Open Peer Review on Qeios

The Born Rule is a Feature of Superposition

David Ellerman[1]

1 University of Ljubljana

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

The Born Rule plays a critical role in quantum mechanics (QM) since it supplies the link between the mathematical formalism and experimental results in terms of probabilities. The Born Rule does not occur in ordinary probability theory. Where then does it come from? This has been a topic of considerable controversy in the literature. We take the approach of asking what is the simplest extension of ordinary probability theory where the Born rule appears. This is answered by showing that the Born Rule appears by adding the notion of superposition events (in addition to the ordinary discrete events) to finite probability theory. Hence the rule does not need any physics-based derivation. It is simply a feature of the mathematics of superposition when only superposition events are added to ordinary probability theory.

David Ellerman

University of Ljubljana, Slovenia

david@ellerman.org

https://orcid.org/0000-0002-5718-618X

Keywords: Born Rule, superposition, density matrices, finite probability theory.

1. Introduction

In quantum mechanics (QM), the Born Rule provides the all-important link between the mathematical formalism (e.g., the wave function) and experimental results in terms of probabilities. The rule does not occur in ordinary classical probability theory. Where does the Born Rule come from? Can it be derived from the other postulates of QM or must it be assumed as an additional postulate? There is a vast and sophisticated literature debating these questions--see [1] and [2] and the articles cited therein.

In this paper, a different approach is taken. What is the simplest extension to classical probability theory where the Born Rule appears? We expand ordinary finite probability theory by introducing superposition events in addition to the usual discrete events (subsets of the outcome space) and then we show that the Born Rule naturally arises in the mathematics of superposition events. A superposition event is a purely mathematical notion--although obviously inspired by the notion of a superposition state in quantum mechanics. As a mathematical notion, it could have been (but was not) introduced centuries before QM. The thesis is that the Born Rule is not a bug that needs to be "explained" or "justified"; it is just a feature of the notion of a superposition event in this minimally expanded probability theory.

2. Superposition events

In classical finite probability theory, the *outcome*} (or *sample*) *space* is a set $U = \{u_1, \ldots, u_n\}$ with point probabilities $p = (p_1, \ldots, p_n)$. An (ordinary) *event S* is a non-empty subset $S \subseteq U$. In an (ordinary) event S, the atomic outcomes or elements of S are considered as perfectly discrete and distinguished from each other; in each run of the "experiment" or trial, there is the probability $\Pr(S)$ occuring and the probability $\Pr(T|S)$ of an event $T \subseteq U$ occurring given the S occurs (including the case of a specific outcome $T = \{u_i\}$).

The intuitive idea of the corresponding superposition state, denoted ΣS, is that the outcomes in the state are not distinguished but are blobbed or cohered together as an indefinite event. In each run of the "experiment" or trial conditioned on ΣS, the indefinite state is sharpened to a less indefinite state which is maximally sharpened to one of the definite outcomes in S. In the case of a singleton event $S = \{u_i\}$, the ordinary event $S = \{u_i\}$ is the same as the superposition event $\Sigma S = \Sigma\{u_i\} = \{u_i\} = S$.

For a suggestive visual example, consider the outcome set U as a pair of isosceles triangles that are distinct by the labels on the equal sides and the opposing angles.

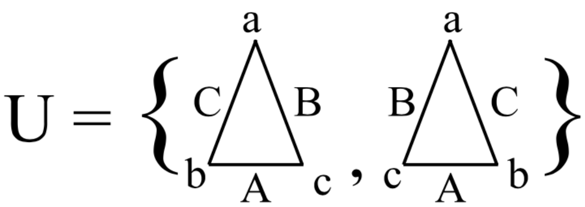

Figure 1. Set of distinct isosceles triangles

The superposition event ΣU is definite on the properties that are common to the elements of U, i.e., the angle a and the opposing side A, but is indefinite where the two triangles are distinct, i.e., the two equal sides and their opposing angles are not distinguished by labels [3].

