International Journal of Quantum Foundations **11** (2025) 57-72

Original Paper

Born Again! The Born Rule as a Feature of Superposition

David Ellerman

Independent researcher, Ljubljana, Slovenia.
E-mail: david@ellerman.org

Received: 28 December 2024 / Accepted: 14 March 2025 / Published: 21 March 2025

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 introducing "superposition events" in addition to the usual discrete events. Two-dimensional matrices (e.g., incidence matrices and density matrices) are needed to mathematically represent the differences between the two types of events. Then it is shown that those incidence and density matrices for superposition events are the (outer) products of a vector and its transpose whose components foreshadow the "amplitudes" of quantum mechanics. The squares of the components of those "amplitude" vectors yield the probabilities of the outcomes. That is how probability amplitudes and the Born Rule arise in the minimal extension of probability theory to include superposition events. This naturally extends to the full Born Rule in the Hilbert spaces over the complex numbers of quantum mechanics. It would perhaps be satisfying if probability amplitudes and the Born Rule only arose as the result of deep results in quantum mechanics (e.g., Gleason's Theorem). But both arise in a simple extension of probability theory to include "superposition events"–which should not be too surprising since superposition is the key non-classical concept in quantum mechanics.

Keywords: Born Rule, superposition, amplitudes, 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. Hence one might ask with Steven Weinberg: "So where does the Born Rule come from?" [21, 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 [16], [20], [15], and the articles cited therein.

In this paper, a different approach is taken. We ask 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. Hence there two versions of events, the usual discrete events as a subset of the outcome or sample space, and associated with the same subset is a superposition event where the outcomes in the event are intuitively blurred, blobbed, or cohered together. The point probabilities for the outcomes are the same in the discrete version and the superposition version of the event. The difference lies in how the outcomes are not distinguished from each other in the superposition event since they are blobbed or superposed together.

If we represent each type of event with a $n \times 1$ column vector of probabilities, then the two vectors are the same. Therefore to mathematically represent the difference between these two types of events, we need to add another dimension to form a $n \times n$ matrix where the difference lies in the off-diagonal terms. If the i^{th} and k^{th} outcome are blobbed together, this is indicated with the i, k off-diagonal element being non-zero. The discrete events are then represented with diagonal $n \times n$ matrices of no non-zero off-diagonal elements. From that simple difference, probability amplitudes and the Born Rule arise.

How does this extension of probability theory to include superposition events with matrix representations lead to the Born rule? We will construct the matrices that exhibit the difference between discrete and superposition events. We will see that if the matrix represents a single superposition event, then it can be constructed as the "outer product" of a *new column vector* times its transpose (a row vector). No such construction is possible for the diagonal matrices representing discrete events (except in the overlapping case of a singleton event). Thus these new vectors contain something new that is not there in the ordinary probability theory for discrete events. The entries in these new vectors are "probability amplitudes" whose squares give the probabilities of the outcomes in the superposition event. And that is the Born Rule. This simple case then extends in a straightforward manner to quantum mechanics where the $n \times n$ matrices are density matrices over the field of complex numbers (so the transpose is then the conjugate transpose and the squares are the absolute squares). That is the "battle plan" to show where the Born Rule comes from.

Why superposition events? 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." [19, 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 along with the Born Rule 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 mathematics of superposition events foreshadowed in this minimally expanded probability theory using only the real numbers–and then extended to the Hilbert spaces over $\mathbb{C}$ in QM. The fact that the Born Rule *works* to give empirical probabilities is an empirically corroborated fact. No one can "derive" that empirical fact. Our task is only to develop the mathematical fact that it arises when probability theory is extended to include a notion of superposition events in addition to the usual discrete events.

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,

---

[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.

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 that S occurs (including the case of a specific outcome $T = \{u_i\}$).

The intuitive idea of the corresponding superposition event, denoted ΣS, is that the outcomes in the event are *not* distinguished from each other but are blobbed or cohered together as an indefinite event. We will see how the concept of superposition yields 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. [18, p. 314]

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. The probabilities of the individual outcomes are assumed to be the same when conditioned by the discrete event or the superposition event: $\Pr(u_i|S) = \Pr(u_i|\Sigma S) = p_i$ where $p = (p_1, ..., p_n)$ are the point probabilities.[2] In the case of a singleton event $S = \{u_i\}$, the ordinary event $S = \{u_i\}$ is the same as the superposition event $\Sigma S = \Sigma\{u_i\} = \{u_i\} = S$.

