Four Postulates of Quantum Mechanics Are Three

Gabriele Carcassi[®], Lorenzo Maccone[®], and Christine A. Aidala[®]¹ ¹Physics Department, University of Michigan, 450 Church Street, Ann Arbor, Michigan 48109-1040, USA ²Dip. Fisica and INFN Sez. Pavia, University of Pavia, via Bassi 6, I-27100 Pavia, Italy

(Received 3 September 2020; revised 2 December 2020; accepted 21 January 2021; published 16 March 2021)

The tensor product postulate of quantum mechanics states that the Hilbert space of a composite system is the tensor product of the components' Hilbert spaces. All current formalizations of quantum mechanics that do not contain this postulate contain some equivalent postulate or assumption (sometimes hidden). Here we give a natural definition of a composite system as a set containing the component systems and show how one can logically derive the tensor product rule from the state postulate and from the measurement postulate. In other words, our Letter reduces by one the number of postulates necessary to quantum mechanics.

DOI: 10.1103/PhysRevLett.126.110402

In this Letter we derive the tensor product postulate (which, hence, loses its status of postulate) from two other postulates of quantum mechanics: the state postulate and the measurement postulate. The tensor product postulate does not appear in all axiomatizations of quantum mechanics: it has even been called "postulate 0" in some literature [1]. A widespread belief is that it is a direct consequence of the superposition principle, and it is hence not a necessary axiom. This belief is mistaken: the superposition principle is encoded into the quantum axioms by requiring that the state space is a linear vector space. This is, by itself, is insufficient to single out the tensor product, as other linear products of linear spaces exist, such as the direct product, the exterior or wedge product, or the direct sum of vector spaces, which is used in classical mechanics to combine state spaces of linear systems. These are all maps from linear spaces to linear spaces but they differ in how the linearity of one is mapped to the linearity of the others [2]. This belief may have arisen from the seminal book of Dirac [3], who introduces tensor products (Chap. 20) by appealing to linearity. However, he adds the seemingly innocuous request that the product among spaces be distributive (rather, bilinear), which is equivalent to postulating tensor products (or linear functions of them). It is not an innocuous request. For example it does not hold where the composite vector space of two linear spaces is described by the direct product, e.g., in classical mechanics, for two strings of a guitar: it is not distributive. (General classical systems, not only linear ones, are also composed through the direct product.) Of course, Dirac is not constructing an axiomatic formulation, so his "sleight of hand" can be forgiven. In contrast, von Neumann (Ref. [4] Chap. VI.2, also Ref. [5]) introduces tensor products by noticing that this is a natural choice in the position representation of wave mechanics (where they were introduced in Refs. [6,7]), and then *explicitly postulates* them in general: "This rule of transformation is correct in any case for the coordinate and momentum operators [...] and it conforms with the [observable axiom and its linearity principles], we therefore postulate them generally." [4]. More mathematical or conceptually-oriented modern formulations (e.g., Refs. [8–11]) introduce this postulate explicitly. An interesting alternative is provided in Refs. [12,13]: after introducing tensor products, Ballentine verifies a posteriori that they give the correct laws of composition of probabilities. Similarly, Peres uses relativistic locality [14]. While these procedures seemingly bypass the need to postulate the tensor product, they do not guarantee that this is the *only* possible way of introducing composite systems in quantum mechanics. In the framework of quantum logic, tensor products arise from some additional conditions [15] which (in contrast to what is done here) are not connected to the other postulates. In Refs. [16,17] tensor products were obtained by specifying additional physical or mathematical requirements.

Let us first provide a conceptual overview of our approach. We start from the natural definition of a composite system as the set of two (or more) quantum systems. The composite system is therefore made of system A and (joined with) system B and nothing else. The first key insight is that the first two postulates of quantum theory (introduced below) already assume that the preparation of one system is independent from the preparation of another (statistical independence). In fact, we cannot even talk about a system in the first place if we cannot characterize it independently. The second key insight is that, using the law of composition of probabilities of independent events, we can easily find a map M that takes the state of the component systems and gives the composite state for the statistically independent case. These insights are enough to characterize mathematically the state space of the composite: the linearity given by the Hilbert space, together with the fact that the composite system is fully described by the observables of A and B, allows us to extend the construction from the statistically independent composite states to the general case (that includes entangled states). So the work consists of two interrelated efforts: a physical argument that starts from the first two postulates and leads to the necessary existence of the composition map M and its properties together with a formal argument that shows how M leads to the tensor product.

