Introduction to Quantum Information: Hilary Term 2020

Luca Mondada

3rd April 2020

Why Born's rule?

Why Born's rule?

At first glance, the formalism of quantum mechanics, and thus of quantum information, appears particularly arbitrary. Its formulation seems far off from our intuitive understanding of classical dynamics and, more generally, from any other physical theory: it is for example still unclear how the mathematics of quantum theory translate into relativistic behaviour.

Among all postulates of quantum mechanics, Born's rule is arguably the most mysterious. While quantum states and operations can elegantly be modelled in Hilbert space, the measurement postulates underlying Born's rule seem to sit awkwardly within this framework. It is hard to justify why Born's rule should make accurate predictions, or indeed why the concept of measurement should play such a prominent and apparently asymmetric role in quantum theory.

This question of quantum foundation becomes all the more relevant when one considers how Born's rule plays a central role in the theory. The measurement postulate is the connection between Schrödinger's mathematical formalism and real world experiments. Without the Born rule, quantum information science cannot make any predictions — it is no science.

In the absence of any straightforward physical argument to justify Born's rule, we will argue in this essay that Born's rule has to be understood within the broader structure of quantum mechanics. By looking closely at the physical principles that underlie the mathematical framework of quantum physics, we will uncover relations between different aspects of the theory that are not evident at first glance. Born's rule will then emerge as an inseparable part of a larger tight knot of physical postulates and mathematical consequences. As we try to construct hypothetical alternatives, we will see that a quantum theory without Born's rule becomes increasingly hard to fathom.

Our essay will proceed in three sections. The first part will lay the foundations for our understanding of quantum mechanics and Born's rule. Our focus will be on the mathematical framework that is needed to model quantum theory as we understand it. We will take interest in the formalism that arises, as well as the necessary physical postulates that underlie the theory. This will bring us to discuss what measurement theory in its broadest sense must entail. We will conclude this first part with a short discussion of quantum logic to give a different justification for the mathematical constructions that we will rely on in the rest of the essay.

In the second part, our discussion will focus on developing the measurement theory around our quantum formalism. We will analyse what properties must hold to be coherent with the rest of the theory and hypothesise a family of plausible measurement theories. Studying the properties of the different measurement postulates will single out Born's rule as the only measurement postulate coherent with the rest of quantum theory.

This will lead us to the third and final section, in which we will attempt to formalise the argument for Born's rule. We will introduce axiomatic reconstructions of quantum theory and will focus in particular on recent work by Masanes and colleagues [10]. Remarkably, they present a derivation of Born's rule relying only on non-measurement related postulates of quantum theory, suggesting strongly that Born's rule can indeed be seen as a consequence of the structure of the rest of quantum theory.

1 The structure of quantum theory

We call the data that defines a physical theory its structure. This mathematical construction gives a framework to study the properties of the theory and to describe and reason about physical systems. From a theoretical perspective, the structure is also essential to express physical postulates and formulate derivations of physical laws from principles.

The structure of quantum theory thus provides the building blocks necessary to any discussion of Born's rule. In this first part of our essay, we would like to not only describe this structure, but also discuss some of its underlying postulates. As we will see in later sections, these axioms play a central role in interpreting Born's rule, as they form the premises on which any derivation of Born's rule is founded.

The Hilbert space structure

Let us start our considerations with a physical system $\mathfrak{S}$. Denote its state space $\mathcal{H}$ and consider two possible states of this system $\psi_1 \in \mathcal{H}$ and $\psi_2 \in \mathcal{H}$. The first key physical property that characterises quantum mechanics is the possibility to create a state that is a superposition of ψ_1 and ψ_2; denote this new state $\psi_1 + \psi_2$. In fact, not only can we superpose them, but any weighted superposition of ψ_1 and ψ_2 should be allowed[1]:

$$\lambda_1\psi_1 + \lambda_2\psi_2, \quad \lambda_1, \lambda_2 \in \mathbb{C}. \qquad [1]$$

This defines addition and scalar multiplication, so that $\mathcal{H}$ forms a complex vector space. In accordance with our intuition, we think of state vectors $\psi \in \mathcal{H}$ and $\psi + \psi$, or, in general, of states $\lambda\psi$ for any $\lambda \in \mathbb{C}$ as representing the same physical state. To formalise this, we define

Definition 1 (Normalised states). *Given a state space $\mathcal{H}$ forming a complex vector space, we define the normalised state space $P\mathcal{H}$ through the identification*

$$P\mathcal{H} := \mathcal{H}/_{\mathbb{C}\setminus\{0\}},$$

ie the quotient space under identification of states with different normalisations. Elements of $P\mathcal{H}$ are often called projective rays of Hilbert space.

For simplicity of notation, we will often refer to states $\psi \in P\mathcal{H}$, in which case we mean a representative from the equivalence class that should be clear from context (typically such that $\|\psi\| = 1$). We call these representatives the normalised states.

