Interpretations

(of Quantum States)

Bharath Ron

What's the most resilient parasite?
A bacteria? A virus? An intestinal worm??

An Idea!

Resilient, highly contagious.
Once an idea has taken hold of the brain,
it's almost impossible to eradicate.

Cobb
Inception

The complications in the interpretation of quantum theory come from the interpretation of probabilities occurring in quantum theory. We follow Adan Cabello's classification of the various interpretations of quantum theory [18]. Interpretations of quantum theory can be broadly classified according to whether they view probabilities of measurement outcomes as determined or not by intrinsic properties of the observed system.

- Interpretations
  - Intrinsic Realism
    - ψ-ontic
      - Many worlds, etc.
    - ψ-epistemic
      - Einstein(?), Ballentine [3]
      - Spekkens [16]
      - Superdeterminism [?]
  - Participatory Realism
    - ψ-doxastic
      - QBism [5],[6],[26]
      - ConSol [15]
      - Relational [11]
    - Knowledge
      - Neo-Copenhagen [23]
      - Wheeler [9],[10]
      - Fuchs & Peres [4]

Depending on the choice of interpretation of quantum states, the physical meaning of experiments is interpreted differently. So, questions like what is quantum theory about, what is the meaning of quantum entanglement, what does teleportation means, etc. depend on the choice of interpretation of quantum states. The participatory realist interpretations are also called Copenhagenish interpretations, due to their similarity with the original Copenhagen interpretation of quantum theory, where both unitary evolution and measurement collapse are treated on an equal footing.

1 | Structure of Physical Theories

Before comparing different interpretations of quantum theory we will quickly describe, without too much mathematical precision, the general structure of physical theories. Physics is an epistemological subject. The quality of other scientific theories are usually evaluated based on whether they can be reduced from the more primitive ideas of physical theories or not. Even though reductionist approach should be the preferred approach, there is no hope of such reduction when it comes to physics. Physics is the most foundational subject, and it has to be epistemological. In this sense physics is all about constructing theories that describe and predict observed phenomena. The first step is to observe nature and then try to speculate ideas about the 'underlying' structure.

In order to avoid getting lost in formality we will only focus on the aspects that are relevant to us, which is the process of going from physical ideas to physical theories. For a more formal treatment regarding the structure of physical theories see [**?**]. The next section is regarding the structure and heirarchy of physical theories and can be skipped.

We will call the collection of all observed phenomena physical domain denoted by $\mathcal{P}$. In order to be able to describe observed phenomena a physical theory should first allow a way to describe the observed phenomena. The collection of all phenomena a physical theory $\mathcal{PT}_\alpha$ considers will be called a fundamental domain of the physical theory denoted by $\mathcal{D}_\alpha$. $\mathcal{D}_\alpha$ consists of sentences in an appropriate mathematical language corresponding to observed phenomena. Any language that can be used to communicate the observed phenomena will be such a theory. This is however not sufficient, we expect physical theories to be predictive.

We will call the collection of all observed phenomena that can be predicted by the physical theory the application domain of the physical theory denoted by $\mathcal{P}_\alpha$. To get predictive theories one has to put additional constraints on what the theory can describe such that it is possible to axiomatise the observed patterns within the framework of the theory. For example, one cannot impose linearity condition if the sets we are working with are not vector spaces. These additional constraints give rise to special mathematical theories. We denote the mathematical theory associated with the physical theory $\mathcal{PT}_\alpha$ by $\mathcal{MT}_\alpha$.

Ideally we want a theory that describes all the observed phenomena and also predicts all the observed phenomena. That is to say,

$$\mathcal{P}_\alpha \equiv \mathcal{D}_\alpha \equiv \mathcal{P}.$$

In order to be able to predict, the phenomena should first be describable by the theory. Hence for a general physical theory we have,

$$\mathcal{P}_\alpha \subseteq \mathcal{D}_\alpha$$

We can now have that,

$$\mathcal{P} \subset \mathcal{P}_\alpha \subseteq \mathcal{D}_\alpha,$$

or

$$\mathcal{P}_\alpha \subseteq \mathcal{D}_\alpha \subset \mathcal{P}$$

In the first case the theory $\mathcal{PT}_\alpha$ predicts more than what is observed. An example would be a theory that predict the existence of some particles that do not actually exist in nature. The second case fails to describe and hence even consider some of the observed phenomenon. For example, classical mechanics fails to consider observed quantum phenomena. Both of these cases are suboptimal physical theories. To understand where the problem lies in each of the cases we need to look at how we construct the physical theories.

1.1 | Heirarchy of Ideas

Every physical theory $\mathcal{PT}_\alpha$ starts off with a collection of physical ideas $\mathcal{I}_\alpha$ speculating the structure of the physical world, a mathematical theory $\mathcal{MT}_\alpha$ which is used to describe the observed phenomenon and also articulates the physical ideas into a mathematical theory. If $\mathcal{I}$ denote the collection of all physical ideas physics can be thought of as the association,

$$\mathcal{I}_\alpha \xleftrightarrow{\mathcal{PT}_\alpha} \mathcal{MT}_\alpha.$$

where $\mathcal{I}_\alpha \subset \mathcal{I}$.

To get the optimal physical theory we need to have an optimal number of physical ideas for ideal predictability, and hence the theory can describe and predict all the observed phenomena. Such an ideal physical theory will be called a universal theory or a theory of everything. We can use $\mathcal{D}_\alpha \equiv \mathcal{P}_\alpha \equiv \mathcal{P}$ as a criterion for ideal physical theory to describe a heirarchy on the collection of subsets of $\mathcal{I}$. In this article we are concerned with four main points of failure of physical theories.

Case I: Bad Predictions. In this case the theory offers predictions for unobserved or sometimes even unobservable phenomena. The first subcase is common to physical theories. The failure of a physical theory when the predictions are unobserved are usually due to having wrong speculation of the structure. This can be due to lack of capabilities to observe when the ideas of the theory are speculated. As technology improves and the capabilities to observe phenomena improves more phenomena can be observed and the theory at hand might not have considered such observations while speculating the structure of the world. In this sense all physical theories are to be expected to predict things that would not be observed. This failure is usually taken as a reason to figure out the problem in the physical theory.