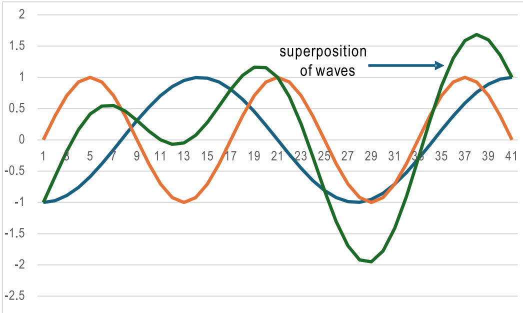
There are two fundamentally different ways to interpret superposition: the classical wave-addition version and the objective indefiniteness version. In the classical wave-addition version, the superposition of two definite waves is another equally definite wave as illustrated in Figure 1. There is no hint of indefiniteness.

The appropriate quantum notion of superposition-as-indefiniteness differs from the classical superposition of electromagnetic waves or even water waves. The ontic difference is that, in quantum superposition, the superposed definite- or eigen-states are rendered

---

[2] This is not a bug but a feature since in QM, the probabilities of the eigenstates in a superposition are the same as the probabilities in the corresponding completely decomposed mixed state.

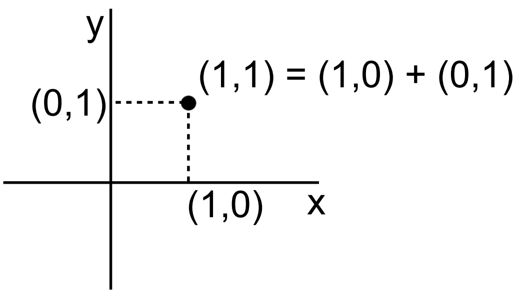
Figure 1. Classical wave-addition notion of superposition



indefinite on how they differ–which is variously described in the literature as superpositions being blurry, unsharp, smudged, blunt, cohered, fuzzy, blob-like, dispersed, smeared-out, indeterminate, spread-out, or indefinite. In contrast, the superposition of two classical waves is just as definite as the summands.

For a visual illustration of the simple indefiniteness or blurred version of superposition, consider the real x, y plane with the two definite basis vectors or eigen (= definite) states, the definite state of "x-ness" $(1, 0)$ and the definite state of "y-ness" $(0, 1)$. Then their superposition $(1, 0) + (0, 1) = (1, 1)$ should be thought of as the state (equally in this case) indefinite between the two definite basis (or eigen) vectors as in Figure 2.

Figure 2. The vector $(1, 1)$ as being indefinite between definite basis vectors $(1, 0)$ and $(0, 1)$

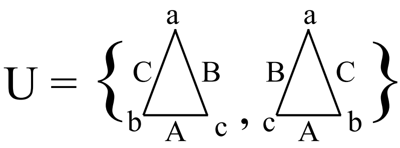


In contrast, no such blurriness or indefiniteness occurs in the classical superposition of, say, water or electromagnetic waves. That is why the standard classroom ripple-tank model of the two-slit experiment is seriously misleading since it represents superposition classically as the addition of matter waves.

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 as in Figure 3.

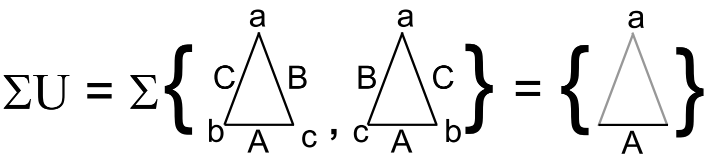
The superposition event ΣU is definite on the properties that are common to the elements of U, i.e., the angle a and the opposing side A, but is indefinite where the two triangles are distinct, i.e., the two equal sides and their opposing angles are not distinguished by labels

Figure 3. Set of distinct isosceles triangles



as illustrated in Figure 4.

Figure 4. The superposition event ΣU



3. Mathematically modeling superposition events (over the reals)

What is a mathematical model that will distinguish between the ordinary discrete event $S \subseteq U$ 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 $\chi_S : U = \{u_1, ..., u_n\} \to \{0, 1\}$ 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 event ΣS by the off-diagonal elements.

