Multi-Observer Quantum Mechanics from Classical Bayesian Inference

Kevin Player*

May 22, 2023

> Solipsism may be logically consistent with present Quantum Mechanics, Monism in the sense of Materialism is not.
>
> Eugene Wigner

Abstract

We present several quintessential quantum ideas and shed them in a classical light. We argue how quantum information theory is a generalization of classical information theory in a nonstandard way (without density matrices). In fact, we show how classical information theory can be embedded in quantum information theory using zero phase wavefunctions. We employ this embedding to motivate a new multi-observer extension of quantum mechanics. Finally, we outline an experiment to test the existence of our multi-observer theory.

1 Mathematical Framework

This section lays out a mathematical foundation for talking about information using an epistemic treatment. The arguments are rather technical and are not required for an appreciation of the main results. The reader is invited to skip ahead to Section 2 during a first read through.

Lets start with a 3-bit example, where each bit is realized by coins which are either heads (H) or tails (T). With these 3 coins in hand, we conduct some thought experiments. The ideas that we encounter all extend without loss of generality to n-bits or any finite set.

1.1 Tribits and Notation

Let us describe the situation once we throw the 3 coins. We move freely between binary bits (0 and 1) and coin states (H and T) mapping between them $H \leftrightarrow 0$ and $T \leftrightarrow 1$. Let

$$x_i \in \{H, T\}$$

be the outcome for the throws $i = 1, 2, 3$ and let $x = (x_1, x_2, x_3)$ denote the resulting sequence or tribit. There are 8 tribits and we will refer to them with the italic integers 0 through 7 in binary order. Let C be the set of 8 possible tribits

$$C = \{HHH, HHT, \cdots, TTH, TTT\} = \{0, \cdots, 7\}$$

---

*kjplaye@gmail.com

1.2   Postulates for Constructing Knowledge Statements

For any set A, let $\mathcal{K}(A)$ be the set of knowledge statements about A. Motivating by example, we construct and present knowledge statements $\mathcal{K}(C)$ for C. We do this by repeatedly applying the following simple postulates.

Postulate 1.1. *For any tribit $x \in C$ we have the knowledge statement $b_x \in \mathcal{K}(C)$, that says we know x with certainty or probability 1. We call this a basic statement.*

There are 8 tribits, so there are 8 basic statements, $b_0, \cdots, b_7$. Each statement is a separate kind of knowledge about the tribit. A physical constraint is that an observer can only hold one knowledge statement at any given time. In contrast, a knowledge statement itself may represent more than one possibility simultaneously, as is the case in the next Postulate.

Postulate 1.2. *A formal superposition can be made for any statements $s_1 \in \mathcal{K}(C)$ and $s_2 \in \mathcal{K}(C)$ and $\alpha \in [0, 1]$*

$$super_\alpha(s_1, s_2) = \alpha s_1 + (1 - \alpha)s_2. \tag{1}$$

This formal combination means to take statement s_1 with probability α and statement s_2 with probability $(1 - \alpha)$; This is a new statement in $\mathcal{K}(C)$.

1.3   Distributions on the Ensemble

For any set A let $\mathcal{P}(A)$ be the set of probability distributions on A. Then $\mathcal{P}(C)$ is the set of probability distributions on tribits. For a fixed $\eta \in \mathcal{P}(C)$ we let $p_i(\eta)$ be the probability of seeing the tribit i under η. For an example, consider the tribit $6 = TTH$. p_6 is the probability of seeing 6 under η.

The eight p_i completely characterize η as the vector $(p_0(\eta), \cdots, p_7(\eta)) \in \mathbb{R}^8$. Call this 8 dimensional ambient space $\mathcal{V}(C)$. In general $\mathcal{V}(A) = \mathbb{R}^A$ for the set A. We identify any $a \in A$ with the basis element b_a. In this way we map $A \hookrightarrow \mathcal{V}(A)$.

Lemma 1.1. $\mathcal{K}(C) = \mathcal{P}(C)$

Proof. Let $\eta \in \mathcal{K}(C)$. The Postulates 1.1 and 1.2, that form η, output probability distributions at every step. So $\eta \in \mathcal{P}(C)$.

