On the necessity for continuous reversible transformations in Lucien Hardy's five axioms for quantum theory

by Jan-Philipp Christ (LMU Munich). Date of submission: August 11, 2023

Lucien Hardy's five axioms for quantum theory are a remarkable attempt to derive the formalism of quantum mechanics from a few intuitive and reasonable principles [1]. These axioms have been subject to avid debate since their publication around 20 years ago. In particular, there have been discussions whether a different or even more minimal choice of axioms exists [2, 3] and about the 'reasonableness' of the axioms [4].

One of the axioms, axiom 5, states that there exist continuous reversible transformations between pure states. This axiom is crucial for ruling out classical probability theory and obtaining quantum theory instead, and seems to stand out from the other axioms, as it is the one that has not been replaced in alternative formulations of the axioms, such as [2, 3]. However, it is also the most debatable of the five axioms, as it seems to rely on some mathematical assumptions that are, at least as set forth in his original paper, not clearly and sufficiently motivated by physical considerations.

To argue for the 'reasonableness' and the motivation of the fifth axiom is the goal of this essay. It is structured as follows: We will first briefly summarize Hardy's axioms, how he builds quantum theory upon them and how *he* motivates his axioms. We will then in particular focus on his fifth axiom and try to argue whether it is indeed 'natural' and 'resonable' by taking into account both physical and information-theoretic considerations.

The setup Hardy introduces is the following: An experimenter has a preparation device, a transformation device, and a measurement device. The preparation device yields a *state*, which is "(that thing represented by) any mathematical object that can be used to determine the probability associated with the outcomes of any measurement that may be performed on a system prepared by the given preparation" [1]. The preparation can be thought of as the set of all interactions of, say, a particle with its environment. If exactly the same preparation is repeated, it is assumed that it will produce the same state. Note that this does not imply that the preparation is independent of the context in which it is performed, because we can include the context of the preparation, e.g. where the preparation device is located, as a property of the preparation device itself.

Hardy further defines the *dimension N*, equal to the maximum number of states distinguishable in a single measurement, and the *number of degrees of freedom K*, equal to the minimum number of measurements to specify the state. Indices for either of these numbers indicate the system they are referring to.

The axioms read in their original formulation:

Axiom 1 (Probabilities) *Relative frequencies (measured by taking the proportion of times a particular outcome is observed) tend to the same value (which we call the probability) for any case where a given measurement is performed on an ensemble of n systems prepared by some given preparation in the limit as n becomes infinite.*

Axiom 2 (Simplicity) *K is determined by a function of N (i.e. $K = K(N)$) where $N = 1, 2, \ldots$ and where, for any given N, K takes the minimum value consistent with the axioms.*

Axiom 3 (Subspaces) *A system whose state is constrained to belong to an M dimensional subspace behaves like a system of dimension M.*

Axiom 4 (Composite Systems) *A composite system consisting of two subsystems A and B having dimension N_A and N_B respectively, and number of degrees of freedom K_A and K_B respectively, has dimension $N = N_A N_B$ and number of degrees of freedom $K = K_A K_B$.*

Axiom 5 (Continuity) *There exists a continuous reversible transformation on a system between any two pure states of the system.*

A few remarks are in order: First, most details of Hardy's mathematical proofs do not heavily rely on the rather experimentalist notion of probability introduced here. It is also possible to work within a Bayesian framework, where probabilities are seen as degrees of belief, with only minor adjustments to

Hardy's line of reasoning [2]. Furthermore, the simplicity axiom turns out to be unnecessary, as Darrigol points out [4], and the composite systems axiom can be relaxed by not requiring $K = K_A K_B$, which is already implied in $N = N_A N_B$, as Hardy shows in his original work [1, Sect. 6.15].

Before defining pure states, as referenced in the fifth axiom, one should convince oneself that the definition of a state combined with the notion of the degree of freedom implies that K appropriately chosen measurements are sufficient (and necessary) to define the state. The corresponding probabilities (and thus the state) will represented by a column vector $\mathbf{p}$. The mixture, i.e. convex combination, of two such states is still a admissible state because it has definite values for the probabilities and the relation between states and probability vectors is one-to-one. Hence, the set of all such probability vectors is convex. The pure states are defined as the non-zero extremal points of this convex set, i.e. those points that cannot be expressed as a convex combination of any other two states. Intuitively, the name 'pure' is clear: Such states cannot be written as a mixture of two other states.

We will go on to outline how Hardy derives important features of quantum theory from the five axioms introduced here. Mathematical details will be skimmed over, but we encourage the reader to work through the details in [4] or [1], for which only very elementary linear algebra is required.

Hardy looks at a system of dimension $N + 1$. By considering a N dimensional subsystem and its one dimensional complement, he argues that $K(N + 1) \geq K(N) + 1$. Moreover, the axiom 4 about composite systems implies that $K(N)$ is completely multiplicative, i.e. $K(N_A N_B) = K(N_A)K(N_B)$. It can be proven that a function with these properties can be written in the form $K(N) = N^\alpha, \alpha \in \mathbb{R}$. The restriction that K is an integer now requires $r := \alpha = 1, 2, \ldots$, and the simplicity axiom makes us choose the minimal value for r which is still compatible with the other axioms.