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 rendering-indefinite, equating, cohering, or blobbing together of the corresponding diagonal elements in a single superposition event. If the blobbing together is thought of as a type of equating, then the superposition is like an equivalence class in an equivalence relation. 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.

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 $\langle t|s\rangle = (|t\rangle)^t|s\rangle = \sum_{i=1}^{n} t_i s_i$.

To explain "where the Born Rule comes from," we have to also show how probability *amplitudes* arise in our extension of classical probability theory. This is foreshadowed even at the level of incidence matrices. The *outer product* $|s\rangle\langle t|$ is the $n \times n$ matrix denoted as $|s\rangle\langle t| = |s\rangle(|t\rangle)^t$. The key result, that foreshadows probability amplitudes, is that the outer product of $|S\rangle$ with its transpose is the incidence matrix representing the superposition event–and vice-versa.

Theorem 1. *Given* $|S\rangle$*, the outer product* $|S\rangle\langle S|$ *is an incidence matrix of the form* $\mathrm{In}(S \times S)$*, and given an incidence matrix with the form* $\mathrm{In}(S \times S)$*, it is obtained as the outer product* $|S\rangle\langle S|$.

Proof: $(|S\rangle\langle S|)_{ik} = \chi_S(u_i)\chi_S(u_k) = 1$ if and only if (iff) $(u_i, u_k) \in S \times S$.□

This is a key result because it directly connects the *outer product of a vector and its transpose with the notion of superposition* (even though the vectors and matrices have only $0, 1$ entries). The vector in the outer product foreshadows the vector of "amplitudes" that is used in the Born Rule and thereby the connection with superposition. Moreover, it might be noted that $|S\rangle$ is an eigenvector of $\mathrm{In}\,(S \times S)$ with the eigenvalue $|S| = \mathrm{tr}\,[\mathrm{In}\,(S \times S)]$ since (where the *trace* $\mathrm{tr}\,[M]$ of a square matrix M is the sum of its diagonal elements):

$$\mathrm{In}\,(S \times S)\,|S\rangle = |S\rangle\langle S||S\rangle = |S|\ |S\rangle = \mathrm{tr}\,[\mathrm{In}\,(S \times S)]\,|S\rangle.$$

These results for incidence matrices foreshadow the corresponding results for pure state density matrices in the mathematics of QM to be considered below.

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:

$$\rho(S) = \frac{\mathrm{In}(\Delta S)}{|S|} \text{ and } \rho(\Sigma S) = \frac{\mathrm{In}(S \times S)}{|S|} = \frac{1}{\sqrt{|S|}}|S\rangle\langle S|\frac{1}{\sqrt{|S|}}$$

over the reals $\mathbb{R}$. In the case of equiprobable outcomes $p_i = \frac{1}{n}$, we already have a special case of the Born Rule for the probability of u_i given the superposition event ΣS:

$$\langle u_i|\frac{1}{\sqrt{|S|}}S\rangle^2 = \frac{1}{|S|}\chi_S(u_i).$$

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, i.e., superposition event $\approx$ pure state in QM.

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 labeled 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 event 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 elements that indicate that the differences between the diamond suite $\diamondsuit$ and the hearts suite $\heartsuit$ are rendered indefinite, blurred, cohered, or superposed in the superposition state $\Sigma\{\diamondsuit, \heartsuit\}$.

The next step is to bring in general point probabilities $p = (p_1, ..., p_n)$ where those two real density matrices $\rho(S)$ and $\rho(\Sigma S)$ defined above 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 point 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 $\rho(U)$ for the discrete event 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: $\Pr(S) = \operatorname{tr}[P_S \rho(U)]$. 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:

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

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 normalized column vector $\frac{1}{\sqrt{|S|}}|S\rangle$ is generalized to the normalized $|s\rangle$ 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. Then $\rho(\Sigma S) = \frac{1}{\sqrt{|S|}}|S\rangle\langle S|\frac{1}{\sqrt{|S|}}$ generalizes to $\rho(\Sigma S) = |s\rangle\langle s|$ which is a pure state since:

$$\rho(\Sigma S)^2 = |s\rangle\langle s|s\rangle\langle s| = \operatorname{tr}[|s\rangle\langle s|]|s\rangle\langle s| = |s\rangle\langle s| = \rho(\Sigma S).$$

