics and gain deeper insights into the nature of our world.

References

- [1] P. A. M. Dirac, *The principles of quantum mechanics* (no. 27). Oxford university press, 1981.
- [2] J. Von Neumann, *Mathematical foundations of quantum mechanics*. Princeton university press, 2018.
- [3] G. Birkhoff and J. Von Neumann, "The logic of quantum mechanics," *Annals of mathematics*, pp. 823-843, 1936.
- [4] G. W. Mackey, *Mathematical foundations of quantum mechanics*. Courier Corporation, 2013.

- [5] C. Piron, "Axiomatique quantique," *Helvetica physica acta*, vol. 37, no. 4-5, pp. 439-&, 1964.
- [6] C. Piron, "Survey of general quantum physics," *Foundations of physics*, vol. 2, no. 4, pp. 287-314, 1972.
- [7] G. Ludwig, "Attempt of an axiomatic foundation of quantum mechanics and more general theories. III," *Communications in Mathematical Physics*, vol. 9, no. 1, pp. 1-12, 1968.
- [8] B. Mielnik, "Geometry of quantum states," *Communications in Mathematical Physics*, vol. 9, no. 1, pp. 55-80, 1968.
- [9] A. Landé, "Albert Einstein and the quantum riddle," *American Journal of Physics*, vol. 42, no. 6, pp. 459-464, 1974.
- [10] D. I. Fivel, "How interference effects in mixtures determine the rules of quantum mechanics," *Physical Review A*, vol. 50, no. 3, p. 2108, 1994.
- [11] L. Accardi, "Can mathematics help solving the interpretational problems of quantum theory?," *Il Nuovo Cimento B (1971-1996)*, vol. 110, no. 5, pp. 685-721, 1995.
- [12] C. Rovelli, "Relational quantum mechanics," *International Journal of Theoretical Physics*, vol. 35, no. 8, pp. 1637-1678, 1996.
- [13] N. Landsman, "Simple new axioms for quantum mechanics," *International journal of theoretical physics*, vol. 37, no. 1, pp. 343-348, 1998.
- [14] A. Zeilinger, "A foundational principle for quantum mechanics," *Foundations of Physics*, vol. 29, no. 4, pp. 631-643, 1999.
- [15] B. Coecke, D. Moore, and A. Wilce, "Current research in operational quantum logic," in *Series Fundamental Theories of Physics*, vol. 111: Springer, 2000.
- [16] L. Hardy, "Quantum theory from five reasonable axioms," *arXiv preprint quant-ph/0101012*, 2001.
- [17] R. Clifton, J. Bub, and H. Halvorson, "Characterizing quantum theory in terms of information-theoretic constraints," *Foundations of Physics*, vol. 33, no. 11, pp. 1561-1591, 2003.
- [18] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of quantum theory," *Physical Review A*, vol. 84, no. 1, p. 012311, 2011.
- [19] L. Masanes and M. P. Müller, "A derivation of quantum theory from physical requirements," *New Journal of Physics*, vol. 13, no. 6, p. 063001, 2011.
- [20] H. F. Hofmann, "Derivation of quantum mechanics from a single fundamental modification of the relations between physical properties," *Physical Review A*, vol. 89, no. 4, p. 042115, 2014.
- [21] J. A. Wheeler, "World as system self-synthesized by quantum networking," in *Probability in the Sciences*: Springer, 1988, pp. 103-129.
- [22] J. A. Wheeler, "Toward "it from bit"," *Quantum Coherence*, p. 281, 1992.
- [23] J. A. Wheeler, "Information, physics, quantum: the search for links," in *Feynman and computation: exploring the limits of computers*, 1999, pp. 309-336.
- [24] C. Brukner and A. Zeilinger, "Malus' law and quantum information," *acta physica slovacca*, vol. 49, pp. 647-652, 1999.
- [25] C. Brukner, M. Zukowski, and A. Zeilinger, "The essence of entanglement," *arXiv preprint quant-ph/0106119*, 2001.
- [26] C. Brukner and A. Zeilinger, "Young's experiment and the finiteness of information," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 360, no. 1794, pp. 1061-1069, 2002.

- [27] Č. Brukner and A. Zeilinger, "Information and fundamental elements of the structure of quantum theory," in *Time, quantum and information*: Springer, 2003, pp. 323-354.
- [28] Č. Brukner, M. Aspelmeyer, and A. Zeilinger, "Complementarity and information in "delayed-choice for entanglement swapping"," *Foundations of Physics*, vol. 35, no. 11, pp. 1909-1919, 2005.
- [29] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656-715, 1949.
- [30] W. Thirring and E. M. Harrell, *A course in mathematical physics 3: Quantum mechanics of atoms and molecules*. Springer, 1981.
- [31] L. E. Ballentine, "The statistical interpretation of quantum mechanics," *Reviews of modern physics*, vol. 42, no. 4, p. 358, 1970.
- [32] C. M. Caves and C. A. Fuchs, "Quantum information: How much information in a state vector?," *arXiv preprint quant-ph/9601025*, 1996.