WLOG we can assume that $p_0 \neq 0$. Let $\eta = (p_i) \in \mathcal{P}(C)$. We inductively extend the distribution

$$\eta_k = \gamma_k(p_0, \cdots, p_{k-1}, 0, \cdots, 0)$$

to

$$\eta_{k+1} = \gamma_{k+1}(p_0, \cdots, p_k, 0, \cdots, 0)$$

using basic b_k and

$$\eta_{k+1} = super_\alpha(\eta_k, b_k)$$

for appropriate choices α and of normalizations γ_k, γ_{k+1}. So we have applied a sequence of Postulates that form η and hence $\eta \in \mathcal{K}(C)$. $\square$

Corollary 1.1.1. For any finite set A,

$$\mathcal{K}(A) = \left\{ (p_a)_{a \in A} \in \mathbb{R}^A \,\middle|\, \sum_{a \in A} p_a = 1 \right\} \subseteq \mathcal{V}(A)$$

where $\mathcal{K}(A)$ is affine codimension 1 in $\mathcal{V}(A)$ and $\mathcal{V}(A)$ is the vectorspace $\mathbb{R}^A$.

Proof. The right hand side is the definition of $\mathcal{P}(A)$ and the claim follows using a generalization of the proof of Lemma 1.1. $\square$

A general knowledge statement $\eta \in \mathcal{K}(C)$ combines the basic statements $b_0, \cdots, b_7$ using coefficients $p_0, \cdots, p_7$,

$$p_0, \cdots, p_7 \in \mathbb{R}_{\geq 0}, \tag{2}$$

so that $\eta = \sum p_i b_i$. To be a statistical distribution we also need it to be normalized with $\sum p_i = 1$. Note that now (1) is more than just a formal relation since it actually holds linearly inside of $\mathcal{V}(C)$.

Let $\mathcal{S}$ be the category of sets. Let $\mathcal{W}$ be the category of real vectorspaces. Then

Lemma 1.2. $\mathcal{V}$ *is a contravariant functor from* $\mathcal{S}$ *to* $\mathcal{W}$ *that takes injections and surjections to surjections and injections respectively.*

Proof. Let $f : A \hookrightarrow B$ be an injection. Consider the map for $b \in B$ to $\mathcal{V}(A)$ to be as follows

$$b \mapsto \begin{cases} b & \text{if } b \in A \\ 0 & \text{if } b \notin A \end{cases}$$

We extend this map linearly to form the orthogonal projection $\mathcal{V}(f) : \mathcal{V}(B) \twoheadrightarrow \mathcal{V}(A)$.

Now let $g : A \twoheadrightarrow B$ be a surjection. Consider the map for each $b \in B$ to the sum over the fiber of b
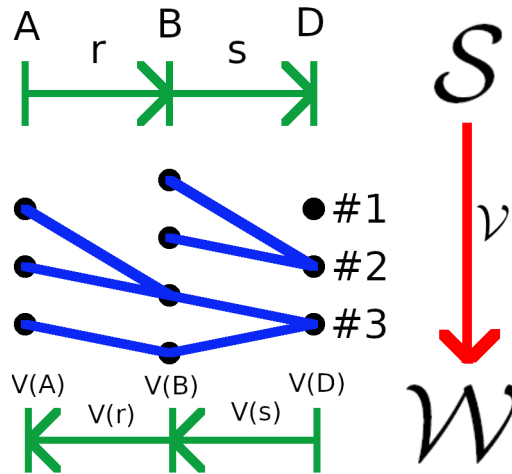
$$b \mapsto \sum_{a \in g^{-1}(b)} a.$$

We linearly extend this to the map $\mathcal{V}(g) : \mathcal{V}(B) \hookrightarrow \mathcal{V}(A)$. Each of these fibers is disjoint so the map is injective.

Consider a general map $h : A \to B$. We write this map as a surjection onto the image followed by an injection $A \twoheadrightarrow Im(h) \hookrightarrow B$. We then send h to

$$\mathcal{V}(A) \hookleftarrow \mathcal{V}(Im(h)) \twoheadleftarrow \mathcal{V}(B).$$

Finally, we examine the composition of $r : A \to B$ and $s : B \to D$. For any $d \in D$ we have three possibilities