The other more serious failure is when predictions cannot be observed at all. This situation is the worst case scenario. An example of such a theory could be a theory which predicts the existence of extra particles which do not interact with anything in the observable world, such a particle cannot be observed because they donot interact with anything. Another such example is the prediction of the existence of a supernatural god or a unicorn. We will refer to [?],[?] for more on the problems with such theories.

Case II: Bad Explanations. In this case the theory can explain the observed phenomena but cannot predict the phenomena. In this case it makes no difference if the ideas behind such explanations were included in the physical theory or not. We lose no predictability if we left out these ideas from the physical theory. These additional ancillary ideas donot offer any predictability and hence serve no purpose, if we view physics as consisting of theories to predict phenomena in the world. These ideas do not belong to $\mathcal{I}$. An example of such a theory is using god to explain observed phenomena. In this case, we can explain all observed phenomena as an act of god, but such a theory cannot predict observed phenomena. These theories shift the goal post and use some other unobservable phenomena to explain observed phenomena. We feel the answers these theories offer are unacceptable scientifically.

Case III: Unexplained Phenomena. In this case, the theory cannot describe observed phenomena. Many physical theories of this type. Each physical theory can explain and predict a collection of observed phenomena, but there does not yet exist a theory of everything which can both explain and predict everything. For example, the mathematical theory of classical mechanics cannot be used to describe continuous and discrete observables if they coexist. The co-existence of discrete and continuous observables is an observed phenomena, but it cannot be explained by classical theory. In this case we have too many physical ideas for the theory constraining its mathematical language too much. Such a physical theory can be fixed by identifying and removing the unnecessary ideas.

Case IV: Unpredictable Phenomena. In this case the theory can describe the observed phenomena, but cannot predict the phenomena. In this case the physical theory has too less physical ideas. Such a physical theory can be fixed by speculating correctly the underlying structure, and articulating this idea within the mathematical theory of the physical theory. These additional ideas add extra constraint to the theory.

2 | Intrinsic Realism

A commonsense belief is that at any given moment any physical quantity must have a value even if we do not know what it is. The intrinsic realist interpretations are those in which probabilities of measurement outcomes are assumed to be determined by the intrinsic properties of the observed system. The notion of reality is closely linked with the values of observables in these interpretations.

- Intrinsic Realism
  - ψ-ontic
    - Many worlds, etc.
  - ψ-epistemic
    - Einstein?, Ballentine [3]
    - Spekkens [16]
    - Superdeterminism? [?]

An intrinsic realist interpretations is called ψ-ontic if quantum state are viewed as an intrinsic property of the system. The interpretation is called ψ-epistemic if they view the quantum state as having information about the intrinsic properties of the system, much like probabilities in classical statistical mechanics. Here we will discuss complications in relation to ψ-epistemic interpretations. ψ-ontic interpretations such as many worlds interpretation will be only briefly discussed.

This intrinsic realist view comes from our belief in classical physics. This is not problematic in classical physics since the underlying structure for observables and states fits this view perfectly. However such a view comes into lots of problems with the quantum formalism.

2.1 | ψ-ontic Interpretations

ψ-ontic interpretations continue to interpret probabilities in quantum theory in the same way as in classical theory. The most frequently discussed ψ-ontic interpretation is the so called many world interpretation due to Everett [?], and popularised by DeWitt [?] in the 1970s, and more recently by popular science personalities. Unlike the Copenhagenish interpretations measurement collapse is not treated on an equal footing as unitary evolution.

The approach takes the unitary evolution seriously and assumes the problem to be coming from the measurement part. The interpretation assumes the existence of a state for the universe which evolves unitarily. In their approach, observer and the universe are treated separately. Interpreting the world as being composed of its components, the quantum state of the world is the product state of its component objects. $|\psi_W\rangle = |\psi_{O_1}\rangle \otimes |\psi_{O_2}\rangle \otimes \cdots$ here O_j are the component objects. The quantum state of the universe is described as the superposition of infinitely many and increasingly divergent quantum worlds.

$$|\psi_U\rangle = \sum_i \alpha_i |\psi_{W_i}\rangle.$$

Different worlds W_i correspond to different classically described states (or values of observables) of at least one object. All states ψ_{W_i} are mutually orthogonal and it is assumed that $\sum_i |\alpha_i|^2 = 1$.

Measurements are treated as a correlation inducing interaction between observer and the universe. This interaction changes the observer's state. The measurement result corresponds to the observer's relative component of the state of the universe. Each measurement splits the world into different worlds, corresponding to different values of the measurement outcomes. The universe is then a superposition of infinitely many and increasingly divergent superposition of quantum worlds. For more details, see Vaidman's review [**?**].

2.1.1 | Sacrificing Predictability

There are many problems with the many world interpretation, some of which are already discussed in [**?**]. For us the many world interpretation offers explanation for everything, and does not offer any predictibility. This is case II in 1 (arguably) bad explanation for observed phenomena.

2.2 | ψ-epistemic Interpretations

In statistical mechanics the states of the system are probability distribution over microstates. The microstates correspond to the dirac delta distribution over the phase space. These are similar to how states are in classical mechanics. The microstates represent the reality, and are the ontic states of the theory. These microstates are variables whose values we donot know accurately for whatever reason. This can be states as saying, 'In statisical mechanics we donot exactly know the microstate' i.e., we donot have complete knowledge about the variables. The macrostates represent this known information about the microstates. This is the idea behind hidden variables, that there exist some ontic 'real' state and the quantum states represent only the knowledge we know about them. For example, a coin flip has a hidden variable determining the probability of outcomes. If we know all the parameters such as force, torque, etc from the finger exerted on the coin and complete information about the system we can in theory know what the result of the experiment will be exactly. These parameters together represent the hidden variable for the coin toss experiment.

The natural question to ask now is if it is possible to find a hidden variable for the quantum case, that is to ask if there exists a hidden variable knowing values of which will determine the values of the measurement. As we will see, the problem comes from the non-commutativity of quantum theory.