We now consider a physical operation on our system $f : \mathfrak{S} \to \mathfrak{S}$. It seems natural to ask that performing the operation f to ψ_1 and ψ_2 and then superposing the obtained

[1] We choose here the complex numbers $\mathbb{C}$ as scalars, because this corresponds to the standard description of quantum theory. Note that a priori there is no way to justify one particular field. See eg. [12] for a discussion.

states should be equivalent to applying f to the superposition 1. In an equation:

$$\lambda_1 f\psi_1 + \lambda_2 f\psi_2 = f\left(\lambda_1\psi_1 + \lambda_2\psi_2\right).$$

This is linearity! These observations thus give us grounds to suspect that physical states form a vector space with superposition and scalar multiplication, and that any physical operation on the system should be linear.

The state space in quantum theory is usually considered as a Hilbert space. To endow our vector space with a Hilbert structure, what is missing is an inner product on that space. This bilinear form (or rather, sesquilinear form in the complex case) is used to formulate another property of quantum mechanics: there are sets of states that can be reliably differentiated by physical operations on the system — call these compatible states — while others cannot. Choose a maximal set of such compatible states and fix their normalisation. Defining this as an orthonormal basis will directly induce an inner product structure on the state space.

However, such sets of compatible states must not necessarily translate into a valid and uniquely defined inner product. We are thus implicitly making a range of assumptions when we accept the Hilbert structure of state space. Firstly, linearity of the inner product implies that one set of maximal compatible states immediately fixes all other such sets. Furthermore, assuming that such a basis of compatible measurements exists boils down to postulating that any physical state of the system can be fully described as a linear superposition of this fixed set of states. This in turns means not only that all sets of compatible states are fixed, but also that the "degree of incompatibility" between any two states is entirely determined and quantified by the inner product structure. Finally, there are also some necessary mathematical postulates which we will not discuss here, in relation with the completeness assumption of the metric space induced by the Hilbert structure.

Some of these postulates and their consequences are not exactly physically indisputable — far from it. In one of the attempts to justify this Hilbert space structure and explore its consequences, a coherent system of propositional logic that departs from the classical Boolean algebras was developed to reason about quantum mechanical systems. This has led to some remarkable statements about the relation between the Hilbert structure of quantum mechanics and its underlying logical algebra that will be briefly presented in the final paragraphs of this section.

For now, we will accept the postulate that state space in quantum theory forms a complex Hilbert space. The natural next step in our discussion towards Born's rule is the introduction of a measurement formalism. We have so far in our treatment of quantum mechanical systems avoided any references to measurements, but the concept has underlain much of our discussion of inner products. The constructions are inseparable in that compatible states, as defined by the inner product, are inevitably related to physical measurements performed on the system.

Measurement theory

In the most general sense, a measurement f given a state should tell if a physical property holds. We will adopt the point of view that measurements may well be unsuccessful if a property does not hold, in which case they do not produce any result. This can be made equivalent to the perhaps more familiar concept in which a measurement returns one of the k possible outcomes in some set $\mathcal{C}$, which we will call a complete measurement. They are given in our formalism by a set of measurements $\{f_1, \dots, f_k\}$, corresponding to each of the k possible outcomes in $\mathcal{C}$.

Given our understanding of the non-deterministic characteristics of measurements in quantum mechanics, it makes sense to model the output of a measurement not as truth/false binary result, but rather as a probability. From our physical understanding of measurement processes in quantum mechanics, we can view these probabilities in the frequentist interpretation as the distribution of the measurement outcome over repeated trials on identically prepared states in the infinity limit.

Definition 2 (Measurement). *A measurement on a quantum state ψ is given by a function on the normalised state space*

$$f : P\mathcal{H} \to [0,1].$$

A complete measurement that yields a unique value from a finite (or countably infinite) set $\mathcal{C}$, $|\mathcal{C}| = k > 0$ is given by a family of measurements $\{f_1, \dots, f_k\}$ such that

$$\sum_{i=1}^{k} f_i(\psi) = 1, \quad \text{for all } \psi \in P\mathcal{H}. \qquad [2]$$

What are potential measurements in our Hilbert theory? As hinted earlier in our discussion, a good first attempt would be to use what we already have at our disposition: the Hilbert space structure and its inner product. Beyond the Hilbert space postulates of quantum mechanics that we discussed earlier, relying on the inner product structure to define measurements also requires a further assumption. We ask that to every state there be a unique measurement that tests for this state and distinguishes it from the other compatible states. Let us make this more precise: we are asking that for any compatible states $\psi_1, \dots, \psi_n$, there be measurements $m_1, \dots, m_n$ that satisfy

$$m_i(\psi_j) = \delta_{ij}, \text{for all } 1 \le i, j \le n. \qquad [3]$$

In other words, we ask that the measurement space $\mathcal{M}$ be the dual of state space: $\mathcal{M} \cong \mathcal{H}^*$. Within our standard Hilbert postulates, we denote state space with $\mathcal{H}$ and measurement space as its dual $\mathcal{H}^*$. Using Dirac notation, we then write measurements as covariant vectors $\langle m| \in \mathcal{H}^*$ and states as contravariant vectors $|\psi\rangle \in \mathcal{H}$.