1. $d \notin Im(s)$

2. $d \in Im(s)$, but $s^{-1}(d) \cap Im(r) = \emptyset$

3. $d \in Im(s \circ r)$

We wish to show the composition property of the functor, $\mathcal{V}(s \circ r) = \mathcal{V}(r) \circ \mathcal{V}(s)$, by considering the three above possibilities on a case by case basis. Consider possibility #1. In this case we have $\mathcal{V}(s)$ sending d to zero. For possibility #2, $\mathcal{V}(s)$ sends d to a sum over the fiber of d under s, but then all of $s^{-1}(d)$ gets sent to zero by $\mathcal{V}(r)$. For possibility #3, $\mathcal{V}(s)$ sends d to the sum over $s^{-1}(d)$ and then that gets sent by $\mathcal{V}(r)$ to the sum over

the nonempty $r^{-1}(s^{-1}(d))$. These cases and maps are the same as directly considering $\mathcal{V}(s \circ r)$ where the cases #1 and #2 have d getting sent to zero and case #3 defines the same map as the composition since the fibers under r of the fibers under s are just the fibers under $s \circ r$.

This demonstrates the composition property of the functor, $\mathcal{V}(s \circ r) = \mathcal{V}(r) \circ \mathcal{V}(s)$. $\square$

Each coin corresponds to a 2 dimensional space $\mathcal{V}(\{H, T\})$ which we will call W. The 8 dimensional space $W^{\otimes 3}$ has a basis with vectors like $T \otimes H \otimes T$. Then $\mathcal{V}(C)$, is identified with the tensor power $W^{\otimes 3}$ by linearly extending

$$(x_1, x_2, x_3) \mapsto x_1 \otimes x_2 \otimes x_3.$$

Lemma 1.3. $\mathcal{V}(A \times B) = \mathcal{V}(A) \otimes \mathcal{V}(B)$ and the n-bit knowledge space $\mathcal{V}(\mathbb{Z}_2^n)$ is $W^{\otimes n} \cong \mathbb{R}^{2^n}$. $\mathcal{V}(A \times B \twoheadrightarrow A)$ identifies $\mathcal{V}(A)$ with $\mathcal{V}(A) \otimes \mathbb{1}_B$. The same is true for B and $\mathbb{1}_A$.

Proof. We map the $|A| \cdot |B|$ basis elements $(a, b) \mapsto a \otimes b$. The linear extension of the map is an isomorphism. For n-bits we repeatedly apply the map to get the formula for $\mathcal{V}(\mathbb{Z}_2^n)$. For the last expression write $\mathbb{1}_B := \sum_{b \in B} b$ then the sum over the fiber of $a \in A$ is $a \otimes \mathbb{1}_B$. The same argument applies when reversing a and A with b and B. $\square$

2   Classical Information Theory

We are ready to enter into the physics surrounding knowledge statements. This means we are going to start talking about measurement and observation. This is part of our epistemic program; we will take a Bayesian viewpoint that knowledge is updated by observation.

2.1   Classical Bayesian Projections

Let us go back to our three coin toss example with eight tribits C. We consider knowledge statements about the three coins $\mathcal{K}(C)$ in an eight dimensional ambient space $\mathcal{V}(C)$ where C is a basis. We don't always get to observe the exact tribit. For instance, we might only get to know that the first coin is T and that the 2nd and 3rd coins are the same. This corresponds to an exemplar subset of C, which we will call S

$$\begin{aligned} S &= \{s \in C \,|\, x_1 = T, x_2 = x_3\} \\ &= \{4, 7\} \qquad \text{(THH, TTT)} \end{aligned}$$

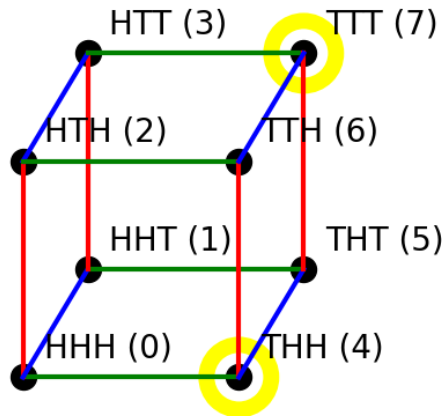We will use S as an example in the remaining text.

Figure 1: Coin sequences with S circled in yellow.