Kochen-Specker Hidden Variable Model

One of the very well understood systems are spin 1/2 systems that we can perform Stern-Gerlach experiments on. In these experiments we can measure the 'spin' of a particle in some direction. To every direction in space we can measure if the particle is 'up' in that direction. The idea is that to every pure state of this particle there corresponds a unique direction in space such that the spin is 'up' in that direction. Therefore its space of pure states is isomorphic to the set of all directions in space. This space of pure states is called Bloch shere, antipodal points on the Bloch sphere correspond to states that have spin up in opposite directions. Bloch ball, whose boundary is the Bloch sphere corresponds to the space of density matrices and can be parametrized as,

$$\rho = \begin{bmatrix} \frac{1}{2}+z & x-iy \\ x+iy & \frac{1}{2}-z \end{bmatrix}$$

This can be treated as an expansion in terms of Pauli matrices, $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, as,

$$\rho = \tfrac{1}{2}\mathbb{I} + \vec{\tau}\cdot\vec{\sigma}$$

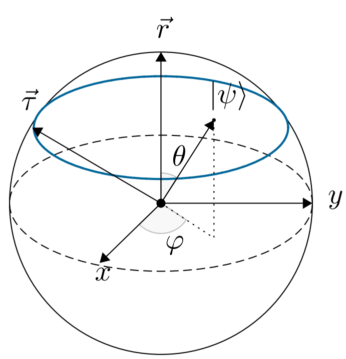
$\vec{\tau}$ is called the Bloch vector such that, $x^2+y^2+z^2 \leq \frac{1}{4}$. These systems are called qubits. A measurement correspond to projector projecting onto the eigenstate corresponding to the eigenvalue ± 1 of spin along some direction $\vec{r}$ is given by,

$$P_{\pm 1,\vec{r}} = \tfrac{1}{2}(\mathbb{I} + (\pm 1)\vec{r}\cdot\vec{\sigma}),$$

where $|\vec{r}| = 1$. The expectation value is given by,

$$Tr(\rho P_{\pm 1,\vec{r}}) = \vec{\tau}\cdot\vec{r}.$$

A qubit can be represented by the following diagram where the direction of measurement is represented by the z axis,



Kochen and Specker gave a hidden variable model for this system. The hidden variable state is given by the Bloch vector $\vec{\tau}$ of the quantum state and a unit vector $\vec{\lambda} = (\theta, \phi)$ uniformly distributed on the sphere given by,

$$\rho(\vec{\lambda})d\vec{\lambda} = \tfrac{1}{4\pi}\sin\theta d\theta d\phi$$

The outcome is deterministically computed as follows,

$$a(\lambda) = \text{sign}[(\vec{\tau} - \vec{\lambda})\cdot\vec{m}].$$

This term is either $+1$ or -1. To every $\vec{\lambda}$ in the cup above blue circle corresponds to a outcome -1. We have,

$$p(-1, \vec{r}) = \frac{1}{4\pi} \int_0^{\alpha} d\theta \sin\theta \int_0^{2\pi} d\psi,$$

where $\alpha = \arccos(\vec{r} \cdot \vec{\tau})$. The complement area will yield $+1$. The mean value is given by,

$$\langle s \rangle_{\vec{r}} = \int \rho(\vec{\lambda}) a(\vec{\lambda}) d\vec{\lambda} = \vec{r} \cdot \vec{\tau}$$

So the Kochen-Specker hidden variable model yields same result as quantum theory. Such a hidden variable model is not possible for higher dimensions however. This is shown by the Kochen-Specker theorem.

2.2.1 | Kochen-Specker Theorem

In classical physics, the states are specified as points in phase space which correspond to the values of position and momentum, the observables of the system. Such states specify the values of every observable that the system can take. The observer can know with certainty the values of all observables. The states in classical mechanics represent a state of reality. If Ω is the state space in classical physics, an observable is a map,

$$X : \Omega \to \mathbb{R}.$$

Each state would fix a value for the observables. If we denote the valuation map associated with a state ρ by λ_ρ then,

$$\lambda_\rho : X \mapsto \lambda_\rho(X).$$

All physical quantities possess a value in any state. If $h : \mathbb{R} \to \mathbb{R}$ is a real-valued measurable function we can construct new observables from old ones, the values of which are $h(X) := h \circ X : \Omega \to \mathbb{R}$. In such cases, we should expect the valuation of the new observable to be,

$$\lambda_\rho(h(X)) = h(\lambda_\rho(X)).$$

The observable $h(X)$ is defined by saying that its value in any state is the result of applying the function h to the value of X.

Effects in the quantum case are projection operators. If a valuation as above exists then the valuation map associated with a state of reality should assign to each projection operator the values 1 or 0 based on whether the system was measured with the said property or not. Such maps are called valuation maps or valuations. If λ is such a valuation map[1] then $\lambda(\mathbb{I}) = 1$. If A and B are self-adjoint operators such that for some real-valued function h, $B = h(A)$ then,

$$\lambda(B) = h(\lambda(A)).$$

Valuation maps represent non-contextual hidden variables i.e., the observables have pre-defined values and the values are independent of measurement context. A valuation map associated with a state ρ is a homomorphism from the algebra of projection operators to the set $\{0, 1\}$.

$$\lambda_\rho : \mathcal{P}(\mathcal{H}) \to \{0, 1\}.$$

[1] Not to be confused with the Lebesgue measure on $\mathbb{R}^n$.

Assuming such a valuation map exists, it must satisfy the valuation conditions $\lambda_\rho(\mathbb{I}) = 1$ and $\lambda_\rho(\sum_i E_i) = \sum_i \lambda_\rho(E_i)$. Such a map satisfies the conditions of Gleason's theorem, hence must take continuous values in $[0, 1]$. Since valuation maps can only take discrete values $\{0, 1\}$, such a map cannot exist.

Theorem 2.1. (Kochen-Specker) *If* $\dim(\mathcal{H}) \geq 3$ *then there exist no valuations.*

We used Gleason's theorem to quickly arrive at the Kochen-Specker theorem. Since the Gleason's theorem makes lots of people uncomfortable we also sketch below a more elementary 'pentagram' proof of the theorem.

