

Original Paper

On Two Recent Approaches to the Born Rule

Blake C. Stacey

QBism Group, Physics Department, University of Massachusetts Boston
100 Morrissey Boulevard, Boston MA 02125, USA

E-mail: blake.stacey@umb.edu

Received: 28 March 2021 / Accepted: 31 March 2021 / Published: 12 April 2021

Abstract: I comment briefly on derivations of the Born rule presented by Masanes et al. and by Hossenfelder.

Keywords: Born rule; Gleason's theorem; noncontextuality

According to the textbooks, the Born rule is the means in quantum theory by which probabilities are calculated, given a quantum state and a representation of a measurement. Apart from certain heterodox individuals [1], the common sentiment has been that positing the Born rule as a premise of the theory is distasteful, and various attempts have been made over the years to derive it from assumptions that are deemed more culturally suitable. I will remark on two such proposals here.

We will be making some historical comparisons to Gleason's theorem of 1957 [2]. Since I've been surprised more than once by physicists not knowing about this theorem, a short recapitulation is in order. Gleason assumes that to each physical system is associated a Hilbert space, and that each possible measurement upon that system corresponds to an orthonormal basis of that space. The concept of probability enters by way of a *frame function*, a map from unit vectors to the unit interval with the property that the values for vectors comprising an orthonormal basis always add up to 1. Importantly, we assume that the probability assigned to a measurement outcome depends on the vector representing that outcome, but not on any choice of basis in which that vector might be embedded. That is, the definition of frame functions makes probability assignments "noncontextual", in the lingo. Gleason's theorem proves that if the dimension of the Hilbert space \mathcal{H} is greater than 2, any frame function $f : \mathcal{H} \rightarrow [0, 1]$ must take the form

$$f(|\phi\rangle) = \langle \phi | \rho | \phi \rangle, \quad (1)$$

where ρ is a positive semidefinite operator of trace 1, or in other words, a density matrix. So, Gleason's theorem gives the set of valid states *and* the rule for calculating probabilities given a state.

It is significantly easier to prove the POVM version of Gleason's theorem, in which a "measurement" is not necessarily an orthonormal basis, but rather any resolution of the identity into positive semidefinite operators, $\sum_i E_i = I$. In this case, the result is that any valid assignment of probabilities to measurement outcomes, or "effects", takes the form $p(E) = \text{tr}(\rho E)$ for some density operator ρ . The math is easier; the conceptual upshot is the same [3,4].

1. The State Space of Quantum Mechanics is Redundant

The first paper I'd like to discuss is "The measurement postulates of quantum mechanics are operationally redundant" by Masanes, Galley and Müller [5]. The short version of my spiel is that they present a condition on *states* that seems more naturally to me like a condition on *measurement outcomes*. Upon making this substitution, the Masanes, Galley and Müller (MGM) result comes much closer to resembling Gleason's theorem than they say it does.

I have a sneaky suspicion that a good many other attempted "derivations of the Born rule" really amount to little more than burying Gleason's assumptions under a heap of dubious justifications. MGM do something more interesting than that. They start with what they consider the "standard postulates" of quantum mechanics, which in their reckoning are five in number. Then they discard the last two and replace them with rules of a more qualitative character. Their central result is that the discarded postulates can be re-derived from those that were kept, plus the more qualitative-sounding conditions.

MGM say that the assumptions they keep are about state space, while the ones they discard are about measurements. But the equations in the three postulates that they keep could just as well be read as assumptions about measurements instead. Since they take measurement to be an operationally primitive notion — fine by me, anathema to many physicists! — this is arguably the better way to go. Then they add a postulate that has the character of noncontextuality: The probability of an event is independent of how that event is embedded into a measurement on a larger system. So, they work in the same setting as Gleason (Hilbert space), invoke postulates of the same nature, and arrive in the same place. The conclusion, if you take their postulates about complex vectors as referring to measurement outcomes, is that "preparations" are dual to outcomes, and outcomes occur with probabilities given by the Born rule, thereupon turning into new preparations.

Let's treat this in a little more detail.

Here is the first postulate of what MGM take to be standard quantum mechanics:

To every physical system there corresponds a complex and separable Hilbert

space \mathbb{C}^d , and the pure states of the system are the rays $\psi \in \mathbb{P}\mathbb{C}^d$.

We strike the words “pure states” and replace them with “sharp effects” — an equally undefined term at this point, which can only gain meaning in combination with other ideas later.

(I spend at least a little of every working day wondering why quantum mechanics makes use of complex numbers, so this already feels intensely arbitrary to me, but for now we’ll take it as read and press on.)