Definition 3 (Measurement rule). *A measurement rule is given by a map from the normalised measurement and state space to probability space*

$$P\mathcal{H}^* \times P\mathcal{H} \to [0,1].$$

A toy model: the inner product as a measurement rule

A more in-depth discussion of some possible measurement rules and their properties is the object of section 2. In this few paragraphs, we would like instead to introduce a toy model of measurement as a first discussion of the concept. As a historical sidenote, we might add that this toy measurement rule corresponded to Born's original suggestion in his 1926 paper [2].

In the light of the above, it seems straightforward to define a measurement rule that is directly induced by the inner product

$$\langle\cdot|\cdot\rangle : \mathcal{H}^* \times \mathcal{H} \to \mathbb{C} \tag{4}$$

that maps a measurement and a state to a scalar. This is already almost of the correct type. The last aspect to ensure is that our rule is properly normalised, that is, for each element $S \in P\mathcal{H}$ and $M \in P\mathcal{H}^*$ we need to single out an element $|\psi\rangle \in S$ and $|m\rangle \in M$ to represent the equivalence class and define the measurement rule on $P\mathcal{H}^* \times P\mathcal{H} \to [0, 1]$. One way of proceeding is to fix the measurement $\langle m|$ such that $\||\langle m|\|\|_2^2 = |\langle m|m\rangle| \leq 1$ and then for each $S \in P\mathcal{H}$ choose $|\psi\rangle \in P$ such that $\||\psi\rangle\|_2^2 = |\langle\psi|\psi\rangle| = 1$ and

$$\langle m|\psi\rangle \in \mathbb{R}^{\geq 0}. \tag{5}$$

With Cauchy-Schwarz it follows that this defines a measurement rule

$$\begin{aligned} \langle\cdot|\cdot\rangle :& P\mathcal{H}^* \times P\mathcal{H} \to [0, 1] \\ & \big(\langle m|, |\psi\rangle\big) \mapsto \langle m|\psi\rangle \end{aligned} \tag{6}$$

for properly normalised states $\langle m| \in \mathcal{H}^*, |\psi\rangle \in \mathcal{H}$.

Remark. *Note that the vector in $|\psi\rangle \in \mathcal{H}$ that represents some quotient group in $P\mathcal{H}$ to define the measurement rule 6 might depend on the measurement $\langle m|$! In particular, this rule is no longer linear.*

Equivalently, we can relax the normalisation condition on states to $\||\psi\rangle\|_2 = 1$ (without the phase normalisation condition 5) and define the measurement rule as

$$\begin{aligned} \langle\cdot|\cdot\rangle :& P\mathcal{H}^* \times P\mathcal{H} \to [0, 1] \\ & \big(\langle m|, |\psi\rangle\big) \mapsto |\langle m|\psi\rangle|. \end{aligned} \tag{7}$$

This measurement model has interesting properties. It exhibits for example interference as we know it in quantum mechanics. Given two orthogonal basis states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, consider states $|\psi_1\rangle = \alpha|\varphi_1\rangle + \beta|\varphi_2\rangle$ and $|\psi_2\rangle = -\alpha|\varphi_1\rangle$ for some $0 < \alpha, \beta \leq 1$ and assume they are normalised. Then, whilst for both states the probability $\langle\varphi_1|\psi_i\rangle$ is non-zero, the prbability of measuring $\langle\varphi_1|$ in their superposition

$$\langle\varphi_1|\psi_1 + \psi_2\rangle = \beta\langle\varphi_1|\varphi_2\rangle = 0.$$

This thus seems to be a reasonable choice of measurement rule! We will pick up this discussion in the next section, where we will study the physical properties of this measurement rule and of others. We will see that the rule as it is has some undesirable properties. We can actually already see issues with the normalisation conditions that we had to impose in the inner product measurement rule. The phase normalisation given in 5 in particular seems clumsy at best. This appears to be addressed by the reformulated rule 7, but we will see when we will discuss its generalisation to the projection operator formalism that in fact this does not resolve the problem.

Before this, we will conclude this section by making a short aside to attempt to justify the Hilbert structure of quantum theory, which is at the centre of our argument.

Why Hilbert spaces: an aside into quantum logic

It has long become a platitude to say that quantum mechanics is puzzling. Quantum logic is a simple yet remarkably powerful idea that was suggested as early as the 1930s by Birkhoff and von Neumann to help to think about quantum mechanics and its structure [1].

The key idea is that any physical theory with measurements defines a propositional logic in which measurements are seen as statements. Indeed, to any measurement corresponds a set of states for which the measured property holds — in the quantum setting, we say that a property holds if it holds with probability 1. We can then combine and negate such properties to form conjunctions, disjunctions and complements of measurements.

