

Derivation of quantum theory from Feynman's rules

Philip Goyal*

University at Albany (SUNY), Albany, New York 12222, USA

(Received 18 November 2013; published 20 March 2014)

Feynman's formulation of quantum theory is remarkable in its combination of formal simplicity and computational power. However, as a formulation of the abstract structure of quantum theory, it is incomplete as it does not account for most of the fundamental mathematical structure of the standard von Neumann–Dirac formalism such as the unitary evolution of quantum states. In this paper, we show how to reconstruct the entirety of the finite-dimensional quantum formalism starting from Feynman's rules with the aid of a single physical postulate, the *no-disturbance postulate*. This postulate states that a particular class of measurements have no effect on the outcome probabilities of subsequent measurements performed. We also show how it is possible to derive both the amplitude rule for composite systems of distinguishable subsystems and Dirac's amplitude-action rule, each from a single elementary and natural assumption, by making use of the fact that these assumptions must be consistent with Feynman's rules.

DOI: [10.1103/PhysRevA.89.032120](https://doi.org/10.1103/PhysRevA.89.032120)

PACS number(s): 03.65.Ca

I. INTRODUCTION

The mathematical formalism of quantum theory has numerous structural features, such as its use of complex numbers, whose physical basis has long been regarded as obscure. In recent years, there has been a growing interest in deriving these features from compelling physical principles inspired by an informational perspective on physical processes [1–4]. The purpose of such derivation is to better understand the differences between quantum and classical physics, to establish the range of validity of the various parts of the quantum formalism, and to identify physical principles whose validity might extend beyond quantum theory itself. Substantial progress has now been made, both in deriving much of the quantum formalism from physical principles [5–13], and in identifying physical principles that account for some of the nonclassical features of quantum physics [14,15].

Almost without exception, the above-mentioned attempts to understand the quantum formalism have focused their attention on the standard Dirac–von Neumann formalism. However, Feynman's formulation of quantum theory provides a strikingly different representation of quantum physics [16,17], and this raises the important question of whether we may be able to gain valuable insights by deriving quantum theory from Feynman's perspective.

Perhaps the most remarkable feature of Feynman's formulation is its formal simplicity. The simplicity is achieved by dispensing with the notion of the state of a system and with operators that represent measurements and temporal evolution or symmetry transformations. Instead, the primary notion is that of a *transition* of a system from one measurement outcome, obtained at some time, to another measurement outcome, obtained at some other time; and a complex-valued *amplitude* is associated with each transition. Feynman's abstract formalism for individual systems consists of what we shall refer to as *Feynman's rules* [16]. These rules, summarized in Fig. 3, stipulate how amplitudes associated with given transitions of a system are combined to yield amplitudes of

more complex transitions of that system, and how probabilities are computed from amplitudes.

As Feynman observed, the rules for combining amplitudes bear a striking resemblance to the rules of probability theory [16,17]. In previous work [10], we seized on this observation to *derive* Feynman's rules using a method similar to that used by Cox to derive the rules of probability theory from Boolean algebra [18,19]. Our derivation provides a particularly clear understanding of why complex numbers are such a fundamental part of the mathematical structure of quantum theory, and provides a precise understanding of the relationship between Feynman's rules and probability theory [20].

Feynman's rules, however, do not, by themselves, constitute a complete formulation of quantum theory. Most importantly, they do not account for most of the fundamental mathematical structure of the standard von Neumann–Dirac formalism [21,22]. For example, a basic property of the standard formalism is that state evolution is unitary, but this property does not follow from Feynman's rules. While it is true that Feynman's rules imply unitarity *given* the form of the classical action and Dirac's amplitude-action rule [16,23], unitarity does not follow as a *direct* consequence of Feynman's rules alone, a problem of which Feynman was aware [24]. Not only is this unsatisfactory on a theoretical level, it is particularly problematic since a corresponding classical action does not always exist for a quantum system [17], and even when one does exist, a mathematically rigorous proof of unitarity on the basis of Feynman's path integral, assuming a general form for the action, is highly nontrivial [23].

Even more fundamentally, the notion of a quantum state itself does not follow naturally from Feynman's rules. Contrary to what is asserted in Refs. [16,17], one cannot simply assume that the state of a system consists of the amplitudes to obtain the possible outcomes of a given measurement irrespective of the prior history of the system, for two reasons. First, one cannot exclude the possibility that the system is entangled with another system, and is therefore in a mixed state rather than a pure state. Second, before these amplitudes can be declared to constitute the state of the system, one must establish that these amplitudes suffice to compute the outcome probabilities of not

*pgoyal@albany.edu

just the given measurement but of *any* measurement that could be performed on the system.

Thus, in order to complete the derivation of quantum theory from the Feynman perspective, it is essential to discover what physical ideas are needed in order to derive the standard quantum formalism given Feynman's rules for individual systems. In this paper, we show that a *single* physical idea, formalized in the *no-disturbance postulate*, suffices. This postulate states that a particular class of measurements—which we refer to as *trivial* measurements—have no effect on the outcome probabilities of subsequent measurements. A trivial measurement has a *single* outcome, this outcome having been obtained by coarse-graining over all of the outcomes of an atomic, repeatable measurement (see Sec. II A for definitions of these terms). For example, a Stern-Gerlach measurement with but a single detector which registers all outgoing systems is a trivial measurement.

The no-disturbance postulate formalizes a key difference between classical and quantum physics. In quantum physics, a trivial measurement is nondisturbing. However, in classical physics, a trivial measurement is generally disturbing. For example, consider a trivial Stern-Gerlach measurement with a single coarse-grained outcome obtained by coarse-graining over two atomic outcomes. From a classical point of view, it is a fact of the matter that the system passed through one of the atomic outcomes, *even though the coarse-grained outcome was not capable of registering this fact*. As we show in Sec. III A, this classical inference leads to a change in the outcome probabilities of subsequent measurements.

One can understand the no-disturbance postulate quite naturally as follows. From an information-theoretic point of view, it is the gain of information about a quantum system that is ultimately responsible for the disturbance of its state [25]. From this viewpoint, it seems eminently plausible that, conversely, there should exist measurements that provide no useful information which do not disturb the state [26]. Informally, one might say that, if a measurement provides no information, then it need not disturb the state of the system [27]. Since a trivial measurement has only one outcome, one gains no information (in the sense of Shannon information) about which outcome was obtained on learning the outcome of the measurement. This is similar to one's predicament on learning that a two-headed coin has landed heads. The no-disturbance postulate asserts that trivial measurements are such nondisturbing, noninformative measurements.

Since the no-disturbance postulate exposes a fundamental difference between classical and quantum physics, it is an excellent candidate to take as a physical postulate in deriving quantum theory. Indeed, we have previously employed a special case of this postulate in our derivation of Feynman's rules [10], and a very similar idea has more recently been employed in Ref. [28] to derive interesting results regarding the state space of general probabilistic theories.

As mentioned above, Feynman's rules concern a given quantum system, so that additional rules must be given if one wishes to treat composite systems of distinguishable or indistinguishable subsystems. A particularly attractive feature of these rules is their formal simplicity. In Ref. [17], the amplitude rules for such composite systems are simply postulated, presumably having been extracted from the

standard formalism. In this paper, we show that the rule for composite systems of distinguishable subsystems can in fact be *derived* from a simple *composition postulate* on the condition that Feynman's rules for individual systems are valid. The composition postulate simply posits that the amplitude of a transition of a composite system consisting of two noninteracting subsystems is a continuous function of the amplitudes of the respective subsystem transitions. Remarkably, the rule for assigning amplitudes to composite systems is uniquely determined by the requirement that the amplitude assignment to the composite system be consistent with Feynman's rules for assigning amplitudes to individual systems. This rule immediately gives rise to the tensor product rule in the standard quantum formalism. The derivation of the symmetrization postulate, which is needed to describe composite systems consisting of indistinguishable subsystems, is detailed elsewhere [29].

Finally, a striking feature of Feynman's formulation is its remarkably direct connection to the Lagrangian formulation of classical physics. More precisely, when a series of transitions of a quantum system in configuration space is well approximated by a continuous trajectory of the corresponding classical system, Dirac's amplitude-action rule associates the amplitude $e^{iS/\hbar}$ to the series of transitions, where S is the action of the corresponding classical trajectory [30]. It is this rule from which much of the computational power of Feynman's rules derives, allowing, for example, the derivation of the Schrödinger equation [16] and quantum electrodynamics [31]. Dirac obtained this rule on the basis of a detailed analogy between transformations in quantum theory and contact transformations in the Lagrangian formulation of classical mechanics. However, given the fundamental importance of this classical-quantum connection, a simpler and more direct derivation is desirable. Here, using only elementary properties of the action, we provide a simple and direct derivation of this rule on the simple assumption that the amplitude of a path in configuration space is a continuous function of its classical action.

