Born Again! The Born Rule as a feature of superposition

David Ellerman
Independent Researcher, Ljubljana, Slovenia
david@ellerman.org
https://orcid.org/0000-0002-5718-618X

September 13, 2024

Abstract

Where does the Born Rule come from? We ask what is the simplest extension of probability theory where the Born rule appears. This is answered by showing that the Born Rule first appears as the square of the (normalized) vector components of the notion of superposition events in the enriched probability theory. The Superposition Principle requires the states to be represented as (normalized) vectors in a vector space with positive, negative, or complex components–so the rule of getting probabilities as the (absolute) squares of normalized vector components generalizes from the simple case in the enriched probability theory.

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

Contents

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. Hence one might ask with Steven Weinberg: "So where does the Born Rule come from?" [15, p. 92] 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 [11], [14], [10], 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 in this enriched probability theory–although obviously inspired by the notion of a superposition state in quantum mechanics.

It is not a coincidence that superposition (including the special case of entanglement) is the key non-classical notion in quantum mechanics.

- For instance, "*superposition*, with the attendant riddles of entanglement and reduction, remains *the* central and generic interpretative problem of quantum theory." [3, p. 27]
- Some writers use the word "entanglement" to mean or include superposition.[1] "The superposition or 'entanglement' of states is a hallmark of quantum mechanics." [2, p. 50]
- "In this sense, one can say that the entanglement arising from summation of probability amplitudes over all possible Feynman paths in the appropriate configuration space is *the* distinctive feature of quantum mechanics, the sole mystery." [13, p. 248]

Dirac was quite clear on this point from the beginning.

> The nature of the relationships which the superposition principle requires to exist between the states of any system is of a kind that cannot be explained in terms of familiar physical concepts. One cannot in the classical sense picture a system being partly in each of two states and see the equivalence of this to the system being completely in some other state. There is an entirely new idea involved, to which one must get accustomed and in terms of which one must proceed to build up an exact mathematical theory, without having any detailed classical picture. [4, p. 12]

As a purely mathematical notion (as developed here), superposition events could have been (but were not) introduced long 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 vectorial mathematics of a superposition event foreshadowed in this minimally expanded probability theory–and then extended to the Hilbert spaces over $\mathbb{C}$ in QM.

2 Intuitively modeling superposition events

In classical finite probability theory, the *outcome* (or *sample*) *space* is a set $U = \{u_1, ..., u_n\}$ (where we assume equal probabilities until different point probabilities are introduced). 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 $\Pr(S)$ occurring 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 from each other but are blobbed or cohered together as an indefinite event. The concept of superposition includes the notion of *amplitudes* as the relative 'strength' of the outcomes in the superposition. It is the rules for dealing with amplitudes that separates quantum probabilities from ordinary probabilities.

> In the two-slit experiment, for example, passage through one slit or the other is only a distinguishable alternative if a counter is placed behind one of the slits; without such a counter, these are indistinguishable alternatives. Classical probability rules apply to distinguishable processes. Nonclassical probability amplitude rules apply to indistinguishable processes. [12, p. 314]

[1] The argument by some that it is only entanglement proper that is characteristic of QM, since there is superposition in classical electromagnetic waves or in water waves, will be addressed below.

Hence we are considering the minimal extension of classical probability theory that includes superposition events and *thus* also the notion of amplitudes in the superposition. No physics is involved in this extension; we are only investigating what emerges naturally from the mathematics with these concepts introduced into otherwise classical probability theory.

In each run of the "experiment" or trial conditioned on ΣS, the indefinite event is sharpened to a less indefinite event 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.

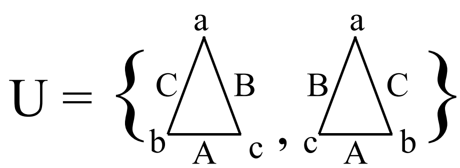

Figure 1: Set of distinct isosceles triangles

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 ([7]; [9]).