This map M acts on the state spaces of the subsystems. Each pure state is identified by a ray ψ , a subspace of the system's Hilbert space comprising all vectors ψ differing by their (nonzero) modulus and phase: a one-dimensional complex subspace (a complex plane). In the same way, constraining the observable X to a particular outcome value x_0 means identifying the subspace comprising all nonnormalized eigenvectors $|x_0\rangle$ of arbitrary phase such that $X|x_0\rangle = x_0|x_0\rangle$. The map M establishes a relationship between the states of the subsystems and the composite, so it is a map between subspaces, not vectors. Therefore, M acts on the projective spaces, where all vectors within the same ray are "collapsed" into a single point (i.e., a quotient space in the equivalence class), removing the unphysical "overspecification" of the phase and of the modulus. The physical requirements on M are such that we can find a bilinear map m between vectors that acts consistently with M in terms of subspaces. This map m is the tensor product.

More in detail, the physical requirements of statistical independence, together with the fact that one can arbitrarily prepare the states of the subsystems, imply three conditions on the map m: (H1) totality: the map is defined on all states of the subsystems; (H2) bilinearity: the map is bilinear thanks to the fundamental theorem of projective geometry; (H3) span surjectivity: the span of the image of the map coincides with the full composite Hilbert space. We then prove that, if these three conditions H1, H2, and H3 hold, then the map m is the tensor product, namely, the Hilbert space of the composite system is the tensor product of the components' Hilbert spaces. The tensor product "postulate" hence loses its status of a postulate. An overview of all these logical implications is given in Fig. 1. The rest of the Letter contains the sketch of this argument, including all the physical arguments outlined above. The Supplemental Material [18] contains the mathematical details.

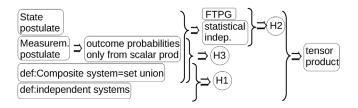


FIG. 1. Schematic depiction of the logical implications used in this Letter [fundamental theorem of projective geometry (FTPG)].

We start with the axiomatization of quantum mechanics based on the following postulates (e.g., Refs. [8–11]): (a) The pure state of a system is described by a ray ψ corresponding to a set of nonzero vectors $|\psi\rangle$ in a complex Hilbert space, and the system's observable properties are described by self-adjoint operators acting on that space. (b) The probability that a measurement of a property X, described by the operator with spectral decomposition $\sum_{x,i} x |x_i\rangle \langle x_i|/\langle x_i|x_i\rangle$ (*i* a degeneracy index), returns a value x given that the system is in state ψ is $p(x|\psi) =$ $\sum_{i} |\langle \psi | x_i \rangle|^2 / \langle x_i | x_i \rangle \langle \psi | \psi \rangle$ (Born rule). (c) The state space of a composite system is given by the tensor product of the spaces of the component systems. (d) The time evolution of an isolated system is described by a unitary operator acting on a vector representing the system state, $|\psi(t)\rangle =$ $U_t|\psi(t=0)\rangle$ or, equivalently, by the Schrödinger equation. The rest of quantum theory can be derived from these axioms. While some axiomatizations introduce further postulates, we will be using only (a) and (b) to derive (c), so the above are sufficient to our aims.

Note that we limit ourselves to kinematically independent systems, where all state vectors $|\psi\rangle$ in the system's Hilbert space ${\cal H}$ describe a valid state, unconditioned on anything else. We call this condition "preparation independence" and it should be noted that the tensor product applies only in this case. For example, the composite system of two electrons is not the tensor product, rather the antisymmetrized tensor product, precisely because the second electron cannot be prepared in the same state of the first. We note that restrictions due to superselection rules arise either from practical (not fundamental) limitations on the actions of the experimenter [25–27] or from the use of ill-defined quantum systems. In the example above, the field is the proper quantum system and the electrons are its excitations [28].