MGM define an “outcome probability function” as a mapping from rays in the Hilbert space \mathbb{C}^d to the unit interval $[0, 1]$. The abbreviation OPF is fine, but let’s read it instead as *operational preparation function*. The definition is the same: An OPF is a function $\mathbf{f} : \mathbb{P}\mathbb{C}^d \rightarrow [0, 1]$. Now, though, it stands for the probability of obtaining the measurement outcome ψ , for each ψ in the space $\mathbb{P}\mathbb{C}^d$ of sharp effects, given the preparation \mathbf{f} . All the properties of OPFs that they invoke can be justified equally well in this reading. If $\mathbf{f}(\psi) = 1$, then event ψ has probability 1 of occurring given the preparation \mathbf{f} . For any two preparations \mathbf{f}_1 and \mathbf{f}_2 , we can imagine performing \mathbf{f}_1 with probability p and \mathbf{f}_2 with probability $1 - p$, so the convex combination $p\mathbf{f}_1 + (1 - p)\mathbf{f}_2$ must be a valid preparation. And, given two systems, we can imagine that the preparation of one is \mathbf{f} while the preparation of the other is \mathbf{g} , so the preparation of the joint system is some composition $\mathbf{f} \star \mathbf{g}$. And if measurement outcomes for separate systems compose according to the tensor product, and this \star product denotes a joint preparation that introduces no correlations, then we can say that $(\mathbf{f} \star \mathbf{g})(\psi \otimes \phi) = \mathbf{f}(\psi)\mathbf{g}(\phi)$. Furthermore, we can argue that the \star product must be associative, $\mathbf{f} \star (\mathbf{g} \star \mathbf{h}) = (\mathbf{f} \star \mathbf{g}) \star \mathbf{h}$, and everything else that the composition of OPFs needs to satisfy in order to make the algebra go.

Ultimately, the same math has to work out, after we swap the words around, because the final structure is self-dual: The same set of rays $\mathbb{P}\mathbb{C}^d$ provides the extremal elements both of the state space and of the set of effects. So, if we take the dual of the starting point, we have to arrive in the same place by the end.

But is either choice of starting point more *natural*?

Beginning with the set of measurement outcomes may help put the mathematics in conceptual and historical context. For instance, when proving a no-hidden-variables theorem of the Kochen–Specker type, the action lies in the choice of measurements, and how the rays that represent one measurement can interlock with those for another [6]. So, from that perspective, putting the emphasis on the measurements and then deriving the state space is the more conceptually clean move.

That said, on a deeper level, I don’t find either choice all that compelling. To appreciate why, we need only look again at that arcane symbol, $\mathbb{P}\mathbb{C}^d$. That is the setting for the whole argument, and it is completely opaque. Why the complex numbers? Why throw away an overall phase? What is the meaning of “dimension”, and why does it scale multiplicatively

when we compose systems? (A typical justification for this last point would be that if we have n completely distinct options for the state of one system, and we have m completely distinct options for the state of a second system, then we can pick one from each set for a total of nm possibilities. But what are these options “completely distinct” with respect to, if we have not yet introduced the concept of measurement? Why should dimension be the quantity that scales in such a nice way, if we have no reason to care about vectors being orthogonal?) All of this cries out for a more fundamental understanding [7].

2. Nothing Isn’t What It Used To Be

The second paper I’d like to discuss is by Hossenfelder, originally titled “Born’s rule from almost nothing” [8]. The claim of this paper is nicely stated in a concise form up front. Let $|\Psi\rangle$ and $|\Phi\rangle$ denote unit-norm elements of a complex vector space \mathbb{C}^N .

Claim: The only well-defined and consistent distribution for transition probabilities $P_N(|\Psi\rangle \rightarrow |\Phi\rangle)$ on the complex sphere which is continuous, independent of N , and invariant under unitary operations is $P_N(|\Psi\rangle \rightarrow |\Phi\rangle) = |\langle\Psi|\Phi\rangle|^2$.

Here, “well-defined” means that P_N always evaluates to a number in the unit interval, and “consistent” means that if $\{|\Phi_i\rangle\}$ is an orthonormal basis, then

$$\sum_{i=1}^N P_N(|\Psi\rangle \rightarrow |\Phi_i\rangle) = 1, \quad (2)$$

and also

$$P_N(|\Phi_i\rangle \rightarrow |\Phi_j\rangle) = \delta_{ij}. \quad (3)$$

Importantly, we have again an assumption of context-independence, since the transition probability P_N is posited to be indifferent to the set in which the final state might be embedded. Whereas in Gleason’s theorem the structure of the possible measurement outcomes was assumed and the state space derived, here both sets are taken as given (and indeed identical).