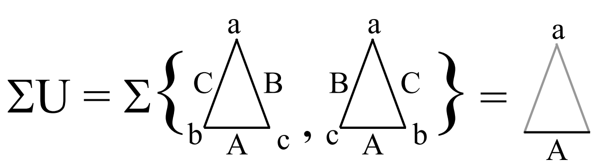

Figure 2: The superposition event ΣU.

For another visual example of superposition, consider a square with labelled corners $\boxed{\begin{smallmatrix} a & b \\ d & c \end{smallmatrix}}$. If it is flipped around the diagonal, then it becomes $\boxed{\begin{smallmatrix} a & d \\ b & c \end{smallmatrix}}$ and then the superposition is $\sum \left(\boxed{\begin{smallmatrix} a & b \\ d & c \end{smallmatrix}}, \boxed{\begin{smallmatrix} a & d \\ b & c \end{smallmatrix}} \right) = \boxed{\begin{smallmatrix} a & ? \\ ? & c \end{smallmatrix}}$. This sort of simple superposition (without magnitudes) is indefinite on where the superposed entities differ and definite on where they are the same.

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 (or silent) on where the elements differ [5]. The two flip-side viewpoints are like seeing a glass half-empty (superposition) or seeing a glass half-full (abstraction).

3 Mathematically modeling superposition events

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 $\chi_S(u_i)$, where $\chi_S : U \rightarrow \{0, 1\}$ is the characteristic function for S, i.e., $\chi_S(u_i) = 1$ if $u_i \in S$, else 0. That vector representation is insufficient to represent whether the elements of S are superposed or not. Hence 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 $(u_j, u_k) \in R$, else 0. The diagonal ΔS is the binary relation consisting of the ordered pairs $\{(u_i, u_i) : u_i \in S\}$ and its incidence matrix $\mathrm{In}(\Delta S)$ is the diagonal matrix with the diagonal elements $\chi_S(u_i)$. 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.[2]

Given two column vectors $|s\rangle = (s_1, ..., s_n)^t$ and $|t\rangle = (t_1, ..., t_n)^t$ 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 $\mathrm{In}(S \times S)$ could be constructed as the outer product $|S\rangle (|S\rangle)^t = |S\rangle \langle S| = \mathrm{In}(S \times S)$.

If we divided $\mathrm{In}(\Delta S)$ and $\mathrm{In}(S \times S)$ through by their trace (sum of diagonal elements) $|S|$, then we obtain two density matrices $\rho(S) = \frac{\mathrm{In}(\Delta S)}{|S|}$ and $\rho(\Sigma S) = \frac{\mathrm{In}(S \times S)}{|S|} = \frac{1}{\sqrt{|S|}} |S\rangle \langle S| \frac{1}{\sqrt{|S|}}$ 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 $\rho = (\rho^*)^t$ in the case of $\mathbb{C}$) with trace $\mathrm{tr}[\rho] = 1$ and all non-negative eigenvalues which sum to 1.

The analogue to a probability theory discrete event in QM is a completely discrete (or decomposed) mixed state. It is not a vector in Hilbert space. A vector in Hilbert space represents a pure state which is in general a superposition in a given basis and thus it is the analogue of a superposition event. One virtue of density matrices is that they represent both mixed and pure states.

A density matrix ρ is *pure* if $\rho^2 = \rho$, otherwise a *mixture*. The existence of the non-zero off-diagonal elements in the incidence matrices and thus in the density matrices indicates the presence of not only superposition but also amplitudes indicating the coherence of the superposed outcomes.

> For this reason, the off-diagonal terms of a density matrix ... are often called *quantum coherences* because they are responsible for the interference effects typical of quantum mechanics that are absent in classical dynamics. [1, p. 177]

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 $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$):

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

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

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

is a mixture since it does not equal its square.

Intuitively, the interpretation of the superposition event represented by $\rho(\Sigma B_1) = \rho(\Sigma \{\diamondsuit, \heartsuit\})$ 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

[2] 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.