In the classical case, this model of logic corresponds exactly to boolean algebras, since states that have a given property form sets and the logical combination of physical properties correspond to intersection, union and complement on these sets. The laws of Boolean algebras do not necessarily hold in quantum mechanics, however. Let us consider the distributivity law of set theory for example:

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C), \quad \text{for any sets } A, B, C.$$

We take the quantum mechanical qubit in the $\mathbb{C}^2$ Hilbert space, with $\{|0\rangle, |1\rangle\}$ forming one orthogonal basis and $\{|+\rangle, |-\rangle\}$ forming the complementary one. Consider the measurements given by

$$a = \langle 0| \qquad\qquad b = \langle 1| \qquad\qquad c = \langle +|$$

The respective eigenspaces A, B, C of these measurements form the statements in our quantum logic. Then we can see that

$$B \wedge C = 0 \qquad\qquad A \vee B = \mathbb{C}^2 \qquad\qquad A \vee C = \mathbb{C}^2.$$

Thus, we get

$$A \vee (B \wedge C) = A \neq \mathbb{C}^2 = (A \vee B) \wedge (A \vee C).$$

Thus quantum logic violates distributivity and is distinct from Boolean logic! Note that if we restrict ourselves to compatible measurements, we recover classical logic, so that quantum logic is strictly more general. A weaker form of distributivity, orthomodularity, can be shown to hold in quantum logic. A formal treatment of quantum logic is given in [14].

The question can then be reversed to attempt to classify all models of quantum logic, in the hope of justifying the Hilbert structure as the unique model of quantum logic. It emerged that few simple axioms for quantum logic bring us remarkably close to Hilbert structure [13]. However, while first steps in quantum logic appeared impressively simple, further research showed that pinpointing the exact logical axioms that would lead to a unique Hilbert model of quantum logic was more challenging. The culmination of years of research in the area was the result by Solèr, giving an explicit set of axioms that could single out the Hilbert model of quantum logic [16].

Showing that the rich structure of quantum mechanics can be completely derived from purely logical and lattice-theoretical considerations is a considerable feat. This means that quantum mechanics can indeed be seen as the mere consequence of a propositional calculus of measurements that differs from our classical intuition, giving us a new perspective on quantum theory. This is a remarkable result and to this day, it provides arguably one of the best justifications for the theory of quantum mechanics.

Nonetheless, the approach has genuine limitations [17]. The axioms postulated by Solèr are no longer easily physically justifiable, making the details of the argument convoluted and the evidence not quite as compelling. Rather than an indisputable proof of the Hilbert structure of quantum mechanics, what emerges from this discussion is further evidence that all aspects of quantum mechanics are intertwined. Quantum information processing as we understand it is a direct consequence of this logical structure, and would not exist in any non-Hilbert model.

2 Deriving Born's rule

In the previous section, we described how state space in quantum mechanics can be seen as a Hilbert space and how measurement theory fits in that frame. We will now examine in more depth what measurement postulates based on the inner product can look like in the hope of justifying and singling out Born's rule as the only coherent measurement theory.

The trouble with the toy model

Recall the "toy model" measurement rule expressed in 7, which we will from now on call R_1 (we will see why this notation makes sense soon):

$$R_1 : (p, \psi) \mapsto |\langle p|\psi\rangle|, \qquad [8]$$

for any normalised measurement $\langle p|$ and state $|\psi\rangle$. Note that in this notation, the normalisation condition simplifies to $\||\psi\rangle\|_2^2 = 1$ and $\|\langle p|\|_2^2 \leq 1$. Before we proceed to generalising our considerations to a wider range of measurement rules, we will attempt to reformulate 8. In fact, given that measurements themselves form a Hilbert space, we can consider the superposition of measurements

$$\langle p| := \alpha \langle p_1| + \beta \langle p_2|,$$

for some normalised measurements $\langle p_1|$ and $\langle p_2|$ and scalars α, β such that $\langle p|$ is normalised. Then because of non-linearity,

$$\begin{aligned} R_1(p,\psi) &= \left|\alpha \langle p_1|\psi\rangle + \beta \langle p_2|\psi\rangle\right| \\ &\neq |\alpha|\left|\langle p_1|\psi\rangle\right| + |\beta|\left|\langle p_2|\psi\rangle\right| = |\alpha| R_1(p,\psi) + |\beta| R_2(p,\psi). \end{aligned} \qquad [9]$$

This is not a problem per se, but it makes our formalism inconvenient, given that we might want to consider the right hand side as a measurement in its own right: the expression represents the physical operation where either measurement in the superposition happens with some probability, and the entire measurement is considered successful if the submeasurement $\langle p_1|$ or $\langle p_2|$ that was performed is successful. Our current notation forbids such measurements as they cannot be expressed in the form of 8.