Someone could make a measurement that tells them that the tribit is in S and nothing else. This would be a specific case of what we will call a Bayesian projection, which we call ρ, and which updates our knowledge to be in S. The process which produces this we call a projection measurement. An initial knowledge statement will project onto the two dimensional space spanned by S. That is, it sends the two S tribits to themselves and the six non-S tribits to zero. Furthermore, we renormalize p_i on the remaining two $i \in S$. Putting this all together for this example on S,

$$\rho : (p_0, \cdots, p_7) \mapsto \left(0, 0, 0, 0, \frac{p_4}{p_4 + p_7}, 0, 0, \frac{p_7}{p_4 + p_7} \right) \tag{3}$$

In more generality we have

Definition 2.1 (Bayesian Projection). For any finite set A with $B \subseteq A$ we consider a measurement of B. It takes

$$p_a \mapsto \begin{cases} 0 & \text{when } a \notin B \\ \frac{p_a}{\sum_{b \in B} p_b} & \text{when } a \in B \end{cases}$$

We call this a Bayesian projection since the ambient spaces follow a projection $\mathcal{V}(B \hookrightarrow A)$.

2.2   Classical Bayesian Inference

A more general type of measurement is a probabilistic measurement. Someone could learn that there is a 95% chance that the tribit is in S. In full generality we will call such a probabilistic observation $\mathcal{O}$. We can figure out how to update our knowledge statement from p_i, to $\hat{p}_i$, using a relative[1] version of Bayes's rule

$$\frac{\hat{p}_i}{\hat{p}_j} = \underbrace{\frac{P(i|\mathcal{O})}{P(j|\mathcal{O})}}_{\text{Posterior}} = \underbrace{\frac{P(\mathcal{O}|i)}{P(\mathcal{O}|j)}}_{\text{Bayes Factor}} \underbrace{\frac{p_i}{p_j}}_{\text{Prior}}$$

Pulling out the Bayes factor we find that we just multiply by the likelihood and renormalize by γ

$$\hat{p}_k = \gamma P(\mathcal{O}|k) \ p_k. \tag{4}$$

Definition 2.2 (Bayesian Inference). A knowledge update of the form (4) is called a Bayesian inference.

Corollary 2.0.1. A special case are two-leveled projections $P(\mathcal{O}|k) \in \{0, v\}$, for a fixed value v determined by normalization. These are the Bayesian projections from Definition 2.1.

2.3   Classical Correlation

We consider the uniform distribution on S.

$$\frac{1}{2}(T \otimes H \otimes H) + \frac{1}{2}(T \otimes T \otimes T) \tag{5}$$

We know the first coin is T, which we give to Eve. She does nothing with her coin.

We have HH or TT with equal probabilities for the second and third coins. If we give the second coin to Alice and the third coin to Bob then we have a classical correlation. If Bob finds that the third coin is T then we know that Alice will also find that the second coin is T (and similarly for F).

The observers all undergo Bayesian projections here. The separate knowledge of Alice and Bob is changing due to observation.

---

[1] In the ratio, the contribution of $P(\mathcal{O})$ is canceled out and accounted for during normalization. $P(\mathcal{O})$ only holds significance prior to its observation and it doesn't require consideration during the update process.

Definition 2.3 (Classical Correlation)**.** In general consider $A \times B$ with surjections f_A and f_B onto A and B respectively. Consider the set of knowledge statements that factor as coordinate-wise products of $\mathcal{V}(f_A)$ and $\mathcal{V}(f_B)$. From Lemma 1.3 we see that such a product must be of the form

$$(\mathbb{1}_A \otimes \eta_B)(\eta_A \otimes \mathbb{1}_B) = \eta_A \otimes \eta_B$$

where $\eta_A \in \mathcal{K}(A)$ and $\eta_B \in \mathcal{K}(B)$. We call all other nonfactorable statements "correlated" since they do not directly come from knowledge statements of A and B independently. A simple example uses two bits with the following knowledge statement

$$\frac{1}{2}(H \otimes H) + \frac{1}{2}(T \otimes T) \tag{6}$$

Up until now we have been treating the coins in only one frame of reference, say the lab technician Lena's frame. But we are now faced with at least two other observers, Alice and Bob. Each of these observers has a different subjective knowledge of the coins. They independently learn through observation and adjust their knowledge accordingly. Also, there is a compatibility condition that they have to be consistent if they get back together.

Eve can observe the superposition of Allice and Bob's knowledge as (5), and Lena can talk with Eve to also know the superposition (5). Then Eve could spy on Allice or Bob to learn the exact bit. This could be done before Lena learns the bit.
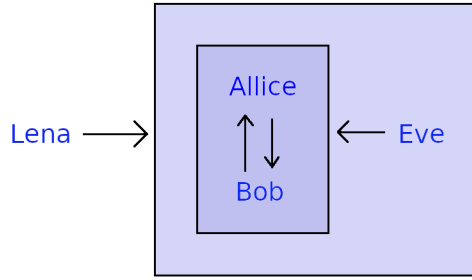

Figure 2: Various observers

2.4   More on Bayesian Projections

Bayesian projections can be thought of as giving the observer some information and reducing the entropy of the knowledge statement by the same amount. The following result shows that the information theoretic entropy always decreases under projection measurement.