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 $p=(p_1,...,p_n)$ 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.

4 Density matrices with general probability distributions

Let the outcome space $U=\{u_1,...,u_n\}$ have the strictly positive probabilities $p=\{p_1,...,p_n\}$. 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 matrix operations to define the conditional probabilities for the superposition events.

The density matrix $\rho(U)$ is the diagonal matrix with the point probabilities down the diagonal. Let P_S be the diagonal (projection) matrix with the diagonal entries $\chi_S(u_i)$. 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 the classical discrete 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 $\rho(S)$. For $\rho(S)$, the eigenvalues are just the conditional probabilities $\Pr(\{u_i\}|S)=\frac{\Pr(\{u_i\}\cap S)}{\Pr(S)}=\frac{p_i}{\Pr(S)}\chi_S(u_i)$ for $i=1,...,n$. 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 $\rho(\Sigma S)=\frac{1}{\sqrt{|S|}}|S\rangle\langle S|\frac{1}{\sqrt{|S|}}$ for the superposition event ΣS 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 $\langle u_i|s\rangle$, is the *amplitude* $\sqrt{\frac{p_i}{\Pr(S)}}$ 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 $\rho(\Sigma S)=|s\rangle\langle s|$.[3]

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:

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

The probabilities computed for the classical and superposition events will be the same–which is a feature, not a bug, since the same thing occurs in quantum mechanics.[4] 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 conditioning 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.

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

[4] 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 $\rho(U)$ [1, p. 176].

A *partition* $\pi = \{B_1, ..., B_m\}$ 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 $\pi = \{B_1, ..., B_m\}$ and $\sigma = \{C_1, ..., C_{m'}\}$, 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 maximum 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 minimum partition or bottom is the *indiscrete partition* $\mathbf{0}_U = \{U\}$ with only one block U where all the elements of U are blobbed together. 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:

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

The same holds if we cut down to any event $S \subseteq U$, i.e., $\rho(\mathbf{1}_S) = \rho(S)$ and $\rho(\mathbf{0}_S) = \rho(\Sigma S) = |s\rangle\langle s|$. Since $\mathbf{0}_S$ represents the blobbing together of the elements of S and $\mathbf{1}_S$ represents the discrete set S, i.e., the event S in ordinary finite probability theory, this result verifies the mathematical treatment of superposition events ΣS as opposed to discrete events S.

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:

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

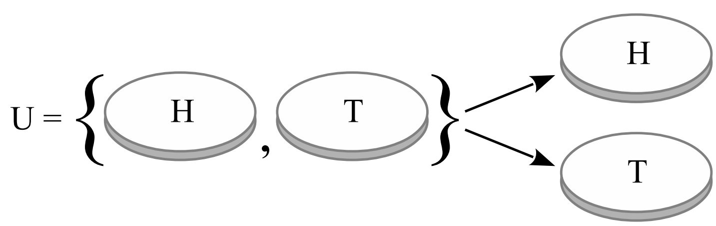

Figure 3: Classical event: A trial picks out heads or tails.

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

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

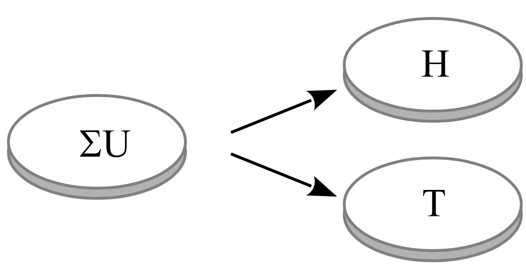

Figure 4: Superposition event: A trial sharpens to heads or tails.

The probability of getting heads in each case is:

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

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

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. As noted, this is a feature, not a bug, since the same thing occurs in quantum mechanics. In QM, they can only be distinguished by measurement in a different observable basis (see [5] for an example).