This motivates the introduction of the familiar projection operator formalism for measurements. To the normalised measurements $\langle p_1|$ and $\langle p_2|$, we associate the projectors $P_1 = |p_1\rangle\langle p_1|$ and $P_2 = |p_2\rangle\langle p_2|$. To express the measurement R_1 in the projection operator formalism, we introduce:

Definition 4 (p-norms)**.** *Fix a basis $\varphi_1, \dots, \varphi_n$ of the Hilbert space $\mathcal{H}$. For each $\psi \in \mathcal{H}$, this defines unique coordinates $(x_1, \dots, x_n)$ given by $x_i = \langle \varphi_i|\psi\rangle$. We can use these to define a family of metric space norms*

$$\|\psi\|_p := \left(|x_1|^p + \cdots + |x_n|^p\right)^{1/p},$$

satisfying positive-definiteness, absolute homogeneity (ie $\|\lambda v\| = |\lambda|v)$ and the triangle inequality.

Given this definition and the projector formalism, we can reformulate the right hand side of 9 as

$$\begin{aligned} |\alpha|\Big|\langle p_1|\psi\rangle\Big| + |\beta|\Big|\langle p_2|\psi\rangle\Big| &= |\alpha|\|P_1|\psi\rangle\|_2 + |\beta|\|P_2|\psi\rangle\|_2 \\ &= |\alpha|\left\|\widetilde{P_1}|\psi\rangle\right\|_1 + |\beta|\left\|\widetilde{P_2}|\psi\rangle\right\|_1 \\ &= \left\|(\alpha\widetilde{P_1} + \beta\widetilde{P_2})|\psi\rangle\right\|_1, \end{aligned}$$

where P_1, P_2 and P must be renormalised in the 1-norm to obtain $\widetilde{P_1}, \widetilde{P_2}$ and $\widetilde{P}$. That is,

$$\widetilde{P_1} = |\tilde{p}_1\rangle\langle p_1| := \frac{1}{\|p_1\|_1}|p_1\rangle\langle p_1|,$$

and similarly for $\widetilde{P_2}$.

This reformulation can be used to generalise our measurement rule R_1 to the projection operator formalism.

Definition 5 (1-norm measurement rule)**.** *For any observable M and state $|\psi\rangle$, the probability of measuring M on $|\psi\rangle$ is given by*

$$R_1 : (M, \psi) \mapsto \|\tilde{M}|\psi\rangle\|_1, \tag{10}$$

where $\tilde{M}$ is the projection M renormalised to the 1-norm, which is given by renormalising the eigenvalue λ of each eigenvector v_λ to $\tilde{\lambda} := \frac{\|v_\lambda\|_2}{\|v_\lambda\|_1}\lambda$.

The renormalisation of the observable M means that $\tilde{M}$ is no longer necessarily a projector. As opposed to the usual quantum formalism, observables in the 1-norm measurement rule are not given by projectors, and a complete set of measurements must not necessarily sum to the identity. More importantly, the renormalisation depends on the choice of one particular basis of Hilbert space! This basis dependency seems hard to justify and will be our subject of discussion as soon as we have formulated some alternatives to our toy model. Luckily for us, our construction is easily adapted to further measurement rules.

Generalising our attempt

Looking at the new formulation of the measurement rule as the 1-norm rule given by 10, there is immediately a wealth of variations of the 1-norm rule that can be considered. In the rest of this section, we will look at the family of rules given by

Definition 6 (p-norm measurement rule)**.** *Fix $p \geq 0$. We define the measurement rule R_p given for any observable M and state $|\psi\rangle$ by the probability*

$$R_p : (M, \psi) \mapsto \|\tilde{M} |\psi\rangle\|_p^p,$$

where the renormalisation $\tilde{M}$ is given by $\tilde{\lambda} := \frac{\|v_\lambda\|_2}{\|v_\lambda\|_p}\lambda$ for each eigenvector v_λ with eigenvalue λ.

Notice that we restrict our considerations to p-norms raised to the power of p. Other alternatives could be discussed as well. Informally, our choice can be somewhat justified by noting that the p-th roots that appear in the norms are not smooth functions at 0 — taking the p-th power resolves this, and means that the considered measurement rules are "more smooth"[2].

Basis-dependency in the measurement rule

We alluded earlier to the fact that the definition of p-norms crucially depends on a fixed basis. In these paragraphs, we would like to detail this basis dependency further.

Consider the identity measurement $M = \mathrm{id}$ that is unchanged by any p-norm renormalisation. Two bases $\mathcal{B}$ and $\mathcal{B}'$ that define the two p-norms $\|\cdot\|_p$ and $\|\cdot\|_p'$ will yield the same predictions for a measurement if their norms coincide

$$\|M |\psi\rangle\|_p = \|M |\psi\rangle\|_p' \quad ie \quad \|\psi\|_p = \|\psi\|_p' \quad \text{for all } \psi \in P\mathcal{H}.$$

Now, changing the basis $\mathcal{B} \to \mathcal{B}'$ amounts to performing a linear transformation U on the coordinates of ψ. Thus we must have

$$\|\psi\|_p = \|U\psi\|_p \quad \text{for all } \psi \in \mathcal{H}.$$

In other words, the change of basis transformation U is an isometry of the p-norm! There is a standard result of p-norms on complex vector spaces that comes in handy at this point.