The work described here significantly improves upon previous attempts to address many of the above-mentioned issues. In particular, a previous attempt to derive unitarity from Feynman's rules makes appeal to the Hilbert space norm, which itself must be independently justified [32]. In contrast, our derivation rests entirely on the no-disturbance postulate. Similarly, two previous derivations of the rule for composite systems of distinguishable subsystems implicitly assume that the functional relationships involved are complex analytic, an assumption which significantly detracts from the physical transparency of the derivations [33,34]. We are able to avoid such assumptions.

The results given here have implications for many issues which have been raised in the literature. For example, a recurrent question is whether linear temporal evolution can be replaced with nonlinear temporal evolution [35–39]. Such replacement has a variety of motivations, such as the desire to incorporate the quantum measurement process into the usual temporal evolution of a quantum state, or to solve the black hole information paradox [40]. Such work has led to attempts to explore whether, taking certain other parts of the quantum formalism as a given, one can derive linearity or unitarity

from a physical principle such as the requirement that there is no instantaneous signaling between separated subsystems [41–43]. Our earlier derivation of Feynman's rules, together with the present derivation of unitarity from the no-disturbance postulate, shows that the linearity and unitarity of temporal evolution is very basic to the structure of quantum theory, and cannot be replaced with nonlinear deterministic evolution, even in a manner that is barely experimentally perceptible, without undermining the entire edifice. Furthermore, since our derivation of unitarity depends only on no-disturbance, and not on any postulate that refers to the behavior (such as no instantaneous signaling) of composite systems, our result shows that such behavior, contrary to that which is suggested by some previous work [41–43], is not necessary to understand unitarity.

The remainder of the paper is organized as follows. In Sec. II, we summarize the experimental framework and notation presented in Ref. [10], extend the framework to deal with composite systems, and summarize Feynman's rules in operational language. In Sec. III, we formulate the no-disturbance postulate, and use this to systematically introduce the notion of a quantum state, show that states evolve unitarily, and show that repeatable measurements are represented by Hermitian operators. In Sec. IV, we introduce a composition operator, formulate its symmetry relations, and show, via the composition postulate, that these lead to the amplitude rule for composite systems. Finally, in Sec. V, we derive Dirac's amplitude-action rule. We conclude in Sec. VI with a discussion of the results and their broader implications.

II. BACKGROUND AND NOTATION

A. Experimental framework

An experimental setup is defined by specifying a source of physical systems, a sequence of measurements to be performed on a system in each run of the experiment, and the interactions between the system and its environment which occur between the measurements (see the example in Fig. 1). In a run of an experiment, a physical system from the source passes through a sequence of measurements $\mathbf{L}, \mathbf{M}, \mathbf{N}, \dots$, which

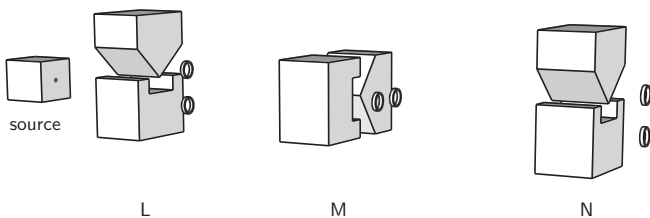


FIG. 1. Schematic representation of a Stern-Gerlach experiment performed on silver atoms. A silver atom from a source (an evaporator) is subject to a sequence of measurements, each of which yields one of two possible outcomes registered by nonabsorbing wire-loop detectors. Between measurements, the atoms interact with a uniform magnetic field. A run of the experiment yields outcomes ℓ, m, n of the measurements $\mathbf{L}, \mathbf{M}, \mathbf{N}$ performed at times t_1, t_2, t_3 , respectively. The probability distribution over the outcomes of \mathbf{M} given an outcome of \mathbf{L} is observed to be independent of any interactions the system had prior to \mathbf{L} , a property to which we refer as closure.

respectively yield outcomes ℓ, m, n, \dots at times t_1, t_2, t_3, \dots . These outcomes are summarized in the *measurement sequence* $[\ell, m, n, \dots]$. In between these measurements, the system may undergo interactions with the environment. We shall label the possible outcomes of a measurement \mathbf{M} as m, m', m'', m''', \dots as far as is needed in each case.

Over many runs of the experiment, the experimenter will observe the frequencies of the various possible measurement sequences, from which the experimenter can estimate the probability associated with each sequence. The probability $P(A)$ associated with sequence $A = [\ell, m, n, \dots]$ is defined as the probability of obtaining outcomes m, n, \dots conditional upon obtaining ℓ ,

$$P(A) = \Pr(m, n, \dots | \ell). \quad (1)$$

A particular outcome of a measurement is either *atomic* or *coarse-grained*. An atomic outcome is one that cannot be more finely divided in the sense that the detector whose output corresponds to the outcome cannot be subdivided into smaller detectors whose outputs correspond to two or more outcomes. An example of atomic outcomes are the two possible outcomes of a Stern-Gerlach measurement performed on a silver atom. A coarse-grained outcome is one that does not differentiate between two or more outcomes, an example being a Stern-Gerlach measurement where a detector's field of sensitivity encompasses the fields of sensitivity of two detectors, each of which corresponds to a different atomic outcome. Abstractly, if a measurement has an outcome which is a coarse-graining of the outcomes labeled m and m' of measurement \mathbf{M} , the outcome is labeled (m, m') , and this notational convention naturally extends to coarse-graining of more than two atomic outcomes. In general, if all of the possible outcomes of a measurement are atomic, we shall call the measurement itself atomic. Otherwise, we say it is a coarse-grained measurement. A coarse-grained measurement with but a single outcome is called a *trivial* measurement as its outcome provides us with no more information than the fact that the measurement has detected a system at a particular time.

As explained in Refs. [10,44], the measurements and interactions which can be employed in a given experiment must satisfy certain conditions if they are to lead to a well-defined theoretical model. The measurements that are employed in an experimental setup must be repeatable, come from the same *measurement set* \mathcal{M} , or be coarsened versions of measurements drawn from this set; and the first measurement in each experiment must be atomic. These conditions ensure that (i) all the measurements are probing the same aspect (say, the spin behavior) of the system, and (ii) the outcome probabilities of all measurements except the first are independent of the history of the system prior to the experiment, a property we refer to as *closure*. Similarly, interactions that occur in the period of time between measurements are selected from a set \mathcal{I} of possible interactions. These interactions preserve closure when they act on the given system between any two measurements in \mathcal{M} . For the operational definitions of sets \mathcal{M} and \mathcal{I} , and further discussion of the above conditions, we refer the reader to Ref. [10].

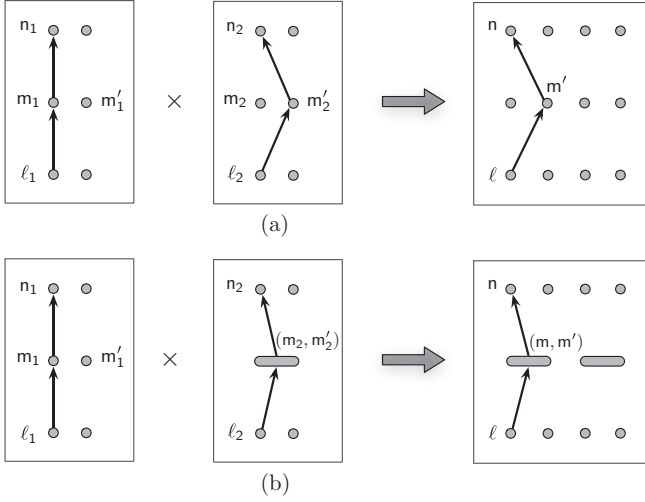


FIG. 2. Composition of sequences belonging to two separate systems. In both examples, the first system undergoes measurements $\mathbf{L}_1, \mathbf{M}_1, \mathbf{N}_1$. The second system undergoes $\mathbf{L}_2, \mathbf{M}_2, \mathbf{N}_2$ in (a) and $\mathbf{L}_2, \mathbf{M}_2, \mathbf{N}_2$ in (b). For concreteness, the atomic measurements are assumed to have two possible outcomes, with m_1, m'_1 labeling the possible outcomes of \mathbf{M}_1 , and so on. In (a), the first system yields sequence $A = [\ell_1, m_1, n_1]$, and the second system yields $B = [\ell_2, m'_2, n_2]$, which are composed (\times) to yield sequence $C = [\ell, m', n] = [(\ell_1 : \ell_2), (m_1 : m'_2), (n_1 : n_2)]$. In (b), the first system yields $A = [\ell_1, m_1, n_1]$, and the second system yields $B = [\ell_2, (m_2, m'_2), n_2]$, which compose to yield $C = [\ell, (m, m'), n] = [(\ell_1 : \ell_2), ((m_1 : m_2), (m_1 : m'_2)), (n_1 : n_2)]$.

Composite systems

Suppose that two distinguishable physical systems, S_1 and S_2 , simultaneously undergo experiments. Operationally, distinguishability means that there is some measurement which can be performed on each system before and after an experiment which determines the identity of the system passing through the experiment. In particular, suppose that system S_1 undergoes an experiment where measurements $\mathbf{L}_1, \mathbf{M}_1, \mathbf{N}_1$ are performed at successive times t_1, t_2, t_3 , while the other system, S_2 , undergoes a separate experiment where measurements $\mathbf{L}_2, \mathbf{M}_2, \mathbf{N}_2$ are performed at these times.