Sketch of Proof

Consider a pentagram with each vertex representing an observable and any two vertices that have an edge between them commute. Denote the observables by A_i, $i \in 1, \cdots, 5$. If there exists a non-contextual hidden variable λ that describes the system then to each of the five observables we assign definite values that's independent of measurement context. So,

$$A_i(\lambda) = \pm 1, \quad i = 1, \ldots, 5.$$

Now some algebra shows us that,

$$-3 \leq A_1(\lambda)A_3(\lambda) + A_3(\lambda)A_5(\lambda) + A_5(\lambda)A_2(\lambda) + A_2(\lambda)A_1(\lambda) \leq 5.$$

So the average should also lie in the same interval, $-3 \leq \langle A_1(\lambda)A_3(\lambda)\rangle + \langle A_3(\lambda)A_5(\lambda)\rangle + \langle A_5(\lambda)A_2(\lambda)\rangle + \langle A_2(\lambda)A_1(\lambda)\rangle \leq 5$. Now to arrive at a contradiction the idea is to choose a system and a set of 5 observables as above and show the expectation value of above expression lies outside the interval required by non-contextual hidden variables. See [**?**] for the full proof. $\square$

The theorem asserts that it is impossible to assign values to all physical observables while simultaneously preserving the functional relations between them. It should, however, be noted that when restricted to commutative subalgebras valuations do exist. Due to the non-commutativity of quantum theory, the values of all the observables canot be known at once and any such notion has to be contextual, value of the observable depends on the experimental context. The states in quantum theories cannot be interpreted completely ontically. Non-contextual hidden variable theories are also not viable. A 'state of reality' is meaningless in quantum theory. Any attempt to view the quantum state as ontic states would require serious mutilation of objects of quantum theory.

2.2.2 | Bell's Theorem

The starting assumptions of quantum theory are quite general. It should be possible to model all observed phenomenon using quantum theory. Now the aim is write down a quantum theory that satisfy the constraints put forth by the theory of relativity of Einstein. The constraint of interest to us is causality. A basic characteristic of physics in the context of relativity is that causal influences on spacetime propagate in timelike or lightlike directions but not spacelike. Communication should only possible at a speed less than that of light according to relativity.

Now the first step to formalize this statement is to formalize the concept of local realism. We start with Bell's definition of locality also called called local causality or Bell locality.

We want to define local realism for two experimenters say, Alice and Bob, in space-like separated regions performing two experiments A and B respectively. Let the values of the observables by A_i and B_j respectively. Bell assumed there existed some hidden variable λ and the probability of observables are distribution over the hidden variable. The probability that the combined system in the hidden variable λ has values A_i and B_j for the obseravbles A and B is given by,

$$\mu(A_i, B_j|\lambda) = \mu(A_i|\lambda)\mu(B_j|\lambda)$$

The product structure comes from the assumption that space-like separated events canot influence each other.

$$\mu(A_i, B_j) = \int_\lambda \mu(A_i, B_j|\lambda)\rho(\lambda)d\lambda = \int_\lambda \mu(A_i|\lambda)\mu(B_j|\lambda)\rho(\lambda)d\lambda,$$

where $\rho(\lambda)$ tells us how the probabilities are distributed over λ. We will call this condition Bell locality. The normalization implies that, $\sum_{A_i,B_j} \mu(A_i, B_j) = 1$. Note that we are already assuming non contextual hidden variables determining the values of observables. We will denote the set of all probabilistic vectors which obey the Bell locality $\mathcal{L}$, call this set local set. It is easy to verify that this is a convex set.

Now for the case of quantum theory the expectations are given by,

$$\mu_\rho(A_i, B_j) = Tr(\rho(P_{A_i} \otimes P_{B_j})),$$

where ρ is the state of the system and P_{A_i} and P_{B_j} are POVMs corresponding to the events associated with values A_i and B_j of observables A and B respectively with $\sum_i P_{A_i} = \mathbb{I}_A$ and $\sum_j P_{B_j} = \mathbb{I}_B$. Denote by,

$$\mathcal{Q} = \{\mu_\rho(A_i, B_j) \mid \rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)\}$$

the set of all probability vectors. Since $Tr(\cdot(P_{A_i} \otimes P_{B_j}))$ is a continuous linear map, it maps convex bounded set to convex bounded set. Now basic topology tells us every continuous function $f : C \to \mathbb{R}$ where C is convex, compact will attain maxima/minima at extreme points of C. This is due to the fact that compact sets get mapped to compact sets, since C is a convex set $f(C)$ is a connected interval, interior of C is mapped to interior of $f(C)$. Our aim is to show that there exist states that do not satisfy Bell locality, or equivalently, the local set is a strict subset of the quantum set.

Theorem 2.2. (Bell's Theorem)

$$\mathcal{L} \subsetneq \mathcal{Q}$$

Here $\mathcal{Q}$ is the collection of all quantum states, and $\mathcal{L}$ is the states which satisfy Bell locality condition. Proofs of Bell's theorem usually involve constructing an entangled state that will not satisfy the Bell locality condition. These proofs donot reveal anything about what mathematical structure of quantum theory is causing the problem. One strategy is to show that states satisfying Bell locality will have certain bounds and then construct a quantum state that disobeys such a bound. We will sketch below one such proof that uses the so called CHSH inequality.

Sketch of Proof

The starting point is the expectation value of the product of the outcomes of the experiment. Assume the observables A and B take two values say ± 1. The expectation of the product of outcomes is given by,

$$E(a,b) = \int A(a,\lambda)B(b,\lambda)\rho(\lambda)d\lambda$$

where $A(a,\lambda)$ is the value of the observable A given the system is in the hidden variable λ and measurement setting a for the instrument and similarly for $B(b,\lambda)$. This is where we have imposed the Bell locality condition. Since possible values are ± 1 we get that $|A| \leq 1$ and $|B| \leq 1$. By considering intruments with two settings a_1, a_{-1} and b_1, b_{-1} respectively, we get the following constraint

$$|S| = |E(a_1, b_1) - E(a_1, b_{-1}) + E(a_{-1}, b_1) + E(a_{-1}, b_{-1})| \leq 2.$$

for all probability vectors in $\mathcal{L}$. The wikipedia article on CHSH inequality has sufficient details and an interested reader should read it.