Hardy proceeds by introducing K probability vectors $\{\mathbf{p}_k\}_{k=1,\ldots,K}$ chosen such that they uniquely specify the state of the system and also K *fiducial measurements* which can be identified with and represented by measurement vectors $\{\mathbf{r}_n\}_{n=1,\ldots,K}$.

It can be shown that the measurement vectors $\mathbf{r}^M$ are related to measurement probability vectors by a linear transformation represented by a real-valued matrix D. Conversely, state probability vectors can be represented by $\mathbf{r}$-type vectors, i.e. $\mathbf{p}^S = D\mathbf{r}^S$ and $\mathbf{p}^M = D^T\mathbf{r}^M$. Here, the superscripts M and S indicate that a vector refers to a measurement or state, respectively. Then the measurement probability is $p_{\text{meas}} = \mathbf{p}^M \cdot \mathbf{r}^S = (\mathbf{r}^M)^T D\mathbf{r}^S$.

On this basis, Hardy starts a proof by contradiction by assuming $K = N$, thus ruling out the case $r = 1$. He does this by considering N fiducial states which are also basis states, implying that D is equal to the identity matrix. However, it can be shown that this implies that the squares of the components of a fiducial vector $\mathbf{p}$ sum to 1. But $\mathbf{p}$ must also be both normalized, i.e. its components must sum to 1.This means that all but one component must be one, as in classical probability theory. But classical probability theory is incompatible with continuity axiom 5, because discrete fiducial vectors (which are also pure) cannot be connected by continuous transformations.

So we must have $r > 1$, and the simplicity axiom then requires $r = 2$, which does not contradict any of the other axioms. Hardy then goes on to show that the $N = 2$ case corresponds to the Bloch sphere. For general N quantum theory can be obtained using the $N = 2$ case. He also obtains the trace formula for calculating probabilities of measurements represented by some operator $\hat{A}$ and systems represented by a density matrix $\hat{\rho}$. Furthermore, he shows that the tensor product offers the right mathematical structure to describe composite systems (axiom 4 already fixes the 'right' dimension of the composite system) and that the trace of the density matrix needs to be invariant under all physically feasable transformations, thus restricting the number of set of all allowed transformations. Again, for the mathematical details we refer to Hardy's original paper, especially to Section 8.7 ff. For the purposes of this essay, it is only necessary to know where the various axioms come into play in recovering quantum theory from them, since the following discussions will be rather conceptual in nature.

It is interesting that Hardy's axioms, which are not defined in a mathematically rigorous way, give rise to rich mathematics, including the Hilbert state structure of quantum mechanics. In fact, Hardy himself stated as the goal of his axioms to show that it would have been *possible* for a 19th century ancestor of Schrödinger or the like to arrive at a quantum theory from reasonable assumptions alone, thus showing the inevitability of quantum theory. To prove this point, the simplicity with which the

axioms can be worded is thus quite intentional, as is the 'reasonableness' of the axioms.

Hardy does not give a clear motivation for his first axiom, suggesting that his choice of definition of probabilities was made for its conceptual simplicity. But he does point out that other notions of probability should lead to the same conclusions. And indeed, only a year after Hardy's publication, it was shown that the axioms also work in a Bayesian framework [2], thus eliminating the criticism that Hardy's insufficient distinction between measured frequencies and probabilities is problematic.

The second axiom is not fully motivated. Hardy appeals to "a certain constancy in nature such that K is a function of N" but does not argue why K should take the smallest value consistent with the axioms, even though he seems to be arguing akin to the spirit of Ockham's Razor, not requiring 'unnecessary' complexity from theories describing nature.

The third axiom is motivated by the fact that "we expect a probability theory pertaining to M propositions to be independent of whether these propositions are a subset or some larger set or not". We can also motivate it from a information-theoretic perspective by looking at an one-dimensional non-interacting spin-chain: If a M dimensional subspace of that system, i.e. M spins with the information content of 1 Bit each, would behave like a system of dimension $M' > M$, it would be possible to store more than M bit in that system. This, however, cannot be done because all other spins, i.e. all other bits, are fixed and cannot store any information. If, on the other hand, a M dimensional subspace of a length $M + N$ spin-chain composed of two non-interacting subsystems would behave like a system of dimension $M' < M$, the total information that could be stored would be $M' + N$ bits which is smaller than the total information capacity of the system ($M + N$ bits).

For the fourth axiom Hardy notes that "if subsystems A and B have N_A and N_B distinguishable states, then there must certainly exist N_A and N_B distinguishable states for the whole system". So the actual assumption of the axiom is not $N = N_A N_B$ but $N \leq N_A N_B$, which can be thought of as allowing only the minimal number of entangled states between the two subsystems, which is again reminiscent of Ockham's razor.