These two experiments can equivalently be viewed as a *single* experiment performed on a system S undergoing a sequence of measurements $\mathbf{L}, \mathbf{M}, \mathbf{N}$, where each of these measurements is viewed as a composite of the corresponding indexed measurements; for example, \mathbf{L} is viewed as a composite of \mathbf{L}_1 and \mathbf{L}_2 . We then say that S is a *composite* system consisting of subsystems S_1 and S_2 . If, say, measurement \mathbf{L}_1 yields outcome l_1 , and measurement \mathbf{L}_2 yields l_2 , this can be described as measurement \mathbf{L} yielding an outcome ℓ which we symbolize as $(\ell_1 : \ell_2)$. The process of composition is illustrated in Fig. 2.

In order that the experiments on S_1 and on S_2 separately satisfy the experimental conditions stated above, the systems cannot be interacting with one another while the experiments are in progress. If the systems are interacting with one another, then only the two experiments viewed as a whole—as an experiment on S —can satisfy these conditions.

B. Operationalization of Feynman's paths and Feynman's rules

Consider an experimental setup in which a physical system is subject to successive measurements $\mathbf{L}, \mathbf{M}, \mathbf{N}$ at successive times t_1, t_2, t_3 , with there possibly being interactions with the system in the intervals between those measurements. Here and subsequently, we assume that the measurements and interactions in any such setup are selected according to the constraints described above. We summarize the outcomes obtained in a given run of the experiment as the *sequence* $C = [\ell, m, n]$. This is the operational counterpart to a Feynman “path.”

We now wish to formalize the idea that setups are interrelated in particular ways. In Ref. [10], we considered two such relationships. First, the above setup could be viewed as a *series concatenation* of two experiments, the first in which measurements \mathbf{L} and \mathbf{M} occur at times t_1 and t_2 , yielding the sequence $A = [\ell, m]$, and the second in which measurements \mathbf{M} and \mathbf{N} occur at times t_2 and t_3 , yielding $B = [m, n]$. In order to ensure that experimental closure is satisfied in the second experiment, measurement \mathbf{M} must be atomic. Formally, we express this concatenation as

$$C = A \cdot B, \quad (2)$$

where “ \cdot ” is the *series* combination operator. More generally, the series operator can be used to concatenate two sequences provided their initial and final measurements are atomic, and the final measurement and outcome of the first sequence is the same as the initial measurement and outcome of the second sequence.

Second, one can consider a setup which is identical to the one above, except that outcomes m and m' of \mathbf{M} have been coarse-grained, so that one obtains the sequence $E = [\ell, (m, m'), n]$. Formally, we express the relationship of this sequence to the sequences $C = [\ell, m, n]$ and $D = [\ell, m', n]$ as

$$E = C \vee D, \quad (3)$$

where \vee is the *parallel* combination operator. More generally, the parallel operator can combine any two sequences which are identical except for differing in the outcome of a single measurement in the setup, provided that this measurement is not the initial or final measurement.

Feynman's rules in operational form

From the above definitions, it follows that the operators \cdot and \vee satisfy several symmetry relations, to which we collectively refer to as an *experimental logic*:

$$A \vee B = B \vee A, \quad (4)$$

$$(A \vee B) \vee C = A \vee (B \vee C), \quad (5)$$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C), \quad (6)$$

$$(A \vee B) \cdot C = (A \cdot C) \vee (B \cdot C), \quad (7)$$

$$C \cdot (A \vee B) = (C \cdot A) \vee (C \cdot B). \quad (8)$$

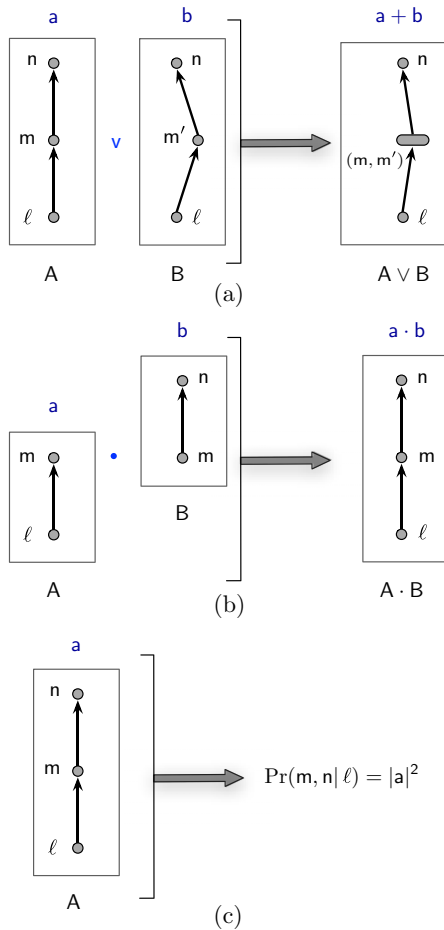


FIG. 3. (Color online) Feynman's rules for individual systems, expressed in operational terms. In each case, the sequence names are denoted A, B, \dots , while their amplitudes are denoted a, b, \dots . (a) Sum rule: $z(A \vee B) = z(A) + z(B)$. (b) Product rule: $z(A \cdot B) = z(A)z(B)$. (c) Probability rule: $\Pr(m, n | \ell) = |a|^2$.

In Ref. [10], it is shown that Feynman's rules are the unique pair-valued representation of this logic consistent with a few additional assumptions. Writing $z(X)$ for the complex-valued *amplitude* that represents sequence X , one finds (see Fig. 3)

$$z(A \vee B) = z(A) + z(B) \quad (\text{amplitude sum rule})$$

$$z(A \cdot B) = z(A)z(B) \quad (\text{amplitude product rule})$$

$$P(A) = |z(A)|^2 \quad (\text{probability rule})$$

These are Feynman's rules for measurements on individual quantum systems.

III. STATE FORMULATION OF QUANTUM THEORY

In the standard, von Neumann–Dirac formulation of quantum theory, one describes a system by specifying its state at a particular time. Temporal evolution of the system is then represented by a unitary operator, and repeatable measurements made on the system are represented by Hermitian operators. When the states of the subsystems of a composite system are given, the state of the composite system is the tensor product of the subsystem states. In this section, starting from Feynman's

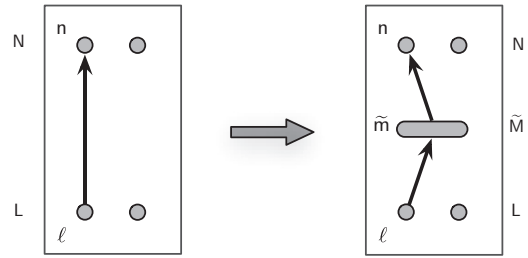


FIG. 4. No-disturbance postulate. Left: A system undergoes measurement L at time t , followed by N at time t' . For illustration, each measurement has two possible outcomes. The sequence of outcomes $[\ell, n]$ has associated probability $\Pr(n | \ell)$. Right: Trivial measurement \tilde{M} , with single outcome \tilde{m} , occurs between L and N . By the no-disturbance postulate, \tilde{M} has no effect on the probability of outcome n given ℓ .

rules and the composite systems rule (derived in Sec. IV), we derive these features with the aid of the no-disturbance postulate.

A. No-disturbance postulate

The no-disturbance postulate asserts that a trivial measurement (as defined in Sec. II A) has no effect on the outcome probabilities of subsequent measurements. For example, in the arrangement shown in Fig. 4, if the trivial measurement \tilde{M} , with single outcome $\tilde{m} = (m, m')$, is inserted between measurements L and N , the probability of outcome n given ℓ is unaffected. That is,

$$\Pr(n | \ell; \tilde{M}) = \Pr(n | \ell), \quad (9)$$

where \tilde{M} in the conditional on the left-hand side indicates that the arrangement containing \tilde{M} is the one under consideration.