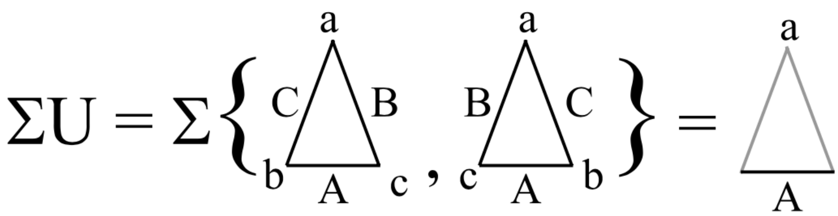

Figure 2. The superposition event ΣU.

It might be noted that this notion of superposition and the notion of abstraction are essentially flip-side viewpoints of the same idea of extracting from a set an entity that is definite on the commonalities of the elements of the set and indefinite on where the elements differ [4]. The two flip-side viewpoints are like seeing a glass half-empty (superposition) or seeing a glass half-full (abstraction).

What is a mathematical model that will distinguish between the ordinary event S and the superposition event ΣS? Using n-ary column vectors in $\mathbb{R}^n$, the ordinary event S could be represented by the column vector, denoted $|S\rangle$, with the i^{th} entry $\chi_S\left(u_i\right)$, where $\chi_S : U \to \{0, 1\}$ is the characteristic function for S, i.e., $\chi_S(u_i) = 1$ if $u_i \in S$, else 0. But to represent the superposition event ΣS we need to add a dimension to use two-dimensional $n \times n$ matrices to represent the blobbing together or cohering of the elements of S in the superposition even ΣS.

An *incidence matrix* for a binary relation $R \subseteq U \times U$ is the $n \times n$ matrix $\mathrm{In}(R)$ where $\mathrm{In}(R)_{jk} = 1$ if $\left(u_j, u_k\right) \in R$, else 0. The diagonal ΔS is the binary relation consisting of the ordered pairs $\left\{\left(u_i, u_i\right) : u_i \in S\right\}$ and its incidence matrix $\mathrm{In}(\Delta S)$ is the diagonal matrix with the diagonal elements $\chi_S\left(u_i\right)$. The superposition state ΣS could then be represented as $\mathrm{In}(S \times S)$, the incidence matrix of the binary relation $S \times S \subseteq U \times U$, where the non-zero off-diagonal elements represent the equating, cohering, or blobbing together of the corresponding diagonal elements.[1] Given two column vectors $|s\rangle = \left(s_1, \ldots, s_n\right)^t$ and $|t\rangle = \left(t_1, \ldots, t_n\right)^t$ in $\mathbb{R}^n$ (where $()^t$ is the transpose), their *inner product* is the sum of the products of the corresponding entries and is denoted $\langle t|s\rangle = (|t\rangle)^t|s\rangle = \sum_{i=1}^n t_i s_i$. Their *outer product* is the $n \times n$ matrix denoted as

$|s\rangle\langle t| = |s\rangle(|t\rangle)^t$. A vector $|s\rangle$ is *normalized* if $\langle s|s\rangle = 1$. That incidence matrix $\ln(S \times S)$ could be constructed as the outer product $|S\rangle(|S\rangle)^t = |S\rangle\langle S| = \ln(S \times S)$.

If we divided $\ln(\Delta S)$ and $\ln(S \times S)$ through by their trace (sum of diagonal elements) $|S|$, then we obtain two density matrices $\rho(S) = \frac{\ln(\Delta S)}{|S|}$ and $\rho(\Sigma S) = \frac{\ln(S \times S)}{|S|} = \frac{1}{\sqrt{|S|}}|S\rangle\langle S|\frac{1}{\sqrt{|S|}}$ over the reals $\mathbb{R}$. In general, a *density matrix* ρ over the reals $\mathbb{R}$ (or the complex numbers $\mathbb{C}$) is a symmetric matrix $\rho = \rho^t$ (or conjugate symmetric matrix $\rho = \left(\rho^*\right)^t$ in the case of $\mathbb{C}$) with trace $\operatorname{tr}[\rho] = 1$ and all non-negative eigenvalues. A density matrix ρ is *pure* if $\rho^2 = \rho$, otherwise a *mixture*.