Remarkably, the fifth axiom, the continuity axiom, is the axiom that sets classical probability theory apart from quantum theory but also seems to be the least motivated amongst the five axioms. Hardy remarks "we expect to be able to transform the state of a system from any pure state to any other pure state [...] in a way that does not extract information about the state and so we expect this can be done by a reversible transformation". This can be considered to be a good motivation for reversibility. It is his argument for the continuity of the transformation to "expect any such transformation to be continuous since there are generally no discontinuities in physics" that can and should be questioned. It is indeed the case that in our macroscopic world usually no discontinuities arise, even if we try to provoke them. Here are two examples that a 19th century physicist might have come up with: When trying to modulate voltage along a square wave, the edges will always be smoothed out; when trying expand a gas very rapidly, you will find that the pressure exerted on the outer walls of the vessel used does change rapidly, but up on closer inspection there will be no jumps or the like. But looking at (discrete!) atomic spectra, for example, it is not clear a priori that the underlying processes are continuous. Since the point of quantum theory is to describe the structure underlying these very processes, it does not seem natural to require these transformations to be continuous.

In what follows, we will argue that it is nevertheless reasonable to assume and require that the transformations between pure states are continuous. We will argue in four ways: First, we will use a correspondence argument involving magnetically ordered systems. Second, we will exploit the fact that pure states can be characterized by their entropy. Third, we will use the continuity of time evolution to argue for the continuity of more general transformations. Finally, as an outlook, we will speculate whether perhaps the requirement of causal order already implies continuity.

This first argument is an argument by classical correspondence based on Darrigol [4]. We expect classical to emerge from quantum behavior when a sufficiently large number of non-interacting copies of the same system in the same pure state is considered. Here we will consider spins in the $\mathbf{u} \in \mathbb{R}^3$ direction where each spin is in the $|+_{\mathbf{u}}\rangle$ state. Magnetic precession, a classically known phenomenon, allows us to transform the total angular momentum $N\hbar\mathbf{u}/2$ to any $N\hbar\mathbf{u}'/2$ by applying an appropriately chosen magnetic field to the system. This precession happens continuously, and its end state requires each individual spin to be in the $|+_{\mathbf{u}'}\rangle$ state. Similar arguments can be made for arbitrary

spin states of the individual spins, thus making the possibility for continuous transformations between (any) two pure states plausible.

For our next argument we will make us of the fact that pure states have zero entropy $S(\mathbf{p}) \equiv -\mathbf{p} \cdot \ln(\mathbf{p})$ (the logarithm is applied component-wise). This concept is independent of the von Neumann entropy and can be motivated from an information theoretic perspective or from statistical mechanics. Pure states, by definition, cannot be written as a mixture of any other states. Thus, for any measurement performed on pure states, our knowledge of the measurement result is maximal and our ignorance, which can be interpreted as entropy, is minimal. In general $S(\mathbf{p}) \geq 0$. Zero is indeed reached as a lower bound because clearly $S((1, 0, \ldots, 0)^T) = 0$. Hence, all pure states have zero entropy. If, on the other hand, a state has zero entropy and it *could* be written as the non trivial mixture of two pure states $\mathbf{p}_a, \mathbf{p}_b$, then by the convexity of $x \mapsto x\ln(x)$ $S(\lambda\mathbf{p}_a + (1 - \lambda)\mathbf{p}_b) > \lambda S(\mathbf{p}_a) + (1 - \lambda)S(\mathbf{p}_b) = \lambda \cdot 0 + (1 - \lambda) \cdot 0 = 0$, contradicting the assumption. Hence, an alternative way to characterize the set of pure states is $\mathbf{p}$ pure $\iff S(\mathbf{p}) = 0$. This equation defines a differentiable manifold of dimensionality $K - 1$ on our set of state vectors. As such, it continuously connects all pure state vectors. It is not clear that all transformations on that manifold are physically admissible. But in combination with a correspondence argument like before we know that some transformations on that manifold can indeed be realized, thus very strongly suggesting that any two pure states are continuously connected by a reversible transformation.

For our third argument,

References

[1] Lucien Hardy. *Quantum Theory From Five Reasonable Axioms*. 2001. arXiv: `quant-ph/0101012` `[quant-ph]`.

[2] Rüdiger Schack. "Quantum Theory from Four of Hardy's Axioms". In: *Foundations of Physics* 33.10 (2003), pp. 1461–1468. DOI: `10.1023/a:1026044329659`. URL: `https://doi.org/10.1023%2Fa%3A1026044329659`.

[3] Lluí s Masanes and Markus P Müller. "A derivation of quantum theory from physical requirements". In: *New Journal of Physics* 13.6 (2011), p. 063001. DOI: `10.1088/1367-2630/13/6/063001`. URL: `https://doi.org/10.1088%2F1367-2630%2F13%2F6%2F063001`.

[4] Olivier Darrigol. "'Shut up and contemplate!': Lucien Hardy's reasonable axioms for quantum theory". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 52 (2015), pp. 328–342. ISSN: 1355-2198. DOI: `https://doi.org/10.1016/j.shpsb.2015.10.002`. URL: `https://www.sciencedirect.com/science/article/pii/S135521981500091X`.