The no-disturbance postulate can be regarded as capturing the essential departure of quantum physics from the mode of thinking embodied in classical physics. The key point is that, from the classical point of view, if outcome \tilde{m} is obtained, one would assert that it is a fact of the matter that the system went either through the field of sensitivity of the detector corresponding to outcome m or through that corresponding to m' , even though neither was, in fact, observed. To see the consequences of this assertion, let us consider the special case where L is repeated at time t' , where t' is immediately after t so that the system undergoes no appreciable temporal evolution in the interim (see Fig. 5). For clarity, we denote outcome ℓ of measurement L at t' by $\hat{\ell}$. Now, according to the classical assertion,

$$\begin{aligned} \Pr(\hat{\ell} | \ell; \tilde{M}) &= \Pr(\hat{\ell}, m | \ell) + \Pr(\hat{\ell}, m' | \ell) \\ &= \Pr(m | \ell) \Pr(\hat{\ell} | m, \ell) + \Pr(m' | \ell) \Pr(\hat{\ell} | m', \ell) \\ &= \Pr(m | \ell) \Pr(\hat{\ell} | m) + \Pr(m' | \ell) \Pr(\hat{\ell} | m'), \end{aligned}$$

where we have used the sum and product rules of probability theory in the first two lines, and closure in the third. If we now assume that transition probabilities are symmetric (an assumption that is independently well supported by experiment), then $\Pr(\hat{\ell} | m) = \Pr(m | \hat{\ell})$ and $\Pr(\hat{\ell} | m') = \Pr(m' | \hat{\ell})$. Setting

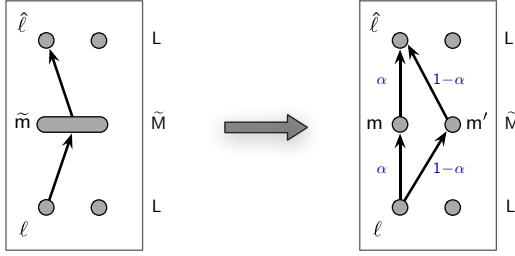


FIG. 5. (Color online) Disturbance of repeatability. Left: A system undergoes measurement \mathbf{L} at time t , and again immediately afterwards at t' , with trivial measurement $\tilde{\mathbf{M}}$ in between. Since measurement \mathbf{L} is a repeatable measurement, the no-disturbance postulate implies that it will yield the same outcome at t' as at t , even though $\tilde{\mathbf{M}}$ is present. Right: From a classical point of view, the occurrence of $\tilde{m} = (m, m')$ implies that either outcome m or m' occurred, but was not observed. The transition probabilities are as indicated, assuming that transition probabilities are symmetric. From these probabilities, it follows that repeatability is disturbed by $\tilde{\mathbf{M}}$ unless α is 0 or 1.

$\alpha = \Pr(m | \ell)$ and noting that $\Pr(m | \ell) + \Pr(m' | \ell) = 1$,

$$\Pr(\hat{\ell} | \ell; \tilde{\mathbf{M}}) = \Pr(m | \ell) \Pr(m | \hat{\ell}) + \Pr(m' | \ell) \Pr(m' | \hat{\ell}) \\ = \alpha^2 + (1 - \alpha)^2.$$

Since \mathbf{L} is a repeatable measurement, $\Pr(\hat{\ell} | \ell; \tilde{\mathbf{M}})$ should be unity. But, for this to be possible, $\alpha^2 + (1 - \alpha)^2 = 1$, which cannot hold true unless α is 0 or 1. Therefore, repeatability cannot be preserved by the insertion of $\tilde{\mathbf{M}}$ except in the special case where one of the outcomes m or m' is certain to occur. That is, on the classical assertion that the occurrence of outcome \tilde{m} implies that either outcome m or m' in fact occurs, insertion of $\tilde{\mathbf{M}}$ will, in general, disturb repeatability of \mathbf{L} . Conversely, if the no-disturbance postulate is true, one must conclude that the classical assertion is, in general, false.

B. Quantum states

We operationally define the mathematical representation of the physical state of a system at any given time as that mathematical object which enables one to compute the outcome probabilities of any measurement (chosen from a given measurement set \mathcal{M}) performed upon the system at that time.

First, suppose that a system is prepared at time t using measurement \mathbf{L} , and that measurement \mathbf{M} is subsequently performed upon it at time t' . Here, and subsequently, we assume that all measurements belong to the same measurement set, and each have N possible outcomes. We label the j th outcome of measurement \mathbf{M} as $m^{(j)}$, where $j \in \{1, 2, \dots, N\}$, and the outcomes of other measurements similarly.

Suppose that measurement \mathbf{L} yields outcome $\ell^{(i)}$. In order to compute the transition probabilities $\Pr(m^{(j)} | \ell^{(i)})$ for every j , it suffices to know the amplitude vector $\mathbf{v} = (v_1, \dots, v_N)$ whose j th component is the amplitude of the sequence $[\ell^{(i)}, m^{(j)}]$. Then, $\Pr(m^{(j)} | \ell^{(i)}) = |v_j|^2$. Since the outcomes of \mathbf{M} are mutually exclusive and exhaustive, $|\mathbf{v}|^2 = \sum_j |v_j|^2 = 1$. Insofar as calculating the outcome probabilities of \mathbf{M} performed at t' , the object \mathbf{v} suffices.

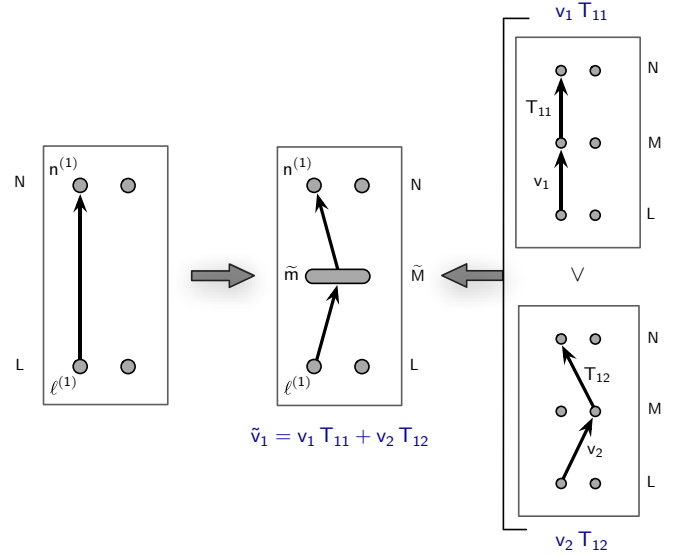


FIG. 6. (Color online) Left: A system undergoes measurement \mathbf{L} at time t , followed by \mathbf{N} at time t' , yielding sequence $[\ell^{(1)}, n^{(1)}]$. Middle: If trivial measurement $\tilde{\mathbf{M}}$ occurs immediately prior to \mathbf{N} , then, by the no-disturbance postulate, it has no effect on the outcome probabilities of \mathbf{N} . Hence, the probability $\Pr(n^{(1)} | \ell^{(1)}) = |\tilde{v}_1|^2$. Right: From the amplitude sum rule, $\tilde{v}_1 = v_1 T_{11} + v_2 T_{12} = (\mathbf{TV})_1$.

Second, suppose that, at time t' , instead of \mathbf{M} , we wish to perform measurement \mathbf{N} , and to compute its outcome probabilities. To do so, we now make use of the no-disturbance postulate, according to which we can insert the trivial form of measurement \mathbf{M} , which we denote $\tilde{\mathbf{M}}$, prior to measurement \mathbf{N} , without changing the outcome probabilities of \mathbf{N} . We insert $\tilde{\mathbf{M}}$ immediately prior to \mathbf{N} in order that the system undergoes no appreciable temporal evolution between $\tilde{\mathbf{M}}$ and \mathbf{N} (see Fig. 6). We can now compute the outcome probabilities of \mathbf{N} in the modified arrangement instead. Now, in this arrangement, the sequence $[\ell^{(i)}, \tilde{m}, n^{(k)}]$, where $\tilde{m} \equiv (m^{(1)}, \dots, m^{(N)})$, can be decomposed as

$$[\ell^{(i)}, \tilde{m}, n^{(k)}] = \bigvee_j [\ell^{(i)}, m^{(j)}, n^{(k)}] \quad (10)$$

$$= \bigvee_j [\ell^{(i)}, m^{(j)}] \cdot [m^{(j)}, n^{(k)}], \quad (11)$$

where \bigvee indicates parallel combination. Hence, given the amplitudes T_{kj} of the sequences $[m^{(j)}, n^{(k)}]$, the amplitude sum and product rules imply that the amplitude \tilde{v}_k of sequence $[\ell^{(i)}, \tilde{m}, n^{(k)}]$ is

$$\tilde{v}_k = \sum_j v_j T_{kj} = (\mathbf{TV})_k, \quad (12)$$

where \mathbf{T} is a matrix with components T_{kj} . Since the system undergoes no appreciable temporal evolution between measurement $\tilde{\mathbf{M}}$ and \mathbf{N} , the matrix \mathbf{T} captures precisely the relationship between $\tilde{\mathbf{M}}$ and \mathbf{N} . We shall refer to it as the transformation matrix from $\tilde{\mathbf{M}}$ to \mathbf{N} . Hence, the transition probability

$$\Pr(\tilde{m}, n^{(k)} | \ell^{(i)}) = |(\mathbf{TV})_k|^2. \quad (13)$$

Using the product rule of probability theory,

$$\Pr(\tilde{m}, n^{(k)} | \ell^{(i)}) = \Pr(n^{(k)} | \ell^{(i)}) \Pr(\tilde{m} | n^{(k)}, \ell^{(i)}), \quad (14)$$

and noting that $\Pr(\tilde{m} | n^{(k)}, \ell^{(i)}) = 1$, we obtain

$$\Pr(n^{(k)} | \ell^{(i)}) = |(\mathbf{T}\mathbf{v})_k|^2. \quad (15)$$

This statement holds for the modified experiment in which $\tilde{\mathbf{M}}$ occurs. But, by the nondisturbance postulate, it also holds true for the original experiment.