To arrive at a contradiction one considers a two qubit system. One starts with measurement basis, $|0\rangle_A, |1\rangle_A$ and $|0\rangle_B, |1\rangle_B$ and then considers Bell entangled states. The expectations will turn out to be $E(a_i, b_j) = \pm 1/\sqrt{2}$ and $|S| = 2\sqrt{2}$ hence violating the CHSH bound and proving the theorem. □

For any self-adjoint operator A, the norm of A is the same as the spectral radius, $\|A\| = \sup_{\lambda\in\sigma(A)}\{|\lambda|\}$. So, for 'binary' observables A_i, B_j with eigenvalues ± 1, consider the operator,

$$S = A_1 \otimes B_1 + A_1 \otimes B_2 + A_2 \otimes B_1 - A_2 \otimes B_2$$

Each $A_i = P_{+1_{A_i}} - P_{-1_{A_i}}$, where $P_{\pm 1_{A_i}}$ is the projection corresponding to the value ± 1. We get, $A_i^2 = \mathbb{I}$, and similarly for B_j. Calculating S^2, we get, $S^2 = 4(\mathbb{I}\otimes\mathbb{I}) + [A_1, A_2] \otimes [B_2, B_1]$. This yields,

$$\|S^2\| \leq 8$$

or equivalently,

$$\|S\| \leq 2\sqrt{2}.$$

This is called the Tsirelson bound, and Bell states attain the maximal value.

A probability distribution is said to satisfy no-signalling condition if,

$$\mu(A_i|\lambda_k) = \sum_{B_j} \mu(A_i, B_j \mid \lambda_k, \lambda_l)$$

and similarly for B_j. Distributions which satisfy this no-signalling condition are denoted by $\mathcal{NS}$. The quantum set satisfies the no-signalling condition.

Theorem 2.3.

$$\mathcal{L} \subsetneq \mathcal{Q} \subsetneq \mathcal{NS}.$$

The proof is a constructive proof like the Bell's theorem, the example construction is called the Popescu-Rorhlich box or PR box which we will describe in the sketch below.

Sketch of Proof

Consider the distribution, $\mu_{PR}(A_i, B_j|\lambda_k, \lambda_l) = \frac{1}{2}$, $\mu_{PR}(A_i, B_i|\lambda_k, \lambda_l) = 1$, $\mu_{PR}(A_i, B_{j\neq i}|1, 1) = 1$ where $i, j = \pm 1$, and $\lambda_{k,l} = \{0, 1\}$. This satisfies the no-signalling condition, and doesnot belong to $\mathcal{Q}$. Suppose $\mu \in \mathcal{Q}$, then we have,

$$\mu(A_i, B_j|\lambda_k, \lambda_l) = Tr(\rho(P^k_{A_i} \otimes P^l_{B_j}))$$

Then the probabilities will be, $\sum_i \mu(A_i, B_j|\lambda_k, \lambda_l) = \sum_i Tr(\rho(P^k_{A_i} \otimes P^l_{B_j})) = Tr(\rho(\mathbb{I} \otimes P^l_{B_j}))$. So, the probability doesnot depend on the choice of the setting λ_k. So, $\mathcal{Q}$ is indeed contained in $\mathcal{NS}$. □

On an operational level, the hidden variable models cannot be distinguished from quantum theory. The question then is about the plausibility of such models and if they are useful as physical theories in terms of predictability, etc. If a model requires an infinite number of hidden variables to describe stuff, it is not a very good model or as Hardy calls it, such models carry "ontological excess baggage".

Let Ω be the set of all hidden variables, let $\mu(\lambda|\varphi)$ denote the probability distribution of the hidden variable corresponding a given state φ. Then we should have,

$$\sum_{\lambda \in \Omega} \mu(\lambda|\varphi) = 1$$

For normalized states we have, $|\langle\varphi|\varphi\rangle|^2 = 1$, so we have, $\sum_{\lambda\in\Omega} \mu(\lambda|\varphi)\mu(\varphi|\lambda) = |\langle\varphi|\varphi\rangle|^2 = 1$. This can happen only if each $\mu(\varphi|\lambda) = 1$. Denote by Ω_φ all $\lambda \in \Omega$ for which $\mu(\lambda|\varphi) > 0$.

Theorem 2.4. (Hardy's Theorem) *Any hidden variable theory that reproduces all measurements of a quantum system must have an infinite number of hidden variable states.*

Proof

Consider a two level system, the idea is that if Ω_φ is a complete set of hidden variables then for all vectors $\varkappa$ we should have $\sum_{\lambda\in\Omega} \mu(\lambda|\varkappa) = 1$. Now consider the set of M states (not orthogonal) given by,

$$|\varphi_i\rangle = \cos\left(\frac{\pi i}{2M}\right)|0\rangle + \sin\left(\frac{\pi i}{2M}\right)|1\rangle.$$

For this set we have, $|\langle\varphi_i|\varphi_j\rangle| < 1$ which implies $\sum_{\lambda\in\Omega} \mu(\lambda|\varphi_i)\mu(\varphi_j|\lambda) = |\langle\varphi_i|\varphi_j\rangle|^2 < 1$ which can happen only if some $\mu(\varphi_i|\lambda) < 1$. So Ω_{φ_i} must have different elements not already in Ω_φ. This means there are atleast M distinct subsets of Ω. Now since M can be arbitrarily large we conclude that Ω must be infinite. □

2.2.3 | PBR Theorem

A non-contextual ontological model of an operational theory is an attempt to provide a causal explanation of the operational statistics. It says that the response of the measurement is determined by the ontic state λ of the system, while preparation procedures determine the distribution over the space of ontic states, Ω, from which λ is sampled. An ontological model associates to each preparation ρ a probability distribution μ_ρ representing the agents' knowledge of the ontic state given the preparation ρ. If we denote the set of such distributions by $\mathcal{M}(\Omega)$, the ontological model specifies a map,

$$\mu : \mathcal{S}(\mathcal{H}) \to \mathcal{M}(\Omega).$$

An ontological model associates to each operational effect a response function on Ω representing the probability assigned to the outcome R_i in a measurement of R if the ontic state of the system fed into the measurement device were known to be $\lambda \in \Omega$. If we denote the set of response functions by $\mathcal{F}(\Omega)$, the ontological model specifies a map,

$$\eta : \mathcal{P}(\mathcal{H}) \to \mathcal{F}(\Omega).$$

These two maps must preserve the convex structure i.e, if ρ is a mixture of ρ_1 and ρ_2 with weights λ and $1-\lambda$ then $\mu_\rho = \lambda\mu_{\rho_1} + (1-\lambda)\mu_{\rho_2}$ and similarly for effects. Furthermore, an ontological model should produce the same probability rule as the operational theory. Assuming Ω is discrete for simplicity, we have,

$$\mu(\rho, R_i) = \sum_{\lambda\in\Omega} \eta_{R_i}(\lambda)\mu_\rho(\lambda).$$

An ontological model of an operational theory is said to satisfy the generalized noncontextuality if every two operationally equivalent procedures have identical representations in the ontological model. That is to say, $\rho \sim \rho' \Rightarrow \mu_\rho = \mu_{\rho'}$ and similarly for effects. The same can be defined for GPTs and it can be showed (see, [?]) that they are equivalent. We will now go back to the quantum case, and the PBR theorem [24].