Consider the partition $\pi = \{B_1, B_2\} = \{\{\diamondsuit, \heartsuit\}, \{\clubsuit, \spadesuit\}\}$ on the outcome set $U = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$ with equiprobable outcomes like drawing cards from a randomized deck. For instance, the superposition event associated with $B_1 = \{\diamondsuit, \heartsuit\}$, has a pure density matrix since (rows and columns labelled in the order $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$):

$$\rho\left(\Sigma B_1\right) = \frac{1}{\sqrt{|B_1|}}\begin{bmatrix}0\\1\\1\\0\end{bmatrix}\begin{bmatrix}0 & 1 & 1 & 0\end{bmatrix}\frac{1}{\sqrt{|B_1|}} = \frac{1}{|B_1|}\begin{bmatrix}0 & 0 & 0 & 0\\0 & 1 & 1 & 0\\0 & 1 & 1 & 0\\0 & 0 & 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0\\0 & \frac{1}{2} & \frac{1}{2} & 0\\0 & \frac{1}{2} & \frac{1}{2} & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

equals its square, but density matrix for the discrete set B_1:

$$\rho\left(B_1\right) = \begin{bmatrix}0 & 0 & 0 & 0\\0 & \frac{1}{2} & 0 & 0\\0 & 0 & \frac{1}{2} & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

is a mixture since it does not equal its square.

Intuitively, the interpretation of the superposition event represented by $\rho\left(\Sigma B_1\right) = \rho(\Sigma\{\diamondsuit, \heartsuit\})$ is that it is definite on the properties common to its elements, e.g., in this case, being a red suite, but indefinite on where the elements differ. The indefiniteness is indicated by the non-zero off-diagonal elements that indicate that the diamond suite $\diamondsuit$ is blurred, cohered, or superposed with the hearts suite $\heartsuit$ in the superposition state $\Sigma\{\diamondsuit, \heartsuit\}$.

The next step is to bring in the point probabilities $p = \left(p_1, \ldots, p_n\right)$ where those two real density matrices $\rho(S)$ and $\rho(\Sigma S)$ defined so far correspond to the special case of the equiprobable distribution on S with 0 probabilities outside of S.

3. Density matrices with general probability distributions

Let the outcome space $U = \left\{u_1, \ldots, u_n\right\}$ have the strictly positive probabilities $p = \left\{p_1, \ldots, p_n\right\}$. The probability of a (discrete) subset S is $\Pr(S) = \sum_{u_i \in S} p_i$ and the conditional probability of $T \subseteq U$ given S is: $\Pr(T|S) = \frac{\Pr(T \cap S)}{\Pr(S)}$. But we have now reformulated both the usual discrete event S and the new superposition event ΣS in matrix terms. Hence we need to reformulate the usual conditional probability calculation in matrix terms and then apply the same matix operations to define the conditional probabilities for the superposition events.

The density matrix $\rho(U)$ is the diagonal matrix with the point probabilties down the diagonal. Let P_S be the diagonal (projection) matrix with the diagonal entries $\chi_S\left(u_i\right)$. Then $\Pr(S)$ can be computed by replacing the summation $\sum_{u_i \in S} p_i$ with the trace formula: $\Pr(S) = \mathrm{tr}[P_S \rho(U)]$. The density matrix $\rho(S)$ for S is defined as the diagonal matrix with diagonal entries $\frac{p_i}{\Pr(S)}$ if $u_i \in S$, else 0, which yields the mixture density matrix $\rho(S)$ (aside from the case of a singleton $S = \left\{u_i\right\}$).

For $\rho(S)$, the eigenvalues are just the conditional probabilities $\Pr\left(\left\{u_i\right\}|S\right) = \frac{\Pr\left(\left\{u_i\right\} \cap S\right)}{\Pr(S)} = \frac{p_i}{\Pr(S)} \chi_S\left(u_i\right)$ for $i = 1, \ldots, n$. Then the conditional probability $\Pr(T|S)$ is reproduced in the matrix format as:

$$\Pr(T|S) = \mathrm{tr}\left[P_T \rho(S)\right].$$

The previously constructed density matrix $\rho(\Sigma S) = \frac{1}{\sqrt{|S|}}|S\rangle\langle S|\frac{1}{\sqrt{|S|}}$ was for the special case of equiprobable outcomes. In the general case of point probabilities, the column vector $\frac{1}{\sqrt{|S|}}|S\rangle$ is generalized to $|s\rangle$ where the i^{th} entry, symbolized $\left\langle u_i|s\right\rangle$, is $\sqrt{\frac{p_i}{\Pr(S)}}$ if $u_i \in S$, else 0, and then $\rho(\Sigma S) = |s\rangle\langle s|$ which is a pure density matrix. For the pure density matrix $\rho(\Sigma S)$, there is one eigenvalue of 1 with the rest of the eigenvalues being zeros (since the sum of the eigenvalues is the trace). Given just $\rho(\Sigma S)$, the vector $|s\rangle$ is recovered as the normalized eigenvector associated with the eigenvalue of 1 and $\rho(\Sigma S) = |s\rangle\langle s|$.[2]

Then applying the same matrix operations to get probabilities as for discrete events, we have $\Pr(\Sigma S) = \mathrm{tr}\left[P_S \rho(\Sigma U)\right]$ and:

$$\Pr(\Sigma T|\Sigma S) = \mathrm{tr}\left[P_T \rho(\Sigma S)\right].$$

It is the interpretation, not the probabilities, that are different for the two types of events. For discrete events, the given discrete event S is reduced by conditionizing to the discrete event $T \cap S$. For superposition events, the given superposition event ΣS is sharpened (i.e., made less indefinite) to the superposition event $\Sigma(T \cap S)$ with the probability $\Pr(\Sigma(T \cap S)|\Sigma S) = \Pr(\Sigma T|\Sigma S)$ given the event ΣS.

A *partition* $\pi = \left\{B_1, \ldots, B_m\right\}$ on U is a set of non-empty subsets, called *blocks*, $B_j \subseteq U$ that are disjoint and whose union is U. Taking each block $B_j = S$, then there is the normalized column vector $|b_j\rangle$ whose i^{th} entry is $\sqrt{\frac{p_i}{\Pr(B_j)}}\chi_{B_j}(u_i)$ and the density matrix $\rho(\Sigma B_j) = |b_j\rangle\langle b_j|$ for the superposition subset ΣB_j. Then the density matrix $\rho(\pi)$ for the partition π is just the probability sum of those pure density matrices for the superposition blocks:

$$\rho(\pi) = \sum_{j=1}^{m} \Pr(B_j)\rho(\Sigma B_j).$$

The eigenvalues for $\rho(\pi)$ are the m probabilities $\Pr(B_j)$ with the remaining $n-m$ values of 0.

Given two partitions $\pi = \{B_1, \ldots, B_m\}$ and $\sigma = \{C_1, \ldots, C_{m'}\}$, the partition π *refines* the partition σ, written $\sigma \precsim \pi$, if for each block $B_j \in \pi$, there is a block $C_{j'} \in \sigma$ such that $B_j \subseteq C_{j'}$. The partitions on U form a partial order under refinement. The maximal partition or top of the order is the *discrete partition* $\mathbf{1}_U = \{\{u_i\}\}_{i=1}^{n}$ where all the blocks are singletons and the minimal partition or bottom is the *indiscrete partition* $\mathbf{0}_U = \{U\}$ with only one block U. Then the density matrices for these top and bottom partitions are just the density matrices for the discrete set U and the superposition set ΣU:

$$\rho(\mathbf{1}_U) = \rho(U) \text{ and } \rho(\mathbf{0}_U) = \rho(\Sigma U).$$

Let us illustrate this result with the case of flipping a fair coin. The classical set of outcomes $U = \{H, T\}$ is represented by the density matrix:

$$\rho(U) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

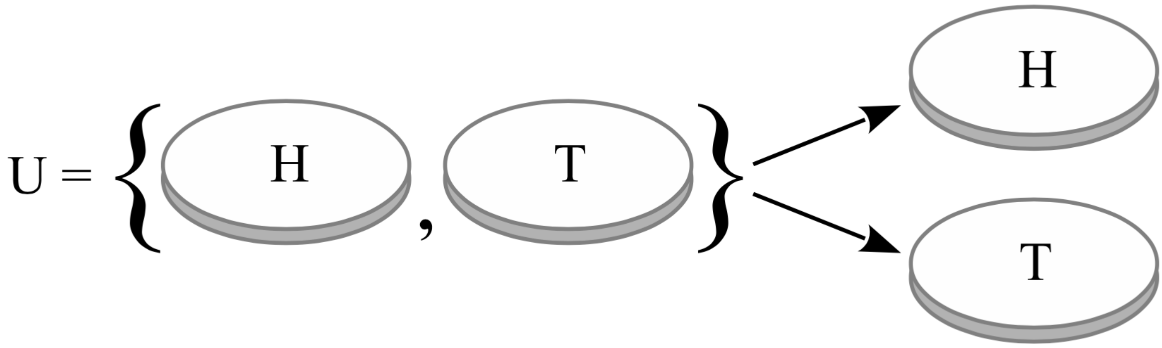

Figure 3. Classical event: trial picks out heads or tails

The superposition event ΣU, that blends or superposes heads and tails, is represented by the density matrix:

$$\rho(\Sigma U) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

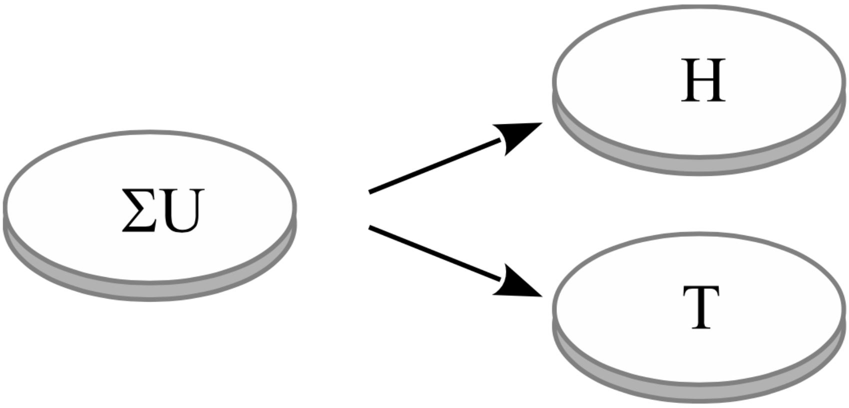

Figure 4. Superposition event: Trial sharpens to heads or tails.

The probability of getting heads in each case is:

$$\Pr\left(H|\rho(U)\right) = \operatorname{tr}\left[P_{\{H\}}\rho(U)\right] = \operatorname{tr}\left[\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right] = \operatorname{tr}\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{2}$$

$$\Pr\left(H|\rho(\Sigma U)\right) = \operatorname{tr}\left[P_{\{H\}}\rho(\Sigma U)\right] = \operatorname{tr}\left[\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right] = \operatorname{tr}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{bmatrix} = \frac{1}{2}$$

and similarly for tails. Thus the two conditioning events U and ΣU cannot be distinguished by performing an experiment or trial that distinguishes heads and tails. This is a feature, not a bug, since the same thing occurs in quantum mechanics. For instance, a spin measurement along, say, the z-axis of an electron cannot distinguish between the superposition state $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ with a density matrix like $\rho(\Sigma U)$ and a statistical mixture of half electrons with spin up and half with spin down with a density matrix like $\rho(U)$ [[5], p. 176].[3]

4. Conclusion: The Born Rule

The Born Rule does not occur in ordinary classical probability theory because that theory does not include superposition events. When superposition events are introduced into the purely mathematical theory, then the Born Rule naturally emerges as a feature of the mathematical treatment of superposition.

The pure density matrix $\rho(\Sigma S)$ can be constructed as the outer product $\rho(\Sigma S) = |s\rangle\langle s|$ where $|s\rangle$ is the n-ary column vector with the i^{th} entry as $\langle u_i | s \rangle = \sqrt{\frac{p_i}{\Pr(S)}}\chi_S(u_i) = \sqrt{\frac{\Pr(\{u_i\} \cap S)}{\Pr(S)}}$. Or starting with the pure density matrix $\rho(\Sigma S)$, then $|s\rangle$ so that $\rho(\Sigma S) = |s\rangle\langle s|$ is obtained as the normalized eigenvector associated with the eigenvalue of 1.