Thus, the object \mathbf{v} , which is specified with respect to \mathbf{M} , not only allows one to compute the outcome probabilities of \mathbf{M} , but *also* to compute the outcome probabilities of any other measurement, \mathbf{N} , provided one is given the transformation matrix \mathbf{T} from \mathbf{M} to \mathbf{N} . Therefore, from the operational point of view stated earlier, \mathbf{v} represents the state of the system at time t' .

C. Representation of measurements

Since the outcomes of \mathbf{N} are mutually exclusive and exhaustive, Eq. (15) becomes

$$\sum_k \Pr(n^{(k)} | \ell^{(i)}) = \sum_k |(\mathbf{T}\mathbf{v})_k|^2 = 1, \quad (16)$$

which implies that $|\mathbf{T}\mathbf{v}|^2 = 1$. But \mathbf{v} can be freely varied by varying the initial measurement \mathbf{L} , its outcome $\ell^{(i)}$, and the interaction with the system in the interval $[t, t']$. Therefore, the transformation matrix \mathbf{T} , which connects \mathbf{M} to \mathbf{N} , is unitary.

To determine the states that can be prepared by measurement \mathbf{N} , we use the fact that, since \mathbf{N} is repeatable, if a system is prepared at time t using measurement \mathbf{N} with outcome $n^{(q)}$, the same outcome is obtained when the measurement is immediately repeated. Therefore, using Eq. (15), the state \mathbf{u}_q that is prepared must be such that

$$\Pr(n^{(k)} | n^{(q)}) = |(\mathbf{T}\mathbf{u}_q)_k|^2 = \delta_{qk}, \quad (17)$$

which implies that $\mathbf{u}_q = (T_{q1}, \dots, T_{qN})^\dagger$ up to a predictively irrelevant overall phase. In terms of the \mathbf{u}_q , one can write $\tilde{\mathbf{v}}_k = \mathbf{u}_k^\dagger \mathbf{v}$, so that Eq. (15) becomes

$$\Pr(n^{(k)} | \ell^{(i)}) = |\mathbf{u}_k^\dagger \mathbf{v}|^2, \quad (18)$$

which is the Born rule with \mathbf{u}_k and \mathbf{v} specified with respect to \mathbf{M} . Therefore, measurement \mathbf{N} can be characterized in terms of the \mathbf{u}_q , which form an orthonormal basis of \mathbb{C}^N . Alternatively, as is more conventional, we can represent \mathbf{N} in terms of the Hermitian matrix $\mathbf{N} = \sum_q a_q \mathbf{u}_q \mathbf{u}_q^\dagger$, where a_q is the value associated with outcome $n^{(q)}$.

In the special case where measurement \mathbf{N} is the same as \mathbf{M} , it follows from the repeatability of measurements that $\Pr(m^{(k)} | m^{(j)}) = \delta_{jk}$. Therefore the transformation matrix \mathbf{T}' that relates \mathbf{M} to itself has the property that $T'_{kj} = \delta_{kj} e^{i\phi_k}$, where the ϕ_k are phases. Now, the states $\mathbf{u}'_q = (T'_{q1}, \dots, T'_{qN})^\dagger$ prepared by \mathbf{M} are predictively relevant only via Eq. (18), whose result is insensitive to the values of the ϕ_k . Therefore, without loss of generality, the ϕ_k can all be set to zero, so that \mathbf{T}' reduces to the identity matrix \mathbf{I} . Hence, measurement \mathbf{M} is represented by a diagonal Hermitian matrix, \mathbf{M} .

D. Relationship between representations

Thus far, we have specified the state of the system \mathbf{v} and the states \mathbf{u}_k that are prepared by measurement \mathbf{N} , with respect to measurement \mathbf{M} . Suppose that we were instead to represent these states as \mathbf{v}' and \mathbf{u}'_k with respect to some other measurement \mathbf{M}' . The no-disturbance postulate can then be used to relate the new representation to the old representation by completely coarse-graining measurement \mathbf{M}' and inserting measurement \mathbf{M} immediately afterwards. If a state is represented by \mathbf{v}' with respect to measurement \mathbf{M}' , and \mathbf{V} is the transformation matrix from \mathbf{M}' to \mathbf{M} , then

$$v_i = \sum_j v'_j V_{ij} = (\mathbf{V}\mathbf{v}')_i, \quad (19)$$

so that $\mathbf{v} = \mathbf{V}\mathbf{v}'$. Since \mathbf{V} is unitary, this can be inverted to give $\mathbf{v}' = \mathbf{V}^\dagger \mathbf{v}$. Similarly $\mathbf{u}'_k = \mathbf{V}^\dagger \mathbf{u}_k$, which implies that the Hermitian operator \mathbf{N}' that represents \mathbf{N} with respect to \mathbf{M}' , is given by $\mathbf{V}^\dagger \mathbf{N} \mathbf{V}$.

If measurement \mathbf{M}' is represented with respect to \mathbf{M} by Hermitian matrix \mathbf{M}' with eigenvectors \mathbf{w}_i , then the transformation matrix \mathbf{V} has components $V_{ij} = (\mathbf{w}_j)_i$.

E. Unitary representation of temporal evolution

Suppose that a system is prepared using measurement \mathbf{L} at time t , and then undergoes measurement \mathbf{M} at time t' . Immediately prior to measurement \mathbf{M} , the system is in state \mathbf{v} . Suppose now that measurement \mathbf{M} is completely coarse-grained, and an additional measurement \mathbf{M} is performed at time $t'' > t'$. In this arrangement, the sequence $[\ell^{(i)}, \tilde{m}, m^{(k)}]$ can be decomposed as

$$[\ell^{(i)}, \tilde{m}, m^{(k)}] = \bigvee_j [\ell^{(i)}, m^{(j)}] \cdot [m^{(j)}, m^{(k)}]. \quad (20)$$

The temporal evolution of the system in interval $[t', t'']$ is represented by the amplitudes U_{kj} of the sequences $[m^{(j)}, m^{(k)}]$. The amplitude sum and product rules accordingly imply that the amplitude \tilde{v}_k of sequence $[\ell^{(i)}, \tilde{m}, m^{(k)}]$ is

$$\tilde{v}_k = \sum_j v_j U_{kj} = (\mathbf{U}\mathbf{v})_k. \quad (21)$$

Therefore, the state of the system $\tilde{\mathbf{v}}$ immediately prior to the last measurement is

$$\tilde{\mathbf{v}} = \mathbf{U}\mathbf{v}. \quad (22)$$

Now, the initial state \mathbf{v} can be arbitrarily varied, but the states \mathbf{v} and $\tilde{\mathbf{v}}$ are normalized. Therefore, \mathbf{U} is unitary.

Since temporal evolution from t' to t'' is represented by \mathbf{U} , temporal evolution from t'' to t' is represented by \mathbf{U}^\dagger . Therefore, if we denote the temporal inverse of the sequence A as A^{-1} , then

$$z(A^{-1}) = z^*(A), \quad (23)$$

to which we shall refer as the *amplitude temporal inversion rule*.

F. Composite systems

At time t , system S_1 is prepared by measurement \mathbf{L}_1 with outcome ℓ_1 , and S_2 by measurement \mathbf{L}_2 with outcome ℓ_2 . Suppose that these systems evolve without interacting with

one another until time t' , at which point they are measured, respectively, by \mathbf{M}_1 and \mathbf{M}_2 , which respectively have N_1, N_2 possible outcomes, $m_1^{(j)}$ and $m_2^{(k)}$ with $j \in \{1, \dots, N_1\}$ and $k \in \{1, \dots, N_2\}$. Then the components of the respective states v' and v'' of the systems immediately prior to t' are given by

$$v'_j = z([\ell_1, m_1^{(j)}]) \quad \text{and} \quad v''_k = z([\ell_2, m_2^{(k)}]). \quad (24)$$

Viewed as a single system S , the system is prepared by measurement \mathbf{L} with outcome $(\ell_1 : \ell_2)$, and subsequently undergoes measurement \mathbf{M} with $N_1 N_2$ possible outcomes $(m_1^{(j)} : m_2^{(k)})$. With respect to \mathbf{M} , the components of the state of the system \mathbf{v} immediately prior to \mathbf{M} is given by

$$v_{(i-1)N_2+j} = z([\ell_1 : \ell_2], (m_1^{(j)} : m_2^{(k)})). \quad (25)$$

But, by the composite systems rule, Eq. (44), which we shall derive in Sec. IV,

$$z([\ell_1 : \ell_2], (m_1^{(j)} : m_2^{(k)})) = z([\ell_1, m_1^{(j)}]) z([\ell_2, m_2^{(k)}]), \quad (26)$$

so that $v_{(i-1)N_2+j} = v'_j v''_k$. Hence,

$$\mathbf{v} = \mathbf{v}' \otimes \mathbf{v}''. \quad (27)$$

G. Summary

In summary, we have derived the following:

- (i) If a system is prepared by some measurement \mathbf{L} at time t , then its state at time t' immediately prior to measurement \mathbf{N} is given by vector \mathbf{v} with respect to reference measurement \mathbf{M} .
- (ii) Measurement \mathbf{N} is represented by Hermitian operator $\mathbf{N} = \sum_q a_q \mathbf{u}_q \mathbf{u}_q^\dagger$, where the \mathbf{u}_q , also specified with respect to \mathbf{M} , are the states prepared by \mathbf{N} . When performed on the system in state \mathbf{v} , the probability of the k th outcome of \mathbf{N} is given by $\Pr(n^{(k)} | \mathbf{v}) = |\mathbf{u}_k^\dagger \mathbf{v}|^2$.
- (iii) If the reference measurement is changed to \mathbf{M}' , then the state \mathbf{v}' of the system with respect to \mathbf{M}' is given by $\mathbf{v}' = \mathbf{V}^\dagger \mathbf{v}$, where \mathbf{V} is the unitary transformation matrix from \mathbf{M}' to \mathbf{M} .
- (iv) The state \mathbf{v} evolves unitarily in the time between measurements.
- (v) If the two subsystems of a composite system are in states \mathbf{v}' and \mathbf{v}'' , then the composite system is in state $\mathbf{v} = \mathbf{v}' \otimes \mathbf{v}''$.