It might be noted that the density matrix $\rho(\mathbf{1}_U)$ for the discrete partition $\mathbf{1}_U$ is the density matrix $\rho\left(U\right)$ for the classical discrete set U which is like a "statistical mixture describing the state of a classical dice before the outcome of the throw." [1, p. 176] In the logic of partitions (or equivalence relations) [8] and its quantitative version, logical information theory based on logical entropy [6], the *distinctions* or dits of a partition $\pi=(B_1,...,B_m)$ on U are the ordered pairs in $U\times U$ whose elements are in different blocks of the partition. The set of all distinctions is the ditset $\operatorname{dit}\left(\pi\right)=\cup_{j,k=1;j\neq k}^{m}B_j\times B_k$. The complementary set in $U\times U$ is the set of all *indistinctions* or *indits* (ordered pairs in the same block) is $\operatorname{indit}\left(\pi\right)=\cup_{j=1}^{m}B_j\times B_j$ which is the equivalence relation associated with the partition π. The classical nature of the discrete partition $\mathbf{1}_U$ and its density matrix $\rho(U)$ is shown by that partition and only that partition satisfying the:

If $(u,u')\in\operatorname{indit}\left(\mathbf{1}_U\right)$, then $u=u'$
Partition logic Principle of Identity of Indistinguishables.

That is, if $u,u'\in U$ are indistinguishable by the discrete partition, i.e., $(u,u')\in\operatorname{indit}\left(\mathbf{1}_U\right)$, then they are identical. This is trivial mathematically since $\operatorname{indit}\left(\mathbf{1}_U\right)=\Delta=\{(u,u):u\in U\}$ and it comports with the fact that all *other* partitions on U contain at least one block with multiple elements, which may thus be interpreted as non-classical superposition events.

At this point, we may return to the question posed in the Introduction about how does quantum notion of superposition differ from the classical superposition of electromagnetic waves or even water waves. The ontic difference is that in quantum superposition (as opposed to the classical examples of wave superposition) is that the superposed definite- or eigen-states are rendered indefinite on how they differ–which is variously described in the literature as superpositions being blurry, unsharp, smudged, blunt, coherent, fuzzy, blob-like, dispersed, smeared-out, indeterminate, spread-out, or indefinite. In contrast, no such blurriness or indefiniteness occurs in the classical superposition of, say, water or electromagnetic waves. That is why the standard classroom ripple-tank model of the two-slit experiment is seriously misleading since it represents superposition classically as the addition of matter waves.

> Another surprising peculiarity of quantons is that they are blurry or fuzzy rather than neat or sharp. Whereas in classical physics all properties are sharp, in quantum physics only a few are: most are blunt or smudged. ...
> The reason for this fuzziness is that ordinarily an isolated quanton is in a "coherent" state, that is, the combination or superposition (weighted sum) of two or more basic states (or eigenfunctions). The superposition or "entanglement" of states is a hallmark of quantum mechanics. [2, pp. 49-50]

This ontic difference comes out mathematically, not in the addition of the $n\times 1$ state vectors, but in the non-zero off-diagonal elements of the $n\times n$ matrix treatment of density matrices (as well as in the prior incidence matrix treatment of superposition events as opposed to classical discrete events). For instance, in the density matrix $\rho(\pi)$ formulation of a partition π, the non-zero elements $\rho(\pi)_{jk} \neq 0$ correspond to the equivalence relation $(u_j, u_k) \in \text{indit}(\pi)$ indicating the elements of the underlying set U that are cohered, blurred, or equated together in the superpositions represented by the equivalence classes or blocks of π. That aspect of the notion of superposition, which comes out in the matrix treatment, is absent in the classical addition (superposition) of waves. The relationship of the $n \times 1$ amplitude vectors $|s\rangle$ to the quantum-like superposition of the $n \times n$ matrix $\rho(\Sigma S) = |s\rangle\langle s|$ brings us to the Born rule.

5 Conclusion: The Born Rule

The Born Rule does not occur in ordinary classical probability theory because that theory does not include superposition events and the accompanying amplitudes. When superposition events are introduced into the purely mathematical theory, then the probability of outcomes can be computed as the *squares* of the coefficients in a normalized superposition event. But that means that the coefficients could be negative which is necessary if the superposition events (as "states") are to form a vector space. And the requirement that the state space be a vector space is precisely what is required by the Superposition Principle as pointed out by Dirac.