Lemma 2.1. A nontrivial projection measurement decreases the entropy.

Proof. Consider a finite set X, the disjoint union $X = X_1 \sqcup X_2$, and a measurement that projects onto X_2. This sends $p_i \to 0$ for each $i \in X_1$ and renormalizes the elements of X_2 by $\sigma = \sum_{i \in X_2} p_i$. Let $H(X)$ be the original entropy and let $H(\pi)$ be the entropy after the projection.

$$H(X) - H(\pi) = \sum_{i \in X} -p_i \log(p_i) + \sum_{i \in X_2} \frac{p_i}{\sigma} \log\left(\frac{p_i}{\sigma}\right)$$

$$= \left(\sum_{i \in X_1} -p_i \log(p_i)\right) + \left(\sum_{i \in X_2} \left(-p_i + \frac{p_i}{\sigma}\right) \log(p_i)\right) - log(\sigma)$$

$$\geq 0$$

Where the equality occurs exactly when $p_i = 0$ for all $i \in X_1$ and $\sigma = 1$, which is when the projection is trivial. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Measurement can be thought of as a context change which extends what we call the observer to include the measurement itself. There is an energy gain in the observer system which corrosponds to the entropy lost in the knowledge. This accounts for the newly stored information, the measurement.

An overly simple example is pictured in Figure 3 where we let a test particle be absorbed into the observer system. Since the test particle is one wavelength away from the boundary, it can be regarded à la Bekenstein to be a part of the system [1], [2], [3], [4]. This limits the change in energy and entropy to the internal degrees of freedom of the particle. This means that we do not need to account for the spatial or temporal degrees of freedom.
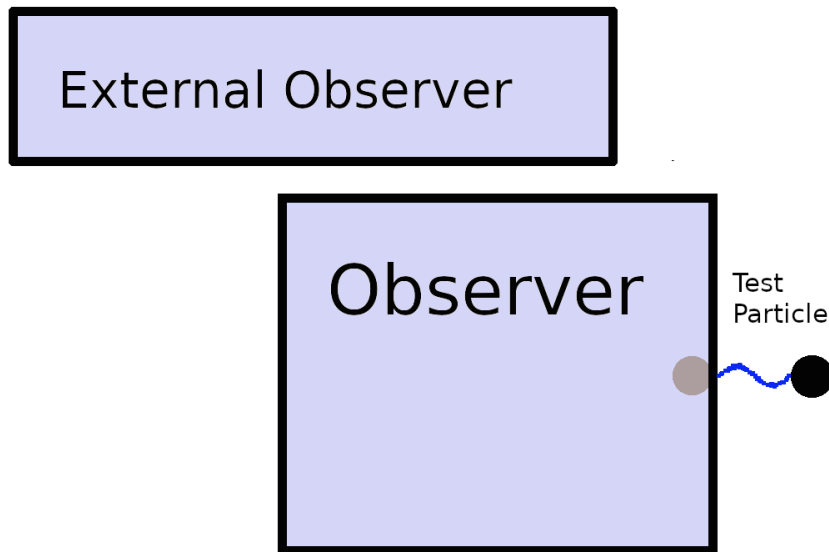

Figure 3: A test particle is one wavelength away from an observer. An external observer sees the internal observer entangled with the test particle.

Other observers outside of this interaction will just see the observer and test particle becoming correlated (later entangled). Outside observers have to measure the system themselves to update their seperate knowledge.