Theorem 7 (Continuous isometries)**.**

(*a*) *For any $1 \leq p \leq \infty$ and $p \neq 2$, a map $\mathcal{U} : \mathbb{C}^n \to \mathbb{C}^n$ is an isometry if and only if $\mathcal{U}$ is a complex permutation matrix, ie a permutation matrix where 1-entries may be any complex number of unit amplitude.*

(*b*) *For $p = 2$, the isometries $\mathcal{U} : \mathbb{C}^n \to \mathbb{C}^n$ are given by the real Lie group of unitary matrices $U(n)$.*

Proof. These are standard results. See for example [9] for part (a), and [6] for part (b). □

[2] Note that if p is odd, the measurement rule R_p might still be not smooth, since $|x|^p$ is not a smooth function.

The difference between $p = 2$ and the other cases couldn't be starker. For $p = 2$, the measurement rule is basis independent, since any unitary transformation is a valid basis change that preserves the measurement rule. In all other cases, no two distinct unordered bases yield the same predictions! Having to singling out a basis to make physical predictions goes against the fundamental space symmetry assumption of physics. In that sense, any measurement rule that cannot be formulated as basis invariant has to be rejected.

Picturing time evolution

Perhaps an even clearer illustration of this space asymmetry is obtained when considering the time evolution of such physical systems. In the Heisenberg representation of quantum mechanics, an evolving state of a quantum mechanical system can dually be seen as a fixed system in a changing measurement basis. Thus the allowed isometries U are in one-to-one correspondence with the allowed evolutions of a physical system. For $p = 2$, the isometries are the Lie group of unitary matrices, so that possible physical evolutions of the system form a smooth manifold. In this case, state evolution can be given by a smooth path in the space as a function of the time.

In contrast, in the case $p \neq 2$, the only physical evolutions allowed form are discrete group. How could a time evolution that is discrete correspond to our understanding of the physical world? Suppose the time evolution operator from time 0 to T is given by U but that there is no continuous path between the identity and U. Then what is the state of the system at some time $0 < t < T$?! And what to make of the infinite velocity, infinite acceleration and infinite forces that this implies?

It seems clear that time evolution of a physical system must be continuous. Thus we conclude that the only measurement rule that can be considered is given by $p = 2$. In terms of information processing, a physical evolution that is limited to permutations of basis states is equivalent to classical computing, in which computations are performed by operating discretely on bits. This means that the case $p = 2$ is the only case where physics give us a new computational model!

With this we finally recover Born's rule:

Definition 8 (Born's rule)**.** *Given an observable M and a normalised state $|\psi\rangle$, the probability of the measurement to be successful is given by*

$$\|M|\psi\rangle\|_2^2 = \left|\langle m|\psi\rangle\right|^2 = \langle\psi|M|\psi\rangle\,. \qquad [11]$$

It is telling that the prediction rule for $p = 2$ takes the elegant form of the right hand side of 11, where the 2-norm does not appear explicitly; in many respects, Born's rule seems to come "naturally" within the Hilbert structure of quantum mechanics. The rule is basis independent, the normalisation conditions simplify to orthonormalisation and observables are given by projectors.

Notice how all these properties are closely related with the inner product structure of quantum theory. In fact, while every other norm requires the data of a basis and the additional norm definition, the 2-norm is already given by the Hilbert structure:

$$\|\psi\|_2^2 = \langle\psi|\psi\rangle\,.$$

Every aspect of the measurement theory as given by Born's rule is already naturally embedded in the Hilbert space structure. We conclude by stating the following result that captures this intuition:

Theorem 9. *The p-norm induces a Hilbert space norm on $\mathbb{C}^n$ if and only if $p = 2$.*

Proof. This is a standard result of geometry of metric spaces. It is straightforward to see that any Hilbert space norm of the form $\|x\| = \sqrt{\langle x|x\rangle}$ satisfies the parallelogram law

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

It can be verified that this identity only holds for the p-norm if $p = 2$ (see [4]). □

Our discussion so far has highlighted how all aspects of quantum theory are closely intertwined. It has hopefully transpired that all characteristics of quantum mechanics, such as state superposition, interference, quantum logic and Born's rule are just different faces of the same physical theory – it seems impossible to build a convincing physical theory with some of these properties but not others. In particular, measurement theory cannot be dissociated from the rest of quantum theory. The aim of the final section will be to formalise this intuition.

3 A formal axiomatic approach

We have so far attempted to convince ourselves from a trial and error approach that Born's law must be the only measurement postulate that is compatible with the rest of the structure of quantum theory. Our aim in this third and final section of the essay will be to formalise this statement and present a rigorous axiomatic reconstruction of measurement theory and Born's rule from physical principles.