Suppose quantum state ρ is a state of knowledge, representing the uncertainty about the real underlying ontic state of the system λ. The quantum state ρ results in a physical state λ with a probability distribution $\mu_\rho(\lambda)$. If the distributions for distinct quantum states do not overlap then the quantum state can be uniquely inferred from the physical state. If the distributions overlap, then the quantum states can be said to only contain some knowledge about the physical state. Suppose we have two quantum states ρ_1 and ρ_2 with overlapping distributions, $\mu_{\rho_1}(\lambda)$ and $\mu_{\rho_2}(\lambda)$ then for any λ in the overlap Δ, there is a $q > 0$ probability that the physical state is compatible with both quantum states. Now consider two uncorrelated systems that are prepared with two copies of the same preparation device. If the physical states λ_1 and λ_2 lie in the overlap Δ, there must be some $q > 0$ such that with q^2 probability the quantum states $\rho_1 \otimes \rho_1$, $\rho_1 \otimes \rho_2$, $\rho_2 \otimes \rho_1$ and $\rho_2 \otimes \rho_2$ will be in this physical state. The measurement on this system can be cleverly chosen such that the first outcome is orthogonal to the first state, the second outcome orthogonal to the second state, and so on.

To arrive at a contradiction consider $\rho_1 = |0\rangle$ and $\rho_2 = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, and choose the measurement which projects onto the following orthogonal vectors,

$$\begin{aligned}
&(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle)/\sqrt{2} \\
&(|0\rangle \otimes |-\rangle + |-\rangle \otimes |0\rangle)/\sqrt{2} \\
&(|+\rangle \otimes |1\rangle + |1\rangle \otimes |+\rangle)/\sqrt{2} \\
&(|+\rangle \otimes |-\rangle + |-\rangle \otimes |+\rangle)/\sqrt{2},
\end{aligned}$$

where $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. The measuring device should have been uncertain at least q^2 of the time about which quantum state was used in the measurement. We have proved the following theorem,

Theorem 2.5. (Pusey-Barret-Rudolph) *Quantum state interpreted as information about an objective physical state cannot reproduce the predictions of quantum theory.*

This will imply a measurement will give an outcome that is predicted to not happen quantum mechanically. Hence, interpreting quantum states as having information about an underlying objective physical state cannot reproduce the predictions of quantum theory. Kochen-Specker theorem rejects non-contextual hidden variable theories and the PBR theorem rejects the existence of a non-contextual ontological model or the statistical interpretation. Frequentist interpretation of probabilities occurring in quantum theory is problematic. Kochen-Spekker theorem can be evaded by arguments of the sort ‘the values are never accurately known’. If the idealisation procedure of quantum mechanics is to be respected then we cannot go down this path.

The interpretation is ψ-epistemic if they view the quantum state as containing knowledge about an underlying reality similar to how we view states in classical statistical mechanics. The point of view given in [3] is the statistical interpretation, the quantum states represent partial knowledge about an underlying state of reality. In classical statistical mechanics, probability distributions are introduced on the phase space. These distributions represent the likelihood of the occurrence of the values. However, if the position and momenta of all the particles are known then we have complete knowledge of the system. These states of complete knowledge of the system correspond to delta distributions which are in a one-to-one correspondence with the points in phase space. The PBR theorem says that this is not possible in the quantum case, and hence the theorem can be thought of as a contradiction to interpreting quantum states statistically.

Superdeterminism

Superdeterminism is a loophole in Bell’s theorem, more generally in Gleason’s theorem itself. The superdeterminists postulate that all systems being measured are correlated with the choices of which measurements to make on them, the assumptions of Gleason’s theorem are no longer fulfilled, and Kochen-Specker theorem and Bell’s theorem need not be true.

For us this lies in Case II theory. Why are the choices of which measurements to make on them correlated? The proponents of superdeterminism, argue that this is perfectly valid scientific theory, and question their opponents as being opposed to determinism itself. We feel trying to answer why the choices are correlated is similar to speculating the nature of a supernatural god after claiming an observed phenomena is due to a supernatural god.

3 | Participatory Realism

If a tree falls in a forest and no one is around to hear it, does it make a sound?

Participatory realist interpretations treat the measurement collapse as fundamental. They do not make comments about the unitary evolution. The problem with these interpretations depend on how they interpret measurement collapse. The most of refined of these interpretaions are QBism and relational quantum mechanics. Both of which take an epistemic interpretation of quantum states. They view is bayesian, unlike the ψ-epistemic interpretations. The information contained in the state is viewed as beliefs of observers about the possible events and not as observer’s knowledge about intrinsic properties. This lets these interpretations avoid all the no-go theorems associated with intrinsic realist interpretations.

- Participatory Realism
  - ψ-doxastic
    - QBism [5][26][6]
    - Relational QM [?]
    - Conviniel Solipsism [15]
  - About Knowledge
    - Neo-Copenhagen [23]
    - Wheeler [9][10]
    - Fuchs & Peres [4]

Every generalised probabilistic theory comes equipped with the notion of information. To motivate what we mean by this, we give here some heuristis. Information is a very intuitive, yet abstract concept. We need to find a way to model it mathematically. Bayes' theorem acts as a good starting point.