If the initial state is $|\Psi\rangle$ and the possible post-transition states are $\{|\Phi_i\rangle\}$, then unitary transformations will leave invariant the 3-vertex Bargmann invariants $\langle\Psi|\Phi_i\rangle\langle\Phi_i|\Phi_j\rangle\langle\Phi_j|\Psi\rangle$. These will vanish if the post-transition states are an orthonormal basis, but if we do not assume Gleason-style context independence, then the Bargmann invariants can be part of the context and the transition probabilities can depend upon them. (These invariants can be quite rich mathematically [9], so it is almost a shame that they don’t appear more directly in calculating probabilities!) In that case, it seems to me, we would have no particular reason to say that orthogonality should mean zero probability. So,

we would have no particular reason to demand that the post-transition states must form an orthonormal basis.

Even if we do not wish to incorporate the lovely 3-vertex information, allowing context dependence opens up the possibilities for non-Born-rule probability assignments. Let g be any continuous, nonnegative function with the property that $g(0) = 0$. Then the generalized Preskill rule [10] given by

$$P_N(|\Psi\rangle \rightarrow |\Phi_i\rangle | \{|\Phi_j\rangle\}) = \frac{g(|\langle\Psi|\Phi_i\rangle|)}{\sum_j g(|\langle\Psi|\Phi_j\rangle|)} \quad (4)$$

satisfies the desiderata that $|\Psi\rangle$ transitions to itself with probability 1 and to an orthogonal $|\Phi_i\rangle$ with probability 0.

Later in the paper, the assumption of continuity is dropped. This raises further possibilities for contextual, non-Born-rule probability assignments, even piecewise-continuous ones. For example, suppose that a measurement corresponds to an orthonormal basis $\{|\Phi_i\rangle\}$, and let a measurement induce a transition from the initial state $|\Psi\rangle$ to whichever $|\Phi_i\rangle$ has the largest absolute overlap $|\langle\Psi|\Phi_i\rangle|$. If there is a tie for largest, we distribute the probability evenly across those outcomes. This rule is unitarily invariant and independent of the dimension N . Moreover, if one of the $\{|\Phi_i\rangle\}$ is equal to $|\Psi\rangle$, then the state will remain unchanged with probability 1, and because for any basis at least one absolute overlap must be greater than zero, a transition will never occur to an orthogonal state.

3. Conclusion

Examining these recent approaches to the Born rule has underlined the significance of Gleason's noncontextuality assumption. Clearly, it is a mathematically potent condition. From one perspective, it is physically rather presumptuous: If one takes the traditional view that going from classical to quantum physics means promoting variables to Hermitian operators, then why should expectation values not depend upon *operators*, rather than merely upon eigenvectors taken one at a time? What conceptual desiderata make a Gleason-style assumption a *natural* move? Finding the answer, I suspect, will require beginning before Hilbert spaces and deriving them in turn.

References

1. C. A. Fuchs and R. Schack, "[Quantum-Bayesian Coherence](#)," *Reviews of Modern Physics* **85** (2013), 1693, [arXiv:1301.3274](#).
2. A. M. Gleason, "[Measures on the closed subspaces of a Hilbert space](#)," *Indiana University Mathematics Journal* **6**, 4 (1957), 885–93.
3. P. Busch, "[Quantum States and Generalized Observables: A Simple Proof of](#)

- Gleason's Theorem," *Physical Review Letters* **91** (2003), 120403, [arXiv: quant-ph/9909073](https://arxiv.org/abs/quant-ph/9909073).
4. C. M. Caves, C. A. Fuchs, K. K. Manne and J. M. Renes, "Gleason-Type Derivations of the Quantum Probability Rule for Generalized Measurements," *Foundations of Physics* **34** (2004), 193–209, [arXiv:quant-ph/0306179](https://arxiv.org/abs/quant-ph/0306179).
 5. L. Masanes, T. D. Galley and M. P. Müller, "The measurement postulates of quantum mechanics are operationally redundant," *Nature Communications* **10** (2019), 1361, [arXiv:1811.11060](https://arxiv.org/abs/1811.11060).
 6. N. D. Mermin, "Hidden variables and the two theorems of John Bell," *Reviews of Modern Physics* **65** (1993), 803–815, [arXiv:1802.10119](https://arxiv.org/abs/1802.10119).
 7. B. C. Stacey, "Quantum Theory as Symmetry Broken by Vitality," [arXiv: 1907.02432](https://arxiv.org/abs/1907.02432) (2019).
 8. S. Hossenfelder, "A derivation of Born's rule from symmetry," *Annals of Physics* **425** (2020), 168394, [arXiv:2006.14175](https://arxiv.org/abs/2006.14175).
 9. M. Appleby, S. Flammia, G. McConnell and J. Yard, "SICs and algebraic number theory," *Foundations of Physics* **47** (2017), 1042–59, [arXiv:1701.05200](https://arxiv.org/abs/1701.05200).
 10. 7 September 1999 note "Two Rabbis in a Bar" to John Preskill, in C. A. Fuchs, *Coming of Age with Quantum Information* (Cambridge University Press, 2011), pp. 399–401.

Copyright © 2021 by Blake C. Stacey. This article is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.