The measurement postulates

One of the tricky aspects of axiomatic reconstruction attempts of physical theories is the formulation and justification of postulates that are necessary for any useful theory. In our previous exposition of quantum theory, we have made several assumptions throughout.

The postulates around measurement theory have always been among the most criticised and disputed. While the general structure of quantum states and superpositions thereof seem rather reasonable and find some justification in the accepted propositional logic of quantum mechanics, some of the postulates surrounding measurement theory and Born's rule seem more arbitrary. Our presentation in the last section gave some intuition about the place of Born's rule within Hilbert structures, but our approach was not formal and open questions remain. Why is the compatibility between different states and measurements given by an inner product structure? Why is it that measurements form a Hilbert space isomorphic to state space in the first place? And while Born's rule seems coherent with the rest of quantum mechanics from our trial-and-error approach, how can we be sure that there is no other coherent rule? These questions turn out to be harder than one might first expect.

A quick history of quantum reconstruction

One of the earliest successful attempts to provide a mathematical derivation of Born's rule from other postulates was Gleason in reference [5]. This result was instrumental in highlighting the place of Born's rule within Hilbert space structure. However, the derivations still relied on strong assumptions on the structure of measurements in quantum theory. It became therefore clear that Born's rule was a consequence of the measurement structure of quantum theory, but the question of the origin of this structure remained unanswered.

Axiomatic approaches from physical principles have since tried to address this. The idea to reconstruct quantum theory from defensible postulates was popularised in work by Hardy [7, 8]. Theorists have since tried to single out the simplest and most "physically justifiable" axioms that lead to quantum theory [11, 15, 17]. In our essay, we will discuss one of the most recent successes of axiomatic reconstruction by Masanes and colleagues [10]. They show for the first time that Born's rule can indeed be recovered from physical postulates that they claim are entirely unrelated to the structure of measurements[3]. This is a profound statement that promises to provide a more formal framework to our discussion on the origin of Born's rule.

Masanes' statement

We will go through the most important assumptions of the derivation. However, we will not detail here the exact postulates nor the derivations made by Masanes. For that, we refer instead to the original publication [10].

The starting point are the usual postulates of the structure of state space in quantum theory as we presented them in section 1. The pure state of a system are given by rays in complex Hilbert space (ie elements of $P\mathcal{H}$); reversible transformations of pure states are unitary transformations $\mathbb{C}^d \to \mathbb{C}^d$; and the joint pure state of two quantum system is given by their tensor product.

Remark. *Note that this treatment of quantum theory holds not only for finite-dimensional Hilbert spaces, but also for countably infinite dimensional ones. All such infinite dimensional spaces are isomorphic and we denote them too with* $\mathbb{C}^d, d = \infty$*.*

Beyond these main postulates, the authors reason about measurements in a similar fashion to the formalism we introduced in the first sections. Note how these postulates remain very general about the structure of measurements.

Measurements: are given by function of the type $f : P\mathcal{H} \to [0, 1]$ with their closure under among others probability mixtures and precomposition with unitaries. Importantly, the closure is assumed to be well-behaved with regards to restriction to subsystems and tensor products (details in the paper).

Possibility of state estimation: any system with finite-dimensional state space should admit a finite list of measurements whose results can characterise any state.

The final statement then reads

Theorem 10 (Masanes [10]). *The only measurement postulate satisfying the "possibility of state*

[3]This can be disputed as they assume the possibility of state estimation: see below.

estimation" is Born's rule: for any measurement $\mathbb{C}^d \to [0,1]$*, there is a linear operator* $F : \mathbb{C}^d \to \mathbb{C}^d$ *such that*

$$f(\psi) = \langle\psi| F |\psi\rangle, \quad \textit{for any normalised state } \psi.$$

F has real eigenvalues $0 \le \lambda_i \le 1$*. It is in particular positive semi-definite.*

This is essentially a formal treatment of our discussion in the second section. It makes apparent that the inner product of the Hilbert space gives rise to a unique measurement rule that is consistent with the other postulates of quantum theory. However, no derivation can entirely justify a measurement theory without making some sort of assumption on what measurements really are.

Beyond the broad definition of measurement as a function from state space into probability space, this result puts another meaning onto measurements. It assumes that finite-dimensional physical systems in principle are entirely defined by a finite set of measurements. This is sensible in so far as properties or states that cannot be measured are impossible to study physically in the first place. Critically, however, this assumption goes further than that, in that it assumes that the set of measurements that can define any possible state of a finite-dimensional system is itself finite. Although this is certainly true in any physical theory that we know of, there is no simple *a priori* justification for this.

While there might never be a derivation of quantum theory from absolutely indisputable axioms, Masanes' result, together with all previous work in the area, seem to make it clear that the measurement postulates given by Born's rule are as coherent with the rest of quantum theory as no other rule is.