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 (up to sign) as the normalized eigenvector associated with the eigenvalue of 1 and $\rho(\Sigma S) = |s\rangle\langle s|$.[3] The "amplitude" vector $|s\rangle$ arises from the representation of $\rho(\Sigma S)$ as an outer product (foreshadowed even for incidence matrices) and the Born Rule mathematically arises in this extended probability theory as:

$$\langle u_i|s\rangle^2 = \frac{p_i}{\Pr(S)}\chi_S(u_i) = \Pr(u_i|S).$$

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

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

[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 measurement in a different basis is necessary to distinguish them.

$T \cap S$. For superposition events, the given superposition event ΣS is sharpened (i.e., made less indefinite) to the superposition event $\Sigma(T \cap S)$.

5. The math of indefiniteness/definiteness: Partitions

Since superposition is interpreted in terms of an indefinite combination of the definite (eigen-) states (not as the addition of definite waves to give a definite wave), we can reformulate the mathematics in terms of partitions. 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. Partitions (or equivalence relations) are the natural mathematics to represent indefiniteness (*within* a block or equivalence class) and definiteness (*between* blocks or equivalence classes). Taking each block B_j as a superposition event, then there is the normalized column vector $|b_j\rangle$ whose i^{th} entry is $\langle u_i|b_j\rangle = \sqrt{\frac{p_i}{\Pr(B_j)}}\chi_{B_j}(u_i)$ and the pure state 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 $\Pr(B_j) = \sum_{u_i \in B_j} p_i$ with the remaining $n - m$ values of 0. Then $\rho(\pi)_{ik} = \sqrt{p_i p_k}$ if (u_i, u_k) is an *indistinction* of π, i.e., $(u_i, u_k) \in B_j \times B_j$ for some $B_j \in \pi$, and otherwise zero. Thus $\rho(\pi)_{ik}$ is an *indistinction (coherence) amplitude* in the sense that its square is the probability in two random draws from U of getting the ordered pair (u_i, u_k) as an indistinction of π.

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. 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. The pure state density matrix $\rho(\Sigma U)$ has an eigenvalue of 1 and the associated eigenvector is $|u\rangle$ so that $\rho(\Sigma U) = |u\rangle\langle u|$.

Theorem 2. $\rho(\mathbf{1}_U) = \rho(U)$ *and* $\rho(\mathbf{0}_U) = \rho(\Sigma U) = |u\rangle\langle u|$.

The same holds if we cut down to any event $B_j \subseteq U$, which comes in the two forms, i.e., $\rho(\mathbf{1}_{B_j}) = \rho(B_j)$ and $\rho(\mathbf{0}_{B_j}) = \rho(\Sigma B_j) = |b_j\rangle\langle b_j|$. 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 *using partitions* verifies the previous mathematical treatment of superposition events ΣS as opposed to discrete events S.

Partitions show the difference. ΣS is the equivalence class that is the single block in the indiscrete partition $\mathbf{0}_S$ and S is the set of discrete elements in the discrete partition $\mathbf{1}_S$ of singletons of the elements of S.

This treatment of superposition events can be further confirmed. Let $supp(\psi) \subseteq U$ be the *support set* of $|\psi\rangle$ (the set of $u_i \in U$ with non-zero coefficients $\langle u_i|\psi\rangle$ in $|\psi\rangle = \sum_{i=1}^{n} \langle u_i|\psi\rangle |u_i\rangle$) and let $supp(\rho)$ be the $0, 1$ matrix giving the pattern of zero and non-zero entries in ρ, i.e., $supp(\rho)_{ik} = 1$ if $\rho_{ik} \neq 0$, else 0. Then we have the following result about the pattern of zero and non-zero entries in any pure state density matrix in QM. The result connects superposition events and pure state density matrices in QM. It confirms our original treatment of superposition events ΣS being represented by $\mathrm{In}(S \times S)$.[5] Let $|\psi\rangle$ be a normalized state vector and $\rho = |\psi\rangle\langle\psi|$ the corresponding pure state density matrix in QM.