The definition of a composite system as containing *only* the collection of the subsystems means that any preparation of both subsystems independently must correspond to the preparation of the composite system. Since states are defined by postulate (a) as rays in the respective Hilbert spaces, there must exist a map $M: \underline{A} \times \underline{B} \to \underline{C}$ that takes a pair of states for the subsystems (\underline{A} and \underline{B} represent the projective space, where each point represents all vectors that identify the same state, and the Cartesian product is the set of all possible pairs) and returns a state in the projective space C for the composite. To visualize the geometrical meaning of M directly within the Hilbert spaces, given a ray (a complex plane) in each of A and B, M returns a ray (a complex plane) in C. Our final goal will be to find a map m: $\mathcal{A} \times \mathcal{B} \to \mathcal{C}$ that acts on vectors in the Hilbert spaces \mathcal{A}, \mathcal{B} , and C consistently with M. Namely, m(a,b) = M(a,b)where the underline indicates the elements in the projective space. Again geometrically, m takes a vector in each of Aand \mathcal{B} , and returns a vector in \mathcal{C} and we want this to be consistent with M such that vectors picked from the same rays will return vectors in the same ray. We will prove that the map m is the tensor product. We focus on pure states here: the argument can be extended to mixed states using standard tools [12].

The map M must be injective: as said above, different states of the subsystems must correspond, by definition of composite system, to different states of the composite. Moreover, preparation independence implies that M, and hence m, must be total maps (condition H1): each subsystem of the composite system can be independently prepared and gives rise to a state of the composite. H1 is not sufficient to identify the tensor product: by itself it does not even guarantee that the map m is linear.

$$p(a \wedge b|\psi \wedge b) = p(a|\psi), \tag{1}$$

$$p(a \wedge b|a \wedge \phi) = p(b|\phi). \tag{2}$$

In other words, since the probability for a measurement on one system depends only on its pure state, the Born rule requires that the measurement of one system is independent from the preparation of the other. We call this property "statistical independence" [30]. It characterizes the map M, since $M(\underline{a},\underline{b})$ corresponds to the composite state where A and B are prepared in the states $|a\rangle$ and $|b\rangle$. Define $M_b(\underline{a}) = M(\underline{a},\underline{b})$. From the Born rule we find

$$|\langle M(\underline{a}, \underline{b})|M(\underline{\psi}, \underline{b})\rangle_{\mathcal{C}}|^{2} = |\langle M_{b}(\underline{a})|M_{b}(\underline{\psi})\rangle_{\mathcal{C}}|^{2}$$
$$= |\langle a|\psi\rangle_{\mathcal{A}}|^{2}, \tag{3}$$

where the first and second terms contain the inner product in the composite space \mathcal{C} . [This is not a new assumption: it follows from the measurement postulate (b) for the composite system.] This means that, when one subsystem is prepared in an eigenstate of what is measured there, the state space of the other is mapped preserving the square of the inner product. This implies orthogonality and the hierarchy of subspaces are preserved through M_b , making M_b a collinear transformation by definition. Geometrically, recall that M_b maps rays to rays. The fact that M_b is collinear means that it also maps higher order subspaces to

higher order subspaces (lines to lines, planes to planes, and so on) while preserving inclusion (if a line is within a plane, the mapped line will be within the mapped plane). In this case, the fundamental theorem of projective geometry [31] applies, which tells us that a unique semilinear map m_b that acts on the vectors exists in accordance with M_b . Moreover, conservation of probability further constrains it to be either linear or antilinear. This tells us that the corresponding m is either linear or antilinear in the first argument. Namely, if Eq. (3) holds then

$$\langle a|\psi\rangle = \langle m(a,b)|m(\psi,b)\rangle$$
 (4)

or
$$\langle a|\psi\rangle = \langle m(\psi,b)|m(a,b)\rangle.$$
 (5)

In this setting, the antilinear case (5) corresponds to a change of convention (much like a change of sign in the symplectic form for classical mechanics) and can be ignored. Given a Hilbert space, in fact, we can imagine replacing all vectors and all the operators with their Hermitian conjugate, mapping vectors into duals $|\psi\rangle^{\dagger} = \langle \psi|$. These changes would effectively cancel out leaving the physics unchanged: the two equations $A|w\rangle = B|z\rangle$ and $\langle w|A^{\dagger} = \langle z|B^{\dagger}$ are equivalent. (For example, in his first papers Schrödinger used both signs in his equation: effectively writing two equivalent equations with complex-conjugate solutions [32]. Also Wigner pointed out this equivalence [33], p. 152). We can repeat the same analysis for the second argument of m to conclude that it is a bilinear map, condition (H2).