To avoid getting lost in the formalism we will state Bayes' theorem in the most simple case. Let E and F be two comeasurable events, and $\mu(E)$ and $\mu(F)$ be their respective probability of occurance. If $E \wedge F$ denotes the event when both E and F happen, then the probability that E happens given F has happened, should be given by,

$$\mu(E \wedge F) = \mu(E|F)\mu(F).$$

Similarly, we have, $\mu(E \wedge F) = \mu(F|E)\mu(E)$. This gives us,

$$\mu(E|F) = \frac{\mu(F|E)\mu(E)}{\mu(F)}.$$

Bayes' theorem says that probability of occurance of an event changes once we have additional information about the system. For us, this means probability measures (possibly non-commutative) act as a good starting point to mathematically model the notion of information. In the participatory realist interpretations, the information is treated as 'beliefs' of the observer regarding the possible events. When we say information, we mean information about possible events. The next question one could ask is information for who? Whenever we talk about information, we intrinsically assume that the information is for the observer measuring the system. We can hence view the states of a GPT as a model of the information an observer has about the possible events. Since an observer is uniquely determined by what they can observe, the observer is uniquely determined by the collection of effects. We can view the space of effects as a model of the observer, and the space of states as the models of the observer's information. In this sense GPTs can be interpreted as mathematical models for observers and the observer's information about the *things* they can observe.

Note that this epistemic reinterpretation of GPTs is mathematically equivalent to the instrumentalist interpretation. At this stage it is important to clear up any confusions that might arise. Since we have identified the space of effects with the observers, this also provides a clear way of interpreting what is meant by an observer. There is no need for going to pre-theories seeking meaning for observers. In this sense, an observer can be interpreted from the instrumentalist formulation of GPTs.

$$\mathcal{E} \equiv \text{equivalence classes of measuring instruments} \equiv \text{observer}$$

The two interpretations, instrumentalist interpretation, and the epistemic interpretation as described here are mathematically equivalent. It offers no additional mathematical advantage. It can however provide us with intuition for thinking about GPTs. The reinterpretation can potentially provide us with different ways to think about and answer questions related to GPTs. One such thing is Bayesian inference.

3.1 | Bayesian Inference

When an observer gains new information about the possible events, they must update their original information. The participatory realist interpretations take this viewpoint and view the act of measurement as a Bayesian updating of the probabilities based on the newly available information. Here we give a characterisation of the measurement collapse from an information theoretic perspective. We need a way to compare the information the observer had prior to performing a measurement and the information post measurement. The starting point then is to quantify the amount of information contained in a state relative to another state.

3.1.1 | Relative Entropy & Inference

The relative entropy of two states ρ and σ is the informational divergence of ρ from σ. Suppose the state σ contains information only about a subsystem $\mathcal{B}$ of $\mathcal{A}$ and E is a projection of norm one of $\mathcal{A}$ onto $\mathcal{B}$ then the state σ should satisfy $\sigma \circ E = \sigma$. In such a case the informational divergence should have two components. First component is the divergence of ρ from σ on the subalgebra $\mathcal{B}$ which is the divergence between the states $\rho|_{\mathcal{B}}$ and $\sigma|_{\mathcal{B}}$. The other component is the remaining information ρ has and this will be the divergence between the states ρ and $\rho \circ E$. If $R(\cdot\,,\cdot)$ is such a function then we should have

$$R(\rho, \sigma) = R(\rho|_{\mathcal{B}}, \sigma|_{\mathcal{B}}) + R(\rho, \rho \circ E).$$

Any automorphism α of the algebra $\mathcal{A}$ should change the information contained in the two states similarly hence the information divergence should be invariant under automorphisms of the algebra,

$$R(\rho, \sigma) = R(\rho \circ \alpha, \sigma \circ \alpha).$$

The informational divergence of a state with respect to itself should be zero $R(\rho, \rho) = 0$. If $R(\cdot\,,\cdot)$ is a real-valued functional satisfying the above conditions then there exists a constant $c \in \mathbb{R}$ such that,

$$R(\rho, \sigma) = c\,\mathrm{Tr}\left(\rho\left(\log\rho - \log\sigma\right)\right).$$

See [?] for a proof of this uniqueness. The relative entropy of the state ρ with respect to σ is defined as,

$$J(\rho, \sigma) = \mathrm{Tr}\left(\rho\left(\log\rho - \log\sigma\right)\right).$$

In the classical case, the Bayes' rule has been shown to be a special case of the constrained maximization of relative entropy [?]. The quantum version of this result is obtained in [?]. We will state the result here.

Suppose an observable A has been subjected to measurement. For simplicity we consider the observable to be a discrete observable. Let A be a discrete observable with effects given by the set $\{A_i\}_{i\in I}$ and the corresponding projection operators $\{E_{A_i}\}_{i\in I}$. If the quantum state of the system after the measurement is σ, it carries information that has to be compatible with the possibility of measuring all eigenvalues of A precisely. Such a situation is given by the condition $[\sigma, A] = 0$. Suppose the result of the measurement is A_k then the probability of measuring A_k again should be $Tr(E_{A_k}\sigma) = 1$. Repeated measurements add no new information. The set of all such states such that $Tr(E_{A_k}\sigma) = 1$ is a convex set. Let $p = \{p_i\}_{i\in I}$ such that $\sum_i p_i = 1$. The set,

$$\mathcal{S}_p = \{\sigma \in \mathcal{S}(\mathcal{H}) \mid [E_{A_i}, \sigma] = 0,\ Tr(\sigma E_{A_i}) = p_i\},$$

encodes the data that the measurement outcome A_i corresponding to the projection E_{A_i} occurs with probability p_i. The commutation condition says that they posses a common eigenbasis and also means that $[\sigma, A] = 0$.

Theorem 3.1. (Hellmann-Kamiński-Kostecki)

$$\arg\inf_{\sigma\in\mathcal{S}_p}\{J(\rho,\sigma)\} = \sum_i p_i E_{A_i}\rho E_{A_i}/Tr(E_{A_i}\rho E_{A_i}).$$

The strong collapse or the Lüders-von Neumann rule of collapse is a limiting case of the above projection with all p_i going to zero except one. By taking the limit $p_i \to 0$ for $i \neq j$ we get the Lüders-von Neumann's rule of collapse,

$$\rho \to E_{A_j}\rho E_{A_j}/Tr(E_{A_j}\rho E_{A_j}).$$

This amounts to selecting the quantum state that is least distinguishable from the original state among all the states that satisfy the constraint.

Idea of Proof

Given a convex subset $\mathcal{V}$ of a finite dimensional topological vector space and $f : \mathcal{V} \to \mathbb{R}$ is a convex function then σ is a global minimum of the function f on $\mathcal{V}$ if and only if all directional derivatives of f at σ are non negative.