> The superposition process is a kind of additive process and implies that states can in some way be added to give new states. The states must therefore be connected with mathematical quantities of a kind which can be added together to give other quantities-of the same kind. The most obvious of such quantities are vectors. ...
>
> We now assume that each state of a dynamical system at a particular time corresponds to a ket vector, the correspondence being such that if a state results from the superposition of certain other states, its corresponding ket vector is expressible linearly in terms of the corresponding ket vectors of the other states, and conversely. [4, pp. 15-16]

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 ket vector with the i^{th} entry as the amplitude $\langle u_i|s\rangle = \sqrt{\frac{p_i}{\Pr(S)}}\chi_S(u_i) = \sqrt{\frac{\Pr(\{u_i\}\cap S)}{\Pr(S)}}$. Or starting with the *pure* density matrix $\rho(\Sigma S)$, then $|s\rangle$ is obtained (up to sign) as the normalized eigenvector associated with the eigenvalue of 1 and $\rho(\Sigma S) = |s\rangle\langle s|$ is obtained as the spectral decomposition of $\rho(\Sigma S)$ as a Hermitian matrix.

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

$$\Pr(\{u_i\}|\Sigma S) = \operatorname{tr}\left[P_{\{u_i\}}\rho(\Sigma S)\right] = \tfrac{p_i}{\Pr(S)}\chi_S(u_i).$$

The point is that this same probability conditioned by the two-dimensional $n \times n$ density matrix $\rho(\Sigma S)$ could also be obtained from the ket vector $|s\rangle$ of amplitudes as the square of the amplitudes:

$$\langle u_i|s\rangle^2 = \operatorname{tr}\left[P_{\{u_i\}}\rho(\Sigma S)\right] = \tfrac{p_i}{\Pr(S)}\chi_S(u_i).$$

The Born Rule (special case)

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 columns of non-negative numbers. The associated $n \times n$ diagonal density matrix $\rho(S)$ for the classical event S is not the outer product of a one-dimensional vector with itself (except when S is a singleton, i.e., the null case of superposition). It has no non-zero off-diagonal elements indicating the blurring or cohering together of the elements of U. Thus the outcomes in a classical discrete event have probabilities, not amplitudes. To accommodate the notion of a superposition event ΣS, it is necessary to use two-dimensional $n \times n$ density matrices $\rho(\Sigma S)$ where the non-zero off-diagonal amplitudes indicate the blobbing or cohering together in superposition of the elements associated with the corresponding diagonal entries.

And mathematically *those* density matrices $\rho(\Sigma S)$ (unlike $\rho(S)$) can be constructed as the outer products $|s\rangle(|s\rangle)^t = |s\rangle\langle s|$ of ket vectors $|s\rangle$ of amplitudes. Then the probability of the individual outcomes $\{u_i\}$ conditioned by the superposition event ΣS is given as the *square* of amplitudes: $\langle u_i|s\rangle^2 = \mathrm{tr}\left[P_{\{u_i\}}\rho(\Sigma S)\right] = \frac{p_i}{\Pr(S)}\chi_S(u_i)$. Of course, the probability of u_i could also be obtained as $\mathrm{tr}\left[P_{\{u_i\}}\rho(S)\right]$ but that method of encoding probabilities does not generalize to vectors in a vector space, e.g., $\mathbb{R}^n$ or $\mathbb{C}^n$, with [positive, negative, or complex components. Only the encoding of probabilities as squared amplitudes has that feature which is the Born Rule.

Thus the Born Rule arises naturally out of the mathematics of probability theory minimally enriched by superposition events and their associated amplitudes. The Superposition Principle [4, Chap. 1] in QM requires that the states be represented as vectors within a full vector space (not just non-negative vectors). The simple treatment of a superposition event in the minimally enriched probability theory shows that the probabilities can then be expressed as the squares of the normalized vector components. That shows how probabilities can be generated with arbitrary normalized vectors, e.g., with positive or negative components since their square is always positive.