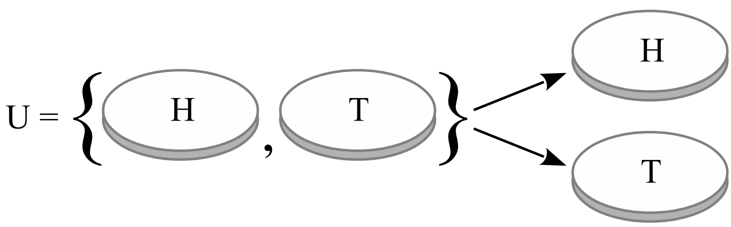
Theorem 3. $supp(\rho) = \mathrm{In}(supp(\psi) \times supp(\psi))$.

Proof: Since the complex numbers are a field (as opposed to only a ring), a product of two entries $\langle u_i|\psi\rangle$ and $\langle\psi|u_k\rangle$ in $\rho = |\psi\rangle\langle\psi|$ is non-zero if and only if the two entries are non-zero if and only if $u_i, u_k \in supp(\psi)$, i.e., $(u_i, u_k) \in supp(\psi) \times supp(\psi)$.□

The situation can be illustrated by considering the case of flipping a fair coin. The classical set of outcomes $U = \{H, T\}$ is represented by the density matrix and Figure 5:

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

Figure 5. 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 and Figure 6:

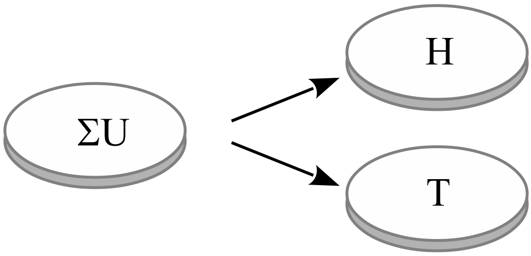
$$\rho(\Sigma U) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The probability of getting heads in each case is:

---

[5] This result also illustrates the connection between full QM and the pedagogical model of "quantum mechanics over support sets" or QM/Sets where superposition sets in QM/Sets are the set version of pure states in QM ([5]; [11]).

Figure 6. Superposition event: A trial sharpens to heads or tails



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

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

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 [6] for an example).

The density matrix $\rho(\mathbf{1}_U)$ for the discrete partition $\mathbf{1}_U$ is the density matrix $\rho(U)$ 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) [9] and its quantitative version, logical information theory based on logical entropy [7], 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 $\mathrm{dit}(\pi) = \cup_{j,k=1;j\neq k}^{m} B_j \times B_k$. The complementary set in $U \times U$ is the set of all *indistinctions* or *indits* (ordered pairs of elements in the same block) is $\mathrm{indit}(\pi) = \cup_{j=1}^{m} B_j \times B_j$ which is the equivalence relation associated with the partition π. In the lattice of partitions on U, the top discrete partition represents classicality while all the partitions below it represent states in the quantum "underworld" since they all involve a non-singleton block, i.e., a superposition set.[12] The classical (non-quantum) 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:

Partition version of Leibniz's Principle of Identity of Indistinguishables

$$\boxed{\text{If } (u, u') \in \mathrm{indit}(\mathbf{1}_U), \text{ then } u = u'.}\,.$$

That is, if $u, u' \in U$ are indistinguishable by the discrete partition, i.e., $(u, u') \in \mathrm{indit}(\mathbf{1}_U)$, then they are identical. This is trivial mathematically since $\mathrm{indit}(\mathbf{1}_U) = \Delta = \{(u_i, u_i) : u_i \in U\}$.

6. 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 (that come from representing the density matrix of a superposition event as an outer product). When superposition events are introduced into the purely mathematical theory (over the reals), then the probability of outcomes can be computed as the *squares* of the coefficients in the normalized amplitude vector $|s\rangle$ associated with the superposition event ΣS.

We have seen that pure state density matrices $\rho(\Sigma S)$ 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 state* density matrix $\rho(\Sigma S) = \rho(\Sigma S)^2$, 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:

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

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:

$$\langle u_i|s\rangle^2 = \Pr(u_i|\Sigma S).$$

The Born Rule (special case of real density matrices)

In classical finite probability theory, the events S are all discrete sets that can be represented by n-ary columns of non-negative numbers, i.e., a $0, 1$-vector to represent the elements of S or the vector of the probabilities $\frac{p_i}{\Pr(S)}\chi_S(u_i)$. The associated $n \times n$ *diagonal* density matrix $\rho(S)$ for the classical discrete 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 S. 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: $\langle u_i|s\rangle^2 = \frac{p_i}{\Pr(S)}\chi_S(u_i)$.