4 Conclusion

From the various perspectives on quantum structure we have discussed, it seems that we can claim with fair confidence that the origin of Born's rule lies in the Hilbert structure of quantum mechanics itself. This essay showed how all aspects of quantum theory from its Hilbert structure to its measurement laws are closely related. In spite of – or perhaps precisely because of – all the theoretical efforts that over the years have gone into studying the foundations of quantum theory, there seems to be no alteration to accepted physical postulates or alternative theory that would be physically defensible. Instead, what emerges is a mathematical construction that is much more coherent and interlinked than it might have seemed.

In terms of information processing, what this means is it seems that no alternative to Born's rule would yield a computational theory as rich as quantum theory. In particular, we saw in our discussion of norm-based measurement postulates that any of the $p \neq 2$ norms yields a theory with a computational power that is equivalent to classical machines.

Nonetheless, many physicists remain unsatisfied with the current theory. Quantum mechanics still lacks a foundational argument that could be compared for example to the derivation of relativity from simplest postulates. In [3], Bub makes the analogy to the Lorentz interpretation of special relativity, that described relativistic behaviour of particles in terms of a fixed ether and complex formulas. The predictions made by the theory were accurate, but theorists had no means of understanding the physical statement of relativity until Einstein provided simple postulates that would justify the predictions.

A similar argument can be made to argue that as long as quantum mechanics remains obfuscated by the operational intricacies of the current theory, the nature of quantum theory will not be understood. The search for the true foundations of quantum mechanics is thus still on. What seems clear, however, is that this breakthrough will require more work than a slight adjustment to Born's rule.

References

[1] Garrett Birkhoff and John von Neumann. "The Logic of Quantum Mechanics". In: *The Annals of Mathematics* 37.4 (Oct. 1936), p. 823. DOI: 10.2307/1968621.

[2] Max Born. "Zur Quantenmechanik der Stoßvorgänge". In: *Zeitschrift für Physik* 37.12 (1926), pp. 863–867. DOI: 10.1007/BF01397477.

[3] Jeffrey Bub. "Quantum Probabilities: An Information-Theoretic Interpretation". In: *Probabilities in Physics*. Oxford University Press, Sept. 2011, pp. 231–261. DOI: 10.1093/acprof:oso/9780199577439.003.0009.

[4] Cyrus D Cantrell. *Modern mathematical methods for physicists and engineers*. Cambridge University Press, 2000.

[5] Andrew M. Gleason. "Measures on the Closed Subspaces of a Hilbert Space". In: *The Logico-Algebraic Approach to Quantum Mechanics*. Vol. 23. 12. Dordrecht: Springer Netherlands, 1975, pp. 123–133. DOI: 10.1007/978-94-010-1795-4_7.

[6] Brian C Hall. *Lie Groups, Lie Algebras, and Representations*. Vol. 222. Graduate Texts in Mathematics C. Cham: Springer International Publishing, 2015, pp. 3–30. DOI: 10.1007/978-3-319-13467-3.

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

[8] Lucien Hardy. "Why Quantum Theory?" In: *Non-locality and Modality*. Dordrecht: Springer Netherlands, Nov. 2002, pp. 61–73. DOI: 10.1007/978-94-010-0385-8_4.

[9] Chi-Kwong Li and Wasin So. "Isometries of ℓp-norm". In: *The American Mathematical Monthly* 101.5 (May 1994), pp. 452–453. DOI: 10.1080/00029890.1994.11996972.

[10] Lluís Masanes, Thomas D. Galley, and Markus P Müller. "The measurement postulates of quantum mechanics are operationally redundant". In: *Nature Communications* 10.1 (Dec. 2019), p. 1361. DOI: 10.1038/s41467-019-09348-x.

[11] Lluís Masanes and Markus P. Müller. "A derivation of quantum theory from physical requirements". In: *New Journal of Physics* 13.6 (June 2011), p. 063001. DOI: 10.1088/1367-2630/13/6/063001.

[12] Asher Peres. "Proposed Test for Complex versus Quaternion Quantum Theory". In: *Physical Review Letters* 42.11 (Mar. 1979), pp. 683–686. DOI: 10.1103/PhysRevLett.42.683.

[13] Constantin Piron. "Axiomatique quantique". In: *Helvetica physica acta* 37.4-5 (1964), p. 439.

[14] Constantin Piron. "On the Foundations of Quantum Physics". In: *Quantum Mechanics, Determinism, Causality, and Particles*. Ed. by M. Flato et al. Dordrecht: Springer Netherlands, 1976, pp. 105–116. DOI: 10.1007/978-94-010-1440-3_7.

[15] 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.

[16] Maria P. Solèr. "Characterization of hilbert spaces by orthomodular spaces". In: *Communications in Algebra* 23.1 (Jan. 1995), pp. 219–243. doi: 10.1080/00927879508825218.

[17] Allen Stairs. "Quantum Logic and Quantum Reconstruction". In: *Foundations of Physics* 45.10 (Oct. 2015), pp. 1351–1361. doi: 10.1007/s10701-015-9879-4.