In our case, $D(\cdot,\cdot) = -J(\cdot,\cdot)$ is a jointly convex function. $D(\rho,\cdot) = -J(\rho,\cdot)$ is a convex function on the state space. Now the problem is a minimization of a convex function. $\mathcal{V} = \mathcal{S}_p \subset \mathcal{S}(\mathcal{H})$. Every element of $\mathcal{S}_p$ can be written as follows,

$$\sigma = U\Lambda U^*,$$

where Λ is a diagonal matrix with positive entries and trace 1 and U is a unitary. Since $[\sigma, P_i] = 0$ for every $\sigma \in \mathcal{S}_p$. Now the idea is to parametrise this and optimise it.

For complete proof and generalization of the result to the algebraic case, see [?],[?] and the references therein. It should however be noted that with this interpretation, the Lüders-von Neumann rule is about calibrating the information with the experimental result and has no predictivity.

Postulate. (Lüders-von Neumann Collapse) *If an observable A, is observed to have the value A_i the state transformation is given by,*

$$\rho \mapsto E_{A_i}\rho E_{A_i}/Tr(E_{A_i}\rho E_{A_i}),$$

where E_{A_i} is the projection corresponding to the value A_i.

The advantage of this approach to arriving at the Lüders-von Neumann rule is that the starting point is information theoretic and can be formulated whenever the GPT has suitable mathematical structure.

3.2 | ψ-Doxastic Interpretations

3.2.1 | QBism & Relational Interpretation

3.3 | Necessary, not Sufficient

- Interpretations
  - Intrinsic Realism
    - ψ-ontic
      - Many worlds, etc.
    - ψ-epistemic
      - Ballentine [3]
      - Spekkens [16]
      - Sabine Hossenfelder?
  - Participatory Realism
    - About Belief (ψ-doxastic)
      - QBism [5],[6],[26]
      - ConSol [15]
    - About Knowledge
      - Neo-Copenhagen [23]
      - Wheeler [9],[10]
      - Relational [11]
      - Fuchs & Peres [4]

Depending on the choice of interpretation of quantum states, the physical meaning of experiments is interpreted differently. So, questions like what is quantum theory about, what is the meaning of quantum entanglement, what does teleportation means, etc. depend on the choice of interpretation of quantum states.

References

[1] H Halvorson, To Be a Realist about Quantum Theory. 2018

[2] R P Kostecki, Quantum theory as inductive inference. http://arxiv.org/abs/1009.2423v4, 2011

[3] L Ballentine, Statistical Interpretation of Quantum Mechanics. Rev. Mod. Phys. 42, 358, 1970

[4] C A Fuchs, A Peres, Quantum Theory Needs No 'Interpretation'. Physics Today 53, 3, 70, 2000

[5] C A Fuchs, Quantum Mechanics as Quantum Information (and only a little more). https://arxiv.org/abs/quant-ph/0205039, 2002

[6] C A Fuchs, N D Mermin, R Schack, An Introduction to QBism with an Application to the Locality of Quantum Mechanics. https://arxiv.org/abs/1311.5253, 2013

[7] A Barzegar, QBism Is Not So Simply Dismissed. Foundations of Physics, https://doi.org/10.1007/s10701-020-00347-3, 2020

[8] R Clifton, J Bub, H Halvorson, Characterizing Quantum Theory in terms of Information-Theoretic Constraints. https://arxiv.org/abs/quant-ph/0211089, 2002

[9] J A Wheeler, Information, physics, quantum: the search for links, Proceedings of the 3rd International Symposiumon Quantum Mechanics, 354-368, Tokyo, 1989

[10] J A Wheeler, "On recognizing 'law without law,' " Oersted Medal Response at the joint APS-AAPT Meeting, New York, 25 January 1983, American Journal of Physics 51, 398, https://aapt.scitation.org/doi/10.1119/1.13224, 1983

[11] C Rovelli, Relational Quantum Mechanics. https://arxiv.org/abs/quant-ph/9609002, 1996

[12] C Rovelli, Agency in Physics. https://arxiv.org/abs/2007.05300v2, 2020

[13] J Pienaar, QBism and Relational Quantum Mechanics compared. https://arxiv.org/abs/2108.13977v1, 2021

[14] J Pienaar, Extending the Agent in QBism. https://arxiv.org/abs/2004.14847v1, 2020

[15] H Zwirn, Is the Past Determined?. Found Phys 51, 57. https://doi.org/10.1007/s10701-021-00459-4, 2021

[16] R Spekkens, In defense of the epistemic view of quantum states: a toy theory. Physical Review A 75, 032110 https://arxiv.org/abs/quant-ph/0401052, 2007

[17] C A Fuchs, On Participatory Realism. https://arxiv.org/abs/1601.04360, 2016

[18] A Cabello, Interpretations of Quantum Theory: A Map of Madness. http://arxiv.org/abs/1509.04711, 2015

[19] J Bell, Against Measurement. Physics World, vol. 3, p. 33, 2015

[20] A Peres, Unperformed experiments have no results. American Journal of Physics, 1978

[21] H Zwirn, The Measurement Problem: Decoherence and Convivial Solipsism. https://arxiv.org/abs/1505.05029, 2015

[22] E P Wigner, Remarks on the mind-body question. In Symmetries and Reflections. pages 171–184. Indiana University Press, 1967

[23] Ĉ Brukner, On the quantum measurement problem, https://arxiv.org/abs/1507.05255, 2015

[24] M F Pusey, An inconsistent friend. Nature Phys 14, 977–978. https://doi.org/10.1038/s41567-018-0293-7, 2018

[25] J Bub, In Defense of a "Single-World" Interpretation of Quantum Mechanics. http://arxiv.org/abs/1804.03267v1, 2018

[26] J B DeBrota, C A Fuchs, Rüdiger Schack, Respecting One's Fellow: QBism's Analysis of Wigner's Friend. https://arxiv.org/abs/2008.03572, 2020

[27] R B Griffiths, Consistent Histories and the Interpretation of Quantum Mechanics. J. Stat. Phys. 36, 219-272, 1984

[28] R Omnes, Quantum Philosophy. Princeton University Press, 1999