Appendix A: Deriving the Probability Measure that Conserves Total Probability in Mappings

To determine the appropriate probability measure for transforming the SM system probabilities via interdependence matrices, we obtained probability measures in the form of (15): $P(\tilde{M}_j^K | \tilde{M}_i^L) = |\rho_{j,i}^{K,L}|^\alpha$, which allow for non-positive probability-intensity values as inputs. To find the probability measures that transform the probabilities consistently between mappings, we first consider those in the form of:

$$P(\tilde{M}_j^K | \tilde{M}_i^L) = (\rho_{j,i}^{K,L})^{2n}, n \in \mathbb{N} \quad (39)$$

The interdependence matrices can be constructed in accord to these probability measures as:

$$\begin{aligned} \Gamma_{j,i}^{L,K} &= \rho_{j,i}^{L,K} \\ \Delta_{j,i}^{K,L} &= \rho_{j,i}^{K,L} \end{aligned} \quad (40)$$

These matrices are subject to the constraints:

$$\begin{aligned} \sum_j (\Gamma_{j,i})^{2n} &= 1 \\ \sum_j (\Delta_{j,i})^{2n} &= 1 \end{aligned} \quad (41)$$

based on the logic that a measurement ultimately results in one outcome, and the probabilities should add up to 1 (cf. (7)). Additionally, the components of these matrices, akin to (11) are confined within the limits:

$$0 \leq (\Gamma_{j,i})^{2n} \leq 1 \quad (42)$$

$$0 \leq (\Delta_{j,i})^{2n} \leq 1$$

which allows for the components to have negative or complex values.

To calculate the probabilities of the SM system, we accordingly define the probability-intensities of the SM system, σ_j^l , in terms of its probabilities, P_j^l , as:

$$P_j^l = (\sigma_j^l)^{2n} \quad (43)$$

where σ_j^l represents the probability-intensity of the SM system for the outcome \tilde{M}_j^l in the measurement \hat{M}^l . The total sum of the system probabilities must equal 1, hence for the probability-intensities we get:

$$\sum_j (\sigma_j^l)^{2n} = \sum_j P_j^l = 1 \quad (44)$$

The interdependence matrices $\Gamma^{L,K}$ and $\Delta^{K,L}$ map the probability-intensities of the system between the two measurements \hat{M}^L and \hat{M}^K . $\Gamma^{L,K}$ maps the probability-intensities from measurement \hat{M}^K to measurement \hat{M}^L , and $\Delta^{K,L}$ maps them from measurement \hat{M}^L to measurement \hat{M}^K , according to:

$$\sigma_j^l = \sum_i \Gamma_{j,i} \sigma_i^k \quad (45)$$

$$\sigma_j^k = \sum_i \Delta_{j,i} \sigma_i^l$$

These mappings should conserve the total probability, i.e.,

$$\sum_j (\sigma_j^l)^{2n} = \sum_j (\sigma_j^k)^{2n} = 1 \quad (46)$$

This is a necessary condition for the consistency of the mappings between measurements and it specifies the value of n required for the consistent probability measure, as will be shown below. Substituting the probability-intensities from (45) returns:

$$\begin{aligned} \sum_j (\sigma_j^l)^{2n} &= \sum_j \left(\sum_i \Gamma_{j,i} \sigma_i^k \right)^{2n} \\ &= \sum_j (\sigma_j^k)^{2n} \end{aligned} \quad (47)$$

which means we should have:

$$\begin{aligned}
\sum_j \left(\sum_i \Gamma_{j,i} \sigma_i^K \right)^2 &= \sum_j \left(\sum_m \Gamma_{j,m}^* \sigma_m^{*K} \right) \left(\sum_l \Gamma_{j,l} \sigma_l^K \right) \\
&= \sum_{m,l} \sigma_m^{*K} \sigma_l^K \sum_j \Gamma_{j,m}^* \Gamma_{j,l} \\
&= \sum_j (\sigma_j^K)^2
\end{aligned} \tag{48}$$

The above relation holds if:

$$\sum_j \Gamma_{j,m}^* \Gamma_{j,l} = \delta_{m,l} \tag{49}$$

Summing over m for a fixed l leads to:

$$\begin{aligned}
\sum_m \sum_j \Gamma_{j,m}^* \Gamma_{j,l} &= \sum_j \Gamma_{j,1}^* \Gamma_{j,l} + \sum_j \Gamma_{j,2}^* \Gamma_{j,l} + \cdots + \sum_j \Gamma_{j,N}^* \Gamma_{j,l} \\
&= \delta_{1,l} + \delta_{2,l} + \cdots + \sum_j \Gamma_{j,l}^* \Gamma_{j,l} + \cdots + \delta_{1,N} \\
&= \sum_j \Gamma_{j,l}^* \Gamma_{j,l} = \sum_j (\Gamma_{j,l})^2 \\
&= \sum_m \delta_{m,l} = 1
\end{aligned} \tag{50}$$

which compared to (41) gives $n = 1$.

To complete this analysis of the proper probability measures, in addition to the ones defined in (39), we should consider all the other possible probability measures that allow non-positive components, i.e., $P(\tilde{M}_j^K | \tilde{M}_i^L) = |\rho_{j,i}^{K,L}|^a$, $a \in \mathbb{R}^+$. For even values of a , these measures are the same as the ones already discussed above. For other values of a , a relationship similar to (47) results in $\Gamma_{j,i} = \pm \delta_{j,i}$ indicating that these *measures* only conserve probabilities in the mappings for the trivial case of identical measurements.

Therefore, the only consistent probability measure that conserves the total probability in the mappings for general cases of measurements is:

$$P(\tilde{M}_j^K | \tilde{M}_i^L) = (\rho_{j,i}^{K,L})^2 \tag{51}$$

which defines the probabilities based on probability-amplitudes.