The last condition, span surjectivity (H3), follows directly from the definition of a composite system. Since it is composed *only* of the component systems, for any state c of the composite system, we must find at least one pair $|a\rangle$, $|b\rangle$ such that $p(a \wedge b|c) \neq 0$. Span surjectivity follows: namely, the span of the map applied to all states in the component systems spans the composite system state space. In other words, the composite does not contain states that are totally independent of (i.e., orthogonal to) the states of the components.

We have obtained the conditions H1, H2, and H3 from the state postulate (a), the measurement postulate (b), and the definitions of composite and independent systems. We now prove that these three conditions imply that the (up to now unspecified) composition rule m is the tensor product. More precisely, given a total, span-surjective, bilinear map $m: \mathcal{A} \times \mathcal{B} \to \mathcal{C}$ that maps the Hilbert spaces \mathcal{A} , \mathcal{B} of the components into the Hilbert space \mathcal{C} of the composite and that preserves the square of the inner product, we find that \mathcal{C} is equivalent to $\mathcal{A} \otimes \mathcal{B}$ and that $m = \otimes$.

Proof.—Step 1: the bases of the component systems are mapped to a basis of the composite system. Because of totality property (H1) and because the square of the inner product is preserved, we can conclude that, given two orthonormal bases $\{|a_i\rangle\} \in \mathcal{A}$ and $\{|b_i\rangle\} \in \mathcal{B}$,

 $|\langle m(a_i,b_j)|m(a_k,b_\ell)\rangle|^2 = \delta_{ik}\delta_{j\ell}$, namely, $\{|m(a_i,b_j)\rangle\}$ is an orthonormal set in \mathcal{C} . Moreover, the surjectivity property (H3) guarantees that in \mathcal{C} no vectors are orthogonal to this set. This implies that it is a basis for \mathcal{C} .

Step 2: use the universal property. The tensor product is uniquely characterized, up to isomorphism, by a universal property regarding bilinear maps: given two vector spaces \mathcal{A} and \mathcal{B} , the tensor product $\mathcal{A} \otimes \mathcal{B}$ and the associated bilinear map $T: \mathcal{A} \times \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$ have the property than any bilinear map $m: \mathcal{A} \times \mathcal{B} \to \mathcal{C}$ factors through T uniquely. This means that there exists a *unique I*, dependent on m, such that $I \circ T = m$. In other words, the following diagram commutes:

$$\mathcal{A} \times \mathcal{B} \xrightarrow{T} \mathcal{A} \otimes \mathcal{B}$$

$$\downarrow^{I}$$

$$C$$

Since $m: \mathcal{A} \times \mathcal{B} \to \mathcal{C}$ is a bilinear operator (property H2), thanks to the universal property of the tensor product we can find a unique linear operator $I: \mathcal{A} \otimes \mathcal{B} \to \mathcal{C}$ such that $m(a,b) = I(a \otimes b)$. The set $\{I(a_i \otimes b_j) \text{ with } |a_i\rangle \text{ and } |b_j\rangle$ orthonormal bases for \mathcal{A} and $\mathcal{B}\}$ forms a basis for \mathcal{C} , since $I(a_i \otimes b_j) = m(a_i,b_j)$ and we have shown above that the latter is a basis. Thus,

$$\langle I(a_i \otimes b_j) | I(a_k \otimes b_\ell) \rangle_{\mathcal{C}} = \langle m(a_i, b_j) | m(a_k, b_\ell) \rangle_{\mathcal{C}}$$