The probability of u_i conditioned on the superposition event ΣS is:

$$\Pr\left(\Sigma\{u_i\} | \Sigma S\right) = \mathrm{tr}\left[P_{\{u_i\}}\rho(\Sigma S)\right] = \frac{\Pr\left(\{u_i\} \cap S\right)}{\Pr(S)}.$$

The point is that this same probability conditioned by the two-dimensional density matrix $\rho(\Sigma S)$ could also be obtained from the one-dimensional vector $|s\rangle$ as (where $\Sigma\{u_i\} = \{u_i\}$):

$$\langle u_i | s \rangle^2 = \frac{p_i}{\Pr(S)}\chi_S(u_i) = \Pr\left(\Sigma\{u_i\} | \Sigma S\right) = \mathrm{tr}\left[P_{\{u_i\}}\rho(\Sigma S)\right]. \text{ The Born Rule}$$

The Born Rule does not occur in classical finite probability theory since the events S are all discrete sets that can be represented by n-ary column vectors. The associated two-dimensional diagonal density matrix $\rho(S)$ is not the outer product of a one-dimensional vector with itself (except when S is a singleton). To accommodate the notion of a superposition event ΣS, it is necessary to use two-dimensional density matrices $\rho(\Sigma S)$ where the non-zero off-diagonal elements indicate the blobbing or cohering together in superposition of the elements associated with the corresponding diagonal entries. And mathematically *those* density matrices $\rho(\Sigma S)$ can be constructed as the outer product $|s\rangle(|s\rangle)^t = |s\rangle\langle s|$ of a one-dimensional vector $|s\rangle$ with itself. Then the probability of the individual outcomes u_i conditioned by the superposition event ΣS is given by the Born Rule: $\Pr\left(\Sigma\{u_i\} | \Sigma S\right) = \langle u_i | s \rangle^2$. Thus the Born Rule arises naturally out of the mathematics of probability theory enriched by superposition events.[4] It does not need any more-exotic or physics-based explanation. No physics was used in the making of this paper. The Born Rule is just a feature of the *mathematics* of superposition.

5. Statements and Declarations

The author did not receive support from any organization for the submitted work. The author has no relevant financial or non-financial interests to disclose. There are no conflicts of interest to disclose.

Footnotes

[1]On the universe set U, the binary relation $U \times U$ is the universal equivalence relation which equates all the elements of U.

Thus $S \times S$ is the universal equivalence relation on S which equates all its elements.

[2]This is by the spectral decomposition of that real density matrix as a Hermitian operator.

[3]In QM, they can only be distinguished by measurement in a different observable basis.

[4]In the vector spaces over the complex numbers $\mathbb{C}$ of quantum mechanics, the square $\langle u_i | s \rangle^2$ is the absolute square $\left| \langle u_i | s \rangle \right|^2$.

References

1. ^*Masanes, Lluís, Thomas D. Galley, and Markus P. Müller. 2019. "The Measurement Postulates of Quantum Mechanics Are Operationally Redundant." Nature Communications 10 (1): 1361. https://doi.org/10.1038/s41467-019-09348-x.*
2. ^*Vaidman, Lev. 2020. "Derivations of the Born Rule." In Quantum, Probability, Logic, edited by Meir Hemmo and Orly Shenker, 567--84. Cham, Switzerland: SpringerNature. https://doi.org/10.1007/978-3-030-34316-3.*
3. ^*Ellerman, David. 2022. "Follow the Math!: The Mathematics of Quantum Mechanics as the Mathematics of Set Partitions Linearized to (Hilbert) Vector Spaces." Foundations of Physics 52 (5). https://doi.org/10.1007/s10701-022-00608-3.*
4. ^*Ellerman, David. 2021. "On Abstraction in Mathematics and Indefiniteness in Quantum Mechanics." Journal of Philosophical Logic} 50 (4): 813--35. https://doi.org/10.1007/s10992-020-09586-1.*
5. ^*Auletta, Gennaro, Mauro Fortunato, and Giorgio Parisi. 2009. Quantum Mechanics. Cambridge UK: Cambridge University Press.*