Collectively, this set of statements is equivalent to von Neumann's postulates for finite-dimensional quantum systems.

IV. COMPOSITE SYSTEMS

A. Composition operator and its symmetries

Suppose that one physical system, denoted S_1 , undergoes an experiment involving measurements $\mathbf{L}_1, \mathbf{M}_1, \mathbf{N}_1$ at successive times t_1, t_2, t_3 , while another system, S_2 , undergoes an experiment where the measurements $\mathbf{L}_2, \mathbf{M}_2, \mathbf{N}_2$ at these same times. The measurements on S_1 yield the outcome sequence $A = [\ell_1, m_1, n_1]$, while the measurements on S_2 yield $B = [\ell_2, m_2, n_2]$. As described earlier, in Sec. II A, one can also describe the situation by saying that measurements \mathbf{L}, \mathbf{M} , and \mathbf{N} are performed on the composite system S , yielding the sequence $C = [(\ell_1 : \ell_2), (m_1 : m_2), (n_1 : n_2)]$. We now symbolize the relationship between A, B , and C by defining a binary operator \times , the *composition* operator, which

here acts on A and B to generate the sequence

$$C = A \times B. \quad (28)$$

Generally, the operator \times combines any two sequences of the same length, each obtained from a different experiment on different physical systems where each measurement in one experiment occurs at the same time as one measurement in the other experiment.

From the definition just given, it follows that \times is associative. To see this, consider the three sequences $A = [\ell_1, m_1]$, $B = [\ell_2, m_2]$, and $C = [\ell_3, m_3]$, obtained from three different experiments satisfying the condition just stated above. We can then combine these to yield the sequence $D = [(\ell_1 : \ell_2 : \ell_3), (m_1 : m_2 : m_3)]$ in two different ways, either as $A \times (B \times C)$ or as $(A \times B) \times C$. Hence,

$$A \times (B \times C) = (A \times B) \times C. \quad (29)$$

Similar considerations show that \times also satisfies the following symmetry relations involving the operators \vee and \cdot :

$$(A \cdot B) \times (C \cdot D) = (A \times C) \cdot (B \times D), \quad (30)$$

$$A \times (B \vee C) = (A \times B) \vee (A \times C), \quad (31)$$

$$(A \vee B) \times C = (A \times C) \vee (B \times C). \quad (32)$$

The cross-multiplicativity and left-distributivity properties expressed in Eqs. (30) and (31) are illustrated in Figs. 7 and 8, respectively.

B. Composite systems rule

If systems S_1 and S_2 are noninteracting, we postulate that, in Eq. (28), the amplitude c of sequence C is determined by the amplitudes a, b of the sequences A, B , so that

$$c = F(a, b), \quad (33)$$

where F is some continuous complex-valued function to be determined. This is the *composition postulate*, given which Eqs. (29), (30), (31), and (32), respectively, imply

$$F(a, F(b, c)) = F(F(a, b), c), \quad (34)$$

$$F(ab, cd) = F(a, c) F(b, d), \quad (35)$$

$$F(a, b + c) = F(a, b) + F(a, c), \quad (36)$$

$$F(a + b, c) = F(a, c) + F(b, c). \quad (37)$$

We can now solve these for F . Due to the cross-multiplicativity equation, Eq. (35),

$$F(u, v) = F(u \times 1, 1 \times v) = F(u, 1) F(1, v). \quad (38)$$

To determine the form of $F(u, 1)$, we use the right-distributivity and cross-multiplicativity equations, Eqs. (37) and (35), respectively, to obtain

$$F(u_1 + u_2, 1) = F(u_1, 1) + F(u_2, 1), \quad (39a)$$

$$F(u_1 u_2, 1) = F(u_1, 1) F(u_2, 1). \quad (39b)$$

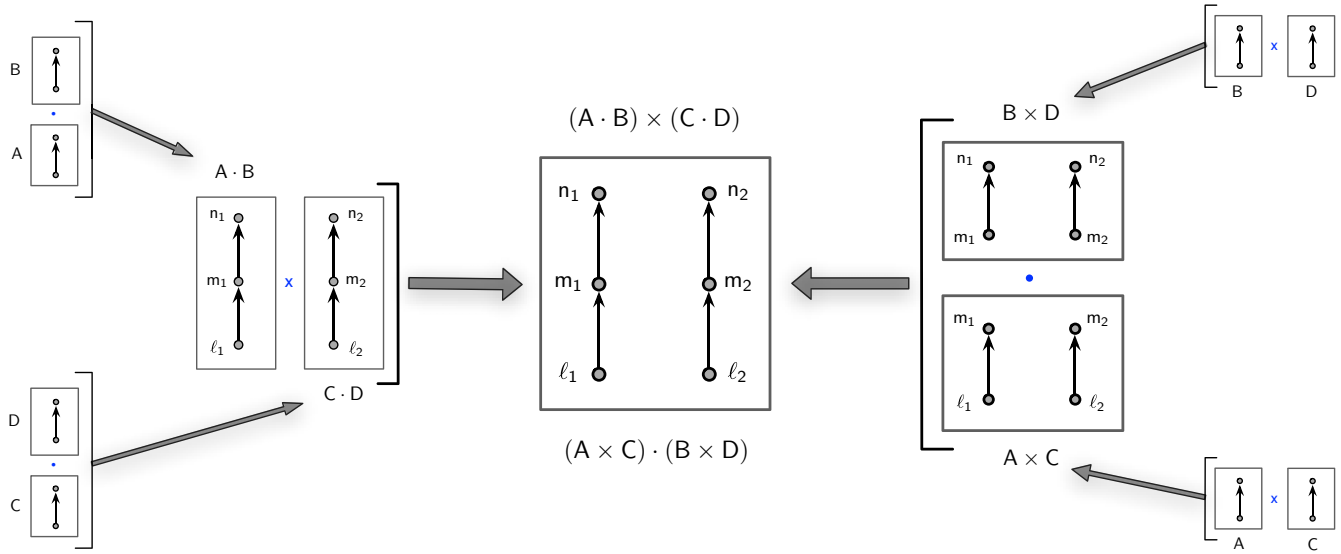


FIG. 7. (Color online) Illustration of the cross-multiplicativity property. The composite sequence $[(\ell_1 : \ell_2), (m_1 : m_2), (n_1 : n_2)]$ can be obtained by combining the sequences $A = [\ell_1, m_1]$, $B = [m_1, n_1]$, $C = [\ell_2, m_2]$, and $D = [m_2, n_2]$ in two different ways, as shown, yielding the cross-multiplicativity relation $(A \cdot B) \times (C \cdot D) = (A \times C) \cdot (B \times D)$, which is Eq. (30).