Thus the Born Rule arises naturally out of the mathematics of probability theory minimally enriched over the reals by superposition events and their associated amplitudes. Developing the probability theory enriched by superposition events over the complex numbers yields the mathematics of probabilities in quantum mechanics. In the Hilbert spaces over the complex numbers $\mathbb{C}$ of quantum mechanics, the components in $|s\rangle$ may be complex (and the bra $\langle s|$ is the conjugate transpose of the corresponding ket) 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.

The fact that the Born Rule *works* in QM is an empirical fact that cannot be derived from any probability theory (enriched with superposition events or not and over the reals or complex numbers).

Does this treatment of the Born Rule have any implications for the "demolition derby" of interpretations of QM? It is not the Born Rule itself but the interpretation of superposition in terms of indefiniteness (rather than classical wave addition), that points toward what Abner Shimony called the "Literal Interpretation."

> Heisenberg [14, p. 53] used the term "potentiality" to characterize a property which is objectively indefinite, whose value when actualized is a matter of objective chance, and which is assigned a definite probability by an algorithm presupposing a definite mathematical structure of states and properties. ...These statements, together with the theses about potentiality, may collectively be called "the Literal Interpretation" of quantum mechanics. This is the interpretation resulting from taking the formalism of quantum mechanics literally, as giving a representation of physical properties themselves, rather than of human knowledge of them, and by taking this representation to be complete. [17, pp. 6-7]

This is the interpretation held implicitly by a non-philosophical quantum physicist who considers a superposition state as having objectively indefinite values of an observable prior to a state reduction. [12]

How does this treatment of the Born Rule compare to the other "derivations" or analyses in the literature, e.g., [16], [20], [15], and the articles cited therein–including Gleason's Theorem [13]? The answer is simply that those treatments work within the full Hilbert space framework of QM whereas our analysis has shown that the Born Rule and probability amplitudes arise at the much more elementary mathematical level of adding superposition events to the usual discrete events in probability theory.

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 emerges as just a somewhat mundane *feature* of the *mathematics* of superposition-enriched probability theory. But superposition is, not coincidentally, the key *non-classical* feature of quantum mechanics so, in short, the answer to Weinberg's question: "So where does the Born Rule come from?" is "superposition."

7. 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 acknowledgments and 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. 2017. "Quantum Mechanics over Sets: A Pedagogical Model with Non-Commutative finite Probability Theory as Its Quantum Probability Calculus." *Synthese* 194 (12): 4863–96.
6. 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.
7. 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.
8. 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.
9. Ellerman, David. 2023. *The Logic of Partitions: With Two Major Applications. Studies in Logic 101*. London: College Publications.
10. 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.

11. Ellerman, David. 2024. "A New Approach to Understanding Quantum Mechanics: Illustrated Using a Pedagogical Model over $\mathbb{Z}_2$." *AppliedMath* 4 (2): 468–94. https://doi.org/10.3390/appliedmath4020025.
12. Ellerman, David. 2024. *Partitions, Objective Indefiniteness, and Quantum Reality: The Objective Indefiniteness Interpretation of Quantum Mechanics*. Cham, Switzerland: Springer Nature. https://doi.org/10.1007/978-3-031-61786-7_2.
13. Gleason, A. M. 1957. "Measures on the Closed Subspaces of a Hilbert Space." *Journal of Mathematics and Mechanics* 6 (6): 885–93.
14. Heisenberg, Werner. 1962. *Physics & Philosophy: The Revolution in Modern Science.* New York: Harper Torchbooks.
15. 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.
16. 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.
17. Shimony, Abner. 1999. "Philosophical and Experimental Perspectives on Quantum Physics." In *Philosophical and Experimental Perspectives on Quantum Physics: Vienna Circle Institute Yearbook 7*, 1–18. Dordrecht: Springer Science+Business Media.
18. 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.
19. 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.
20. 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.
21. Weinberg, Steven. 2015. *Lectures on Quantum Mechanics 2nd Ed.* Cambridge UK: Cambridge University Press.