In the Hilbert spaces over the complex numbers $\mathbb{C}$ of quantum mechanics, the components in $|s\rangle$ may be complex so the square $\langle u_i|s\rangle^2$ is then the *absolute* square $|\langle u_i|s\rangle|^2$. But that introduces nothing new in principle over what we have shown here with real matrices arising from the extension of ordinary probability theory with superposition events.

Given the 'mystery' that surrounds QM, it would perhaps be gratifying if the Born Rule was some deep theorem (like the Spin-Statistics Theorem). But the Born Rule does not need any more-exotic or physics-based explanation. Perhaps it is something of a 'disappointment' that the Born Rule is just a somewhat mundane feature of the mathematics of superposition. If an alleged "derivation of the Born Rule" does not start with probability theory, then the probabilities are likely "put in by hand." In the approach taken here, the extension of ordinary probability theory to model superposition events shows how probabilities can be encoded in (normalized) vectors of amplitudes as the (absolute) squares of those amplitudes. And superposition is, not coincidentally, the key *non-classical* feature of quantum mechanics as foreshadowed by the non-zero off-diagonal elements in $\rho(\Sigma S)$ as opposed to $\rho(S)$. In short, the answer to the question: "Where does the Born Rule come from?" is superposition.

6 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.

References

[1] Auletta, Gennaro, Mauro Fortunato, and Giorgio Parisi. 2009. *Quantum Mechanics.* Cambridge UK: Cambridge University Press.

[2] Bunge, Mario. 2010. *Matter and Mind: A Philosophical Inquiry. Boston Studies in the Philosophy of Science #287.* Dordrecht: Spring Publications.

[3] Cushing, James T. 1988. "Foundational Problems in and Methodological Lessons from Quantum Field Theory." In *Philosophical Foundations of Quantum Field Theory*, edited by Harvey R. Brown and Rom Harre, 25–39. Oxford: Clarendon Press.

[4] Dirac, P.A.M. 1958. *The Principles of Quantum Mechanics (4th Ed.).* Oxford: Clarendon Press.

[5] 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.

[6] Ellerman, David. 2021. *New Foundations for Information Theory: Logical Entropy and Shannon Entropy.* Cham, Switzerland: SpringerNature. https://doi.org/10.1007/978-3-030-86552-8.

[7] 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.

[8] Ellerman, David. 2023. *The Logic of Partitions: With Two Major Applications. Studies in Logic 101.* London: College Publications.

[9] Ellerman, David. 2024. "A New Logic, a New Information Measure, and a New Information-Based Approach to Interpreting Quantum Mechanics." *Entropy Special Issue: Information-Theoretic Concepts in Physics* 26 (2). https://doi.org/10.3390/e26020169.

[10] Hossenfelder, Sabine. 2021. "A Derivation of Born's Rule from Symmetry." *Annals of Physics* 425 (February):168394. https://doi.org/10.1016/j.aop.2020.168394.

[11] 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.

[12] Stachel, John. 1986. "Do Quanta Need a New Logic?" In *From Quarks to Quasars: Philosophical Problems of Modern Physics*, edited by Robert G. Colodny, 229–347. Pittsburgh: University of Pittsburgh Press.

[13] Stachel, John. 1997. "Feynman Paths and Quantum Entanglement: Is There Any More to the Mystery?" In *Potentiality, Entanglement and Passion-at-a-Distance/Quantum-Mechanical Studies for Abner Shimony, Vol. 2*, edited by Robert S. Cohen, Michael Horne, and John Stachel, 245–56. Dordrecht/Boston: Kluwer Academic.

[14] 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.

[15] Weinberg, Steven. 2015. *Lectures on Quantum Mechanics 2nd Ed.* Cambridge UK: Cambridge University Press.