Writing $f(z) = F(z, 1)$, these two equations can be written as a pair of functional equations,

$$f(z_1 + z_2) = f(z_1) + f(z_2), \quad (40a)$$

$$f(z_1 z_2) = f(z_1) f(z_2), \quad (40b)$$

whose continuous solutions in the domain $|z| \leq 1$ are $f(z) = z$, $f(z) = z^*$, or $f(z) = 0$ (see the Appendix). The zero solution implies $F(u, v) = 0$ for all u, v , and is therefore

inadmissible. Therefore,

$$F(u, 1) = \begin{cases} u & (41a) \\ u^* & (41b) \end{cases}$$

To eliminate the possibility $F(u, 1) = u^*$ we make use of the associativity equation, Eq. (34), which implies

$$F(u, F(1, 1)) = F(F(u, 1), 1). \quad (42)$$

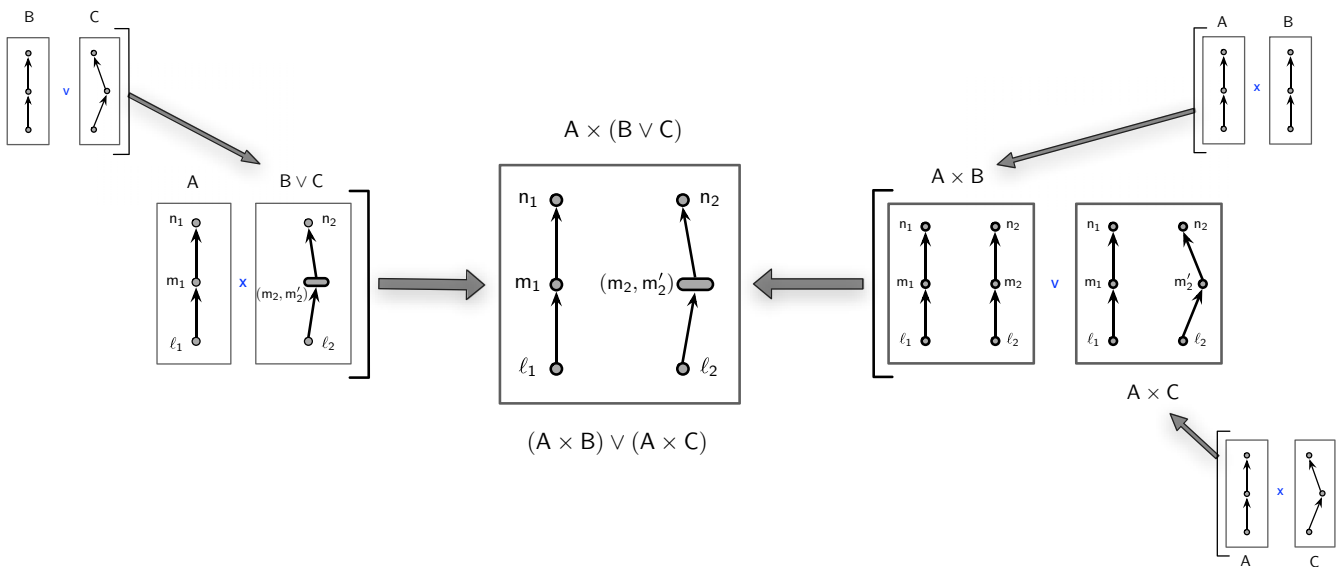


FIG. 8. (Color online) Illustration of left-distributivity of \times over \vee . The composite sequence $[(\ell_1, \ell_2), (m_1 : (m_2, m_2')), (n_1, n_2)]$ can be obtained by combining the sequences $A = [\ell_1, m_1, n_1]$, $B = [\ell_2, m_2, n_2]$, and $C = [\ell_2, m_2', n_2]$ in two different ways, as shown, yielding the relation $A \times (B \vee C) = (A \times B) \vee (A \times C)$, which is Eq. (31).

Now, from the cross-multiplicativity equation, Eq. (35), $F(u \times 1, v \times 1) = F(u, v) F(1, 1)$, which implies $F(1, 1) = 1$ since the zero solution for $F(u, v)$ is inadmissible. Therefore, Eq. (42) becomes

$$F(u, 1) = F(F(u, 1), 1). \quad (43)$$

But this is incompatible with $F(u, 1) = u^*$ since $u^* = F(u, 1) \neq F(F(u, 1), 1) = F(u^*, 1) = u$. We are therefore left with $F(u, 1) = u$, which is compatible with Eq. (43). A parallel argument establishes that $F(1, v) = v$. Therefore, from Eq. (38),

$$F(u, v) = uv. \quad (44)$$

This is the amplitude rule for distinguishable, noninteracting composite systems. We refer to it as the *composite system rule*.

V. DERIVATION OF DIRAC'S AMPLITUDE-ACTION RULE

Consider a quantum system that is subject to position measurements at successive times. Suppose that the intervals between these successive times are sufficiently small that the resulting measurement sequence is well approximated by a continuous classical trajectory (or simply “path”) of the same system as treated according to the framework of classical physics. Dirac's amplitude-action rule asserts that the amplitude associated with the measurement sequence is given by $e^{iS/\hbar}$, where S is the classical action associated with the corresponding classical path. We now derive the form of this rule up to \hbar from two elementary properties of the classical action:

Additivity. If sequence $C = A \cdot B$, then $S_C = S_A + S_B$, where S_X is the classical action of the path corresponding to sequence X .

Inversion. The action $S_{A^{-1}}$ associated with sequence A^{-1} is $-S_A$.

Our assumption is that the amplitude $z(A)$ of sequence A , with corresponding classical action S_A , is given by $f(S_A)$, where f is a continuous, complex-valued function.

The additivity and inversion properties of the classical action induce two functional equations in f . First, the amplitude $z(C)$ of sequence $C = A \cdot B$ can be computed in two ways (see Fig. 9), either using the action additivity property,

$$z(C) = f(S_C) = f(S_A + S_B),$$

or using the amplitude product rule,

$$z(C) = z(A)z(B) = f(S_A)f(S_B),$$

so that

$$f(x + y) = f(x)f(y). \quad (45)$$

Second, the amplitude $z(A^{-1})$ of sequence A^{-1} can be computed either using the amplitude temporal inversion rule, Eq. (23),

$$z(A^{-1}) = z^*(A) = f^*(S_A),$$

or using the action inversion property,

$$z(A^{-1}) = f(S_{A^{-1}}) = f(-S_A),$$

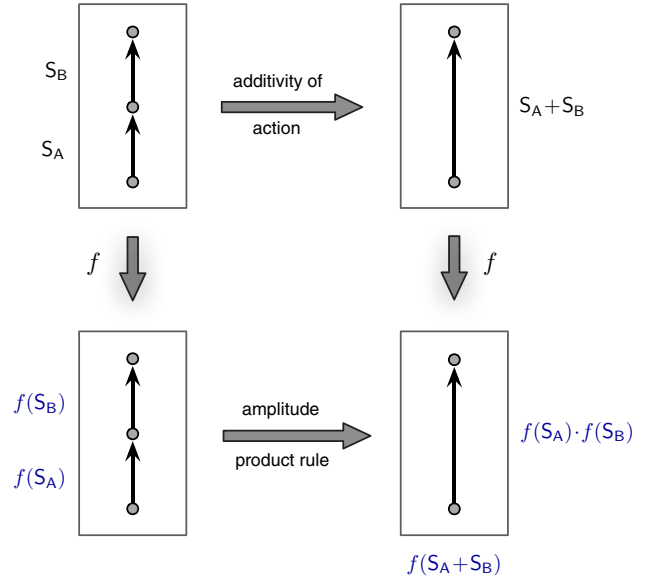


FIG. 9. (Color online) The amplitude of the path $C = A \cdot B$ can be obtained from the classical actions S_A, S_B of paths A, B in two different ways: (i) obtain the action $S_C = S_A + S_B$, whose corresponding amplitude is $f(S_A + S_B)$; and (ii) use f to obtain the amplitudes $f(S_A), f(S_B)$ and then compose these to obtain amplitude $f(S_A)f(S_B)$. Hence, $f(S_A + S_B) = f(S_A)f(S_B)$.

so that

$$f^*(x) = f(-x). \quad (46)$$

We can solve Eq. (45) by writing $f(x) = R(x)e^{i\Phi(x)}$, with R, Φ real, to obtain, for integer n ,

$$R(x + y) = R(x)R(y), \quad (47)$$

$$\Phi(x + y) = \Phi(x) + \Phi(y) + 2\pi n. \quad (48)$$

The second equation can be transformed via $\tilde{\Phi}(x) = \Phi(x) + 2\pi n$ to give

$$\tilde{\Phi}(x + y) = \tilde{\Phi}(x) + \tilde{\Phi}(y). \quad (49)$$

Equations (47) and (49) are two of Cauchy's standard function equations, with general solutions $R(x) = e^{\beta x}$ and $\tilde{\Phi}(x) = \alpha x$, where $\alpha, \beta \in \mathbb{R}$ [45]. Hence, $\Phi(x) = \alpha x - 2\pi n$, and

$$f(x) = e^{\beta x} e^{i\alpha x}. \quad (50)$$

But Eq. (46) then implies that $\beta = 0$. Therefore,

$$z(A) = e^{i\alpha S_A}, \quad (51)$$

where the constant α has dimensions of \hbar^{-1} . This is Dirac's amplitude-action rule up to \hbar .

VI. DISCUSSION

In this paper, we have shown that it is possible to systematically build Feynman's rules into a complete formulation of finite-dimensional quantum theory. The key physical ingredient in this process has been the no-disturbance postulate, which expresses the singularly nonclassical fact that a trivial measurement does not disturb the outcome probabilities of

subsequent measurements on the system. This postulate allows us to introduce the concept of the state of a system in a systematic way, and to prove the unitarity of temporal evolution and the Hermiticity of measurement operators. We have also derived the composite system rule and Dirac's amplitude-action rule, each from a single elementary and natural assumption, by making use of the fact that these assumptions must be consistent with Feynman's rules.

The work described here, in concert with our earlier derivation of Feynman's rules, constitutes a complete derivation of the finite-dimensional quantum formalism. The derivation has a number of important implications for our understanding of quantum theory in addition to those mentioned in the Introduction.