= $\delta_{ik} \delta_{j\ell} = \langle a_i \otimes b_j | a_k \otimes b_\ell \rangle_{\otimes},$ (6)

where we used the orthonormality of the bases and the fact that $|a_i \otimes b_j\rangle$ is a basis of the tensor product space $\mathcal{A} \otimes \mathcal{B}$. Since the function I is a linear function that maps an orthonormal basis of $\mathcal{A} \otimes \mathcal{B}$ to an orthonormal basis of \mathcal{C} , I is a an isomorphism (a bijection that preserves the mathematical structure) between $\mathcal{A} \otimes \mathcal{B}$ and \mathcal{C} . As $\mathcal{C} \cong \mathcal{A} \otimes \mathcal{B}$ are isomorphic as Hilbert spaces, they are mathematically equivalent: $c \in \mathcal{C}$ and $I^{-1}(c)$ represent the same physical object. In this sense, we can loosely say that I is the identity, as it connects spaces that are physically equivalent. So we can directly use the tensor product to represent the composite state space. This means that the map $m: \mathcal{A} \times \mathcal{B} \to \mathcal{C}$ is equivalent to the map $\otimes: \mathcal{A} \times \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$ in the sense that $m \circ I^{-1} = \otimes$.

A few comments on the proof: it is based on the universal property of the tensor product, which uniquely characterizes it. In step 1 we show that the bilinear map m maps subsystems' bases into the composite system basis. We also know that there exists a tensor product map $T = \otimes$ that can compose the vectors in \mathcal{A} and \mathcal{B} . In step 2 we use the universal property: since m is a bilinear map, we are assured that there exists a unique I such that $I \circ T = m$. Since we show that I is an isomorphism, then I bijectively maps vectors in \mathcal{C} onto vectors in the tensor product space. Namely, $m = T = \otimes$.

We conclude with some general comments. The tensor product structure of quantum systems is not absolute, but depends on the observables that are accessible [26,27]. This is due to the fact that an agent that has access to a set of observables will define quantum systems differently from an agent that has access to a different set of observables. Where one agent sees a single system, an agent that has access to less refined observables (and is then limited by some superselection rules) can consider the same system as composed of multiple subsystems.

It has been pointed out before that the quantum postulates are redundant: in Refs. [9,34] it was shown that the measurement postulate (b) can be derived from the others (a), (c), (d). Here instead we have shown how the tensor product postulate (c) can be logically derived from the state postulate (a), the measurement postulate (b), and a reasonable definition of independent systems, and we have described the logical relations among them. Of course, we do not claim that this is the *only* way to obtain the tensor product postulate from the others.

L. M. acknowledges useful discussions with M. Ozawa, P. Zanardi, S. Lloyd, D. Zeh, G. Auletta, A. Aldeni, and funding from the MIUR Dipartimenti di Eccellenza 2018-2022 Project No. F11I18000680001, Attract project through the Eu Horizon 2020 research and innovation programme under Grant Agreement No. 777222. This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Information Science Research Centers. Ouantum Superconducting Quantum Materials and Systems Center (SQMS) under the contract No. DE-AC02-07CH11359. G. C. and C. A. A. would like to thank M. J. Greenfield for reviewing the mathematical details and acknowledge funding from the MCubed program of the University of Michigan. G. C. and C. A. A.'s contribution to this work is part of a larger project, Assumptions of Physics [35], which aims to identify a handful of physical principles from which the basic laws can be rigorously derived.

^[1] W. H. Zurek, Quantum Darwinism, Nat. Phys. 5, 181 (2009).

^[2] For example, in the tensor product $a \otimes (b+c) = a \otimes b + a \otimes c$ while in the direct product $a \times (b+c) = a \times b + 0 \times c$, where 0 is the zero vector. Also, in the tensor product $r(a \otimes b) = (ra) \otimes b = a \otimes (rb)$ while in the direct product $r(a \times b) = ra \times rb$, where r is a scalar.

^[3] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1966).

^[4] J. von Neumann, Mathematical Foundations of Quantum Mechanics (Princeton University Press, Princeton, NJ, 1955).

^[5] J. M. Jauch, in *Foundations of Quantum Mechanics* (Addison-Welsey, Reading, MA, 1968), p. 176.