First, most other attempts to derive the quantum formalism from physically motivated postulates (such as [5,11,46]) depend upon postulates (such as purifiability [11] or local tomography [5,46]) that concern the behavior of composite systems in order to derive the quantum formalism for *individual* systems. This tends to suggest that the behavior of composite systems is in some sense fundamental to the structure of the quantum formalism. However, in the present derivation, there is no such dependency. Instead, we have shown that it is possible to derive the formalism for individual systems *without* assumptions that overtly concern composite systems, and then to derive the rule for composite systems on the basis of a simple assumption merely by requiring consistency with the formalism for individual systems. Therefore, the present derivation strongly implies that the behavior of composite systems is a secondary feature of quantum theory, not a primary one.

Second, one of the most remarkable features of Feynman's formulation of quantum theory is the absence of a state concept, and the absence of any distinction between dynamics, on the one hand, and the relationship between measurements, on the other. We have shown here that these features *can* be recovered, but at the cost of an additional physical postulate which has nontrivial physical content.

Finally, we have shown that Dirac's amplitude-action rule follows from elementary properties of the classical action via the simple assumption that the amplitude of a sequence is determined by the corresponding action. In contrast with Dirac's argument, our approach does not depend on the particular form of the classical Lagrangian or on the existence or form of Lagrange's equations of motion, but only on two elementary properties (additivity, inversion) of the action. Hence, we have shown that Dirac's rule has a very general validity, and arises as soon as one attempts to establish a

quantitative connection between the notion of action in the Lagrangian formulation of classical mechanics, and the notion of amplitude in Feynman's formulation of quantum theory.

We conclude with two open questions. First, is the no-disturbance postulate related in any way with other informational ideas that have been proposed, such as in Refs. [11,47,48]? Second, is there a direct, general path from Dirac's amplitude-action rule to the unitary form $\exp(-i\hat{H}\Delta t/\hbar)$ of the temporal evolution operator?

ACKNOWLEDGMENT

This publication was made possible, in part, through the support of a grant from the John Templeton Foundation.

APPENDIX: SOLUTION OF A PAIR OF FUNCTIONAL EQUATIONS

We solve Eqs. (40a) and (40b) with the aid of one of Cauchy's standard functional equations,

$$h(x_1 + x_2) = h(x_1) + h(x_2), \quad (\text{A1})$$

where h is a real function and $x_1, x_2 \in \mathbb{R}$. Its continuous solution is $h(x) = ax$ with $a \in \mathbb{R}$ [45].

Setting $z_1 + z_2 = x + iy$, with $x, y \in \mathbb{R}$, in Eq. (40a) gives

$$f(x + iy) = f(x) + f(iy).$$

Applying Eq. (40a) again on $f(x_1 + x_2)$ and $f(iy_1 + iy_2)$ then implies

$$\begin{aligned} f(x_1 + x_2) &= f(x_1) + f(x_2), \\ f(iy_1 + iy_2) &= f(iy_1) + f(iy_2). \end{aligned}$$

The real and imaginary parts of both of these equations each have the form of Eq. (A1), and therefore have solutions

$$f(x) = \alpha x \quad \text{and} \quad f(iy) = \beta y$$

with $\alpha, \beta \in \mathbb{C}$, so that

$$f(x + iy) = \alpha x + \beta y. \quad (\text{A2})$$

From Eq. (40b),

$$f(1 \cdot 1) = f(1)f(1) \quad \text{and} \quad f(i \cdot i) = f(i)f(i),$$

which, due to Eq. (A2), imply

$$\alpha = \alpha^2 \quad \text{and} \quad -\alpha = \beta^2.$$

These have solutions $(\alpha, \beta) = (0, 0)$, $(1, i)$, and $(1, -i)$, which correspond to $f(z) = 0$, $f(z) = z$, and $f(z) = z^*$.

-
- [1] J. A. Wheeler, in *Foundations of Quantum Mechanics in the Light of New Technology: Proceedings of the 3rd International Symposium* (Physical Society of Japan, Tokyo, 1990).
 [2] A. Zeilinger, *Found. Phys.* **29**, 631 (1999).
 [3] C. A. Fuchs, in *Quantum Theory: Reconsideration of Foundations*, edited by A. Khrennikov (Växjö University Press, Växjö, Sweden, 2002), pp. 463–543.
 [4] A. Grinbaum, *Br. J. Philos. Sci.* **58**, 387 (2007).

- [5] L. Hardy, [arXiv:quant-ph/0101012](https://arxiv.org/abs/quant-ph/0101012).
 [6] R. Clifton, J. Bub, and H. Halvorson, *Found. Phys.* **33**, 1561 (2003).
 [7] G. M. D'Ariano, in *Foundations of Probability and Physics, 4*, edited by A. Y. K. G. Adenier and C. A. Fuchs (AIP, New York, 2007), p. 79.
 [8] P. Goyal, *New J. Phys.* **12**, 023012 (2010).
 [9] M. Reginatto, *Phys. Rev. A* **58**, 1775 (1998).

- [10] P. Goyal, K. H. Knuth, and J. Skilling, *Phys. Rev. A* **81**, 022109 (2010).
- [11] G. Chiribella, G. M. D’Ariano, and P. Perinotti, *Phys. Rev. A* **84**, 012311 (2011).
- [12] L. Masanes and M. Müller, *New J. Phys.* **13**, 063001 (2011).
- [13] B. Dakic and C. Brukner, in *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, edited by H. Halvorson (Cambridge University Press, Cambridge, 2011), pp. 365–392.
- [14] J. Barrett, *Phys. Rev. A* **75**, 032304 (2007).
- [15] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, *Nature (London)* **461**, 1101 (2009).
- [16] R. P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948).
- [17] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, 1st ed. (McGraw-Hill, New York, 1965).
- [18] R. T. Cox, *Am. J. Phys.* **14**, 1 (1946).
- [19] R. T. Cox, *The Algebra of Probable Inference* (The Johns Hopkins University Press, Baltimore, 1961).
- [20] P. Goyal and K. H. Knuth, *Symmetry* **3**, 171 (2011).
- [21] P. A. M. Dirac, *Principles of Quantum Mechanics*, 1st ed. (Oxford University Press, New York, 1930).
- [22] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).
- [23] G. Johnson and M. L. Lapidus, *The Feynman Integral and Feynman’s Operational Calculus* (Oxford University Press, New York, 2000).
- [24] In Ref. [16], Sec. 11, Feynman states, “One of the most important characteristics of quantum mechanics is its invariance under unitary transformations.... Of course, the present formulation, being equivalent to ordinary formulations, can be mathematically demonstrated to be invariant under these transformations. However, it has not been formulated in such a way that it is *physically* obvious that it is invariant.”
- [25] C. A. Fuchs and A. Peres, *Phys. Rev. A* **53**, 2038 (1996).
- [26] I am grateful to Paulo Perinotti for suggesting this point of view.
- [27] One cannot rule out measurements that provide no useful information and yet still disturb the state.
- [28] C. Pfister and S. Wehner, *Nat. Commun.* **4**, 1851 (2013).
- [29] P. Goyal, [arXiv:1309.0478](https://arxiv.org/abs/1309.0478).
- [30] P. A. M. Dirac, *Phys. Z. Sowjetunion* **3**, 64 (1933).
- [31] R. P. Feynman, *Phys. Rev.* **76**, 769 (1949).
- [32] A. Caticha, *Found. Phys.* **30**, 227 (2000).
- [33] Y. Tikochinsky, *Phys. Rev. A* **37**, 3553 (1988).
- [34] A. Caticha, *Phys. Rev. A* **57**, 1572 (1998).
- [35] P. Pearle, *Phys. Rev. D* **13**, 857 (1976).
- [36] I. Bialynicki-Birula and J. Mycielski, *Ann. Phys.* **100**, 62 (1976).
- [37] A. Shimony, *Phys. Rev. A* **20**, 394 (1979).
- [38] S. Weinberg, *Phys. Rev. Lett.* **62**, 485 (1989).
- [39] S. Weinberg, *Ann. Phys. (N.Y.)* **194**, 336 (1989).
- [40] S. B. Giddings, [arXiv:hep-th/9508151](https://arxiv.org/abs/hep-th/9508151).
- [41] C. Simon, V. Bužek, and N. Gisin, *Phys. Rev. Lett.* **87**, 170405 (2001).
- [42] M. Ferrero, D. Salgado, and J. L. Sánchez-Gómez, *Phys. Rev. A* **70**, 014101 (2004).
- [43] M. Ferrero, D. Salgado, and J. L. Sánchez-Gómez, *Phys. Rev. A* **73**, 034304 (2006).
- [44] P. Goyal, *Phys. Rev. A* **78**, 052120 (2008).
- [45] J. Aczél, *Lectures on Functional Equations and their Application* (Academic Press, New York, 1966).
- [46] L. Hardy, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. Spekkens (unpublished), [arXiv:1303.1538](https://arxiv.org/abs/1303.1538).
- [47] Č. Brukner and A. Zeilinger, *Phys. Rev. Lett.* **83**, 3354 (1999).
- [48] C. Brukner and A. Zeilinger, *Found. Phys.* **39**, 677 (2009).