^[6] H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig, 1928); translated by H. P. Robertson, in *The Theory*

- of Groups and Quantum Mechanics (Methuen, London, 1931); reprinted by Dover, p. 91.
- [7] A. Einstein, B. Podolsky, and N. Rosen, Can quantummechanical description of physical reality be considered complete?, Phys. Rev. 47, 777 (1935).
- [8] M. Ozawa, Uncertainty relations for noise and disturbance in generalized quantum measurements, Ann. Phys. (Amsterdam) 311, 350 (2004).
- [9] L. Masanes, T. D. Galley, and M. P. Müller, The measurement postulates of quantum mechanics are operationally redundant, Nat. Commun. 10, 1361 (2019).
- [10] W. K. Wootters, Optimal information transfer and real-vector-space quantum theory, edited by G. Chiribella and R. Spekkens, in *Quantum Theory: Informational Foundations and Foils*, Fundamental Theories of Physics Vol. 181 (Springer, Dordrecht, 2016).
- [11] M. A. Nielsen and I. L. Chuang, *Quantum Computation* and *Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
- [12] L. E. Ballentine, Quantum Mechanics, a Modern Development (World Scientific, Singapore, 2014).
- [13] L. E. Ballentine, Probability theory in quantum mechanics, Am. J. Phys. 54, 883 (1986).
- [14] A. Peres, Classical interventions in quantum systems. II. Relativistic invariance, Phys. Rev. A 61, 022117 (2000).
- [15] T. Matolcsi, Tensor product of Hilbert lattices and free orthodistributive product of orthomodular lattices, Acta Sci. Math. (Szeged) 37, 263 (1975).
- [16] F. M. Ciaglia, A. Ibort, and G. Marmo, On the notion of composite system, edited by F. Nielsen and F. Barbaresco, in *Geometric Science of Information*, Lecture Notes in Computer Science Vol. 11712 (Springer, New York, 2019).
- [17] D. Aerts and I. Daubechies, Physical justification for using the tensor product to describe two quantum systems as one joint system, Helv. Phys. Acta 51, 661 (1979).
- [18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.126.110402 for the mathematical details of the proof. It includes Refs. [19–24].
- [19] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).

- [20] S. B. Giddings, Hilbert space structure in quantum gravity: An algebraic perspective, J. High Energy Phys. 12 (2015) 001.
- [21] H. Roos, Independence of local algebras in quantum field theory, Commun. Math. Phys. **16**, 238 (1970).
- [22] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, MIP*=RE, arXiv:2001.04383.
- [23] P. Teller, An Interpretive Introduction to Quantum Field Theory (Princeton University Press, Princeton, NJ, 1997).
- [24] M. Redhead and P. Teller, Particles, particle labels, and quanta: The toll of unacknowledged metaphysics, Found. Phys. 21, 43 (1991).
- [25] Y. Aharonov and L. Susskind, Charge superselection rule, Phys. Rev. 155, 1428 (1967).
- [26] P. Zanardi, Virtual Quantum Subsystems, Phys. Rev. Lett. 87, 077901 (2001).
- [27] P. Zanardi, D. A. Lidar, and S. Lloyd, Quantum Tensor Product Structures are Observable Induced, Phys. Rev. Lett. 92, 060402 (2004).
- [28] We emphasize that the kinematic independence is inequivalent to dynamical independence (or isolation). Indeed if two systems interact, their interaction may lead to dynamical restrictions in the state spaces. Here we will not consider dynamical evolution, which is contained in postulate (d).
- [29] R. T. Cox, The Algebra of Probable Inference (J. Hopkins Press, 1961).
- [30] One can also prove that the measurements on the components are independent as well (see Supplemental Material [18]), but we only strictly need preparation here.
- [31] E. Artin, Geometric Algebra (Interscience Publishers Inc., New York, 1957).
- [32] E. Schrödinger, Ann. Phys. (Berlin) **102**, 81 (1926); English translation in E. Schrödinger, *Collected papers on Wave Mechanics* (Blackie & Son, London, 1928).
- [33] E. Wigner, On unitary representations of the inhomogeneous Lorentz group, Ann. Math. 40, 149 (1939).
- [34] J.B. Hartle, Quantum mechanics of individual systems, Am. J. Phys. **36**, 704 (1968).
- [35] https://assumptionsofphysics.org/.