3   Quantum Information Theory

3.1   Classical to Quantum Embedding

General quantum information[2] about 3-qubits can be expressed as a wave function.

$$q_0, \cdots, q_7 \in \mathbb{C}$$

with $\sum |q_i|^2 = 1$; compare to the positive real case in equation (2). The Born rule is ostensibly a map $q_i \mapsto |q_i|^2 = p_i$ to classical probability. In the following subsections we will instrument all of the classical information theoretic constructs from Section 2 by restricting the phase[3] of q_i to be zero. We can map backward $p_i \mapsto \sqrt{p_i} = q_i$. This commutes with the Born rule (subject to normalization):

---

[2]We are purposely leaving out density matrices and POVMs and just dealing with pure states.
[3]We consider negative values as 180 degrees out of phase, so zero phase means non-negative real.

$$\text{(Quantum)} \qquad \mathbb{C}^n \rightleftarrows (\mathbb{R}_{\geq 0})^n \qquad \text{(Classical)}$$

In doing so, we will see that quantum information theory needs to instrument the classical. In particular we will see that quantum measurement and entanglement restrict to classical measurement and classical correlation for zero phase.

3.2   Quantum Bayesian Projection

We illustrate zero phase Bayesian projection with an example. For the running example we define a projection which sends the 6 tribits that are not in S to zero, but sends the 2 tribits in S to themselves. In quantum notation we can write this as an operator π

$$\pi = |TFF\rangle \langle TFF| + |TTT\rangle \langle TTT|$$

which we compare to Section 2.1. Then let A be any operator with an eigenspace, with eigenvalue λ, equal to the image of π. An observation of λ would correspond to an application of π, which is a Bayesian projection as in Definition 2.1.

Theorem 3.1. The set inclusion $B \subseteq A$ yields a Bayesian Projection as in Definition 2.1. Then a zero phase wavefunction instruments the classical measurement.

Proof. Let $q_i = \sqrt{p_i}$ be a zero phase wavefunction corresponding to $(p_i) \in \mathcal{K}(A)$. Let

$$\lambda_i = \begin{cases} 1 \text{ if } i \in B \\ 2 \text{ if } i \notin B \end{cases}$$

and define an operator

$$D = \sum_{i \in A} \lambda_i |i\rangle \langle i|$$

We take a measurement of (p_i) with respect to D. If we measure 1, then the wave function collapses by projecting onto the eigenspace of 1. The wave function is then renormalized. This projection on the B-space and normalization of q_i exactly matches the classical Bayesian projection and normalization of p_i. $\qquad\square$

3.3   Quantum Bayesian Inference

Theorem 3.2. Let $P(\mathcal{O}|i)$ define a Bayesian inference as in Definition 2.2. Then a zero phase wavefunction instruments the classical measurement.

Proof. Consider, in the fashion of [5], a more general measurement with matrix $M_\mathcal{O}$ which is diagonal with entries

$$M_\mathcal{O}(i, i) = \sqrt{P(\mathcal{O}|i)}$$

This measurement is identical to the classical Bayesian inference for zero phase wavefunctions. $\qquad\square$

3.4   Quantum Entanglement

Theorem 3.3. Classical correlation knowledge statements become entangled zero phase wavefunction.

Proof. The example (6) in Definition 2.3 maps to

$$\frac{1}{\sqrt{2}}(H \otimes H) + \frac{1}{\sqrt{2}}(T \otimes T) \tag{7}$$

which is an entangled Bell state. In general, the correlated knowledge statements are exactly the zero phase entangled states. $\qquad\square$

Figure 3 diagrams how measurement between an observer and test particle become an entanglement from an outside observer's perspective.

3.5   Entropy Issues

Von Neumann entropy is left out of this note since it does not reduce to classical entropy in the embedding. This is because the embedded classical wavefunctions are pure states which all have zero Von Neumann entropy, but the classical entropy is nonzero for all non-basic knowledge statements.

4   Overview of the Generalization

The quantum mechanical wave function collapse, entanglements and ensembles generalize the classical information theoretic concepts of Bayesian inference, correlation and statistical ensembles respectively. Wave function collapsing instruments Bayesian inference and classical measurement. Entanglement instruments classical correlation. We are not claiming that wavefunctions are "just" knowledge statements, but that they are "at least" knowledge statements. This is enough to motivate new ways of doing quantum mechanics, generalizations which we will encounter in the next sections.

4.1   Multi-Observer Quantum Mechanics

We focus on quantum wave function collapse as a generalization of classical Bayesian inference. We can implement the classical Bayesian inference within zero phase quantum mechanics as outlined above. So the classical Bayesian theory has a direct tie in; that measurement occurs in tandem with a change in knowledge[4].

In the classical theory, knowledge is local to the observer[5], multiple observers each have their own knowledge. For instance, Wigner's friend and Wigner each have their own classical knowledge. It would seem then that Wigner's friend and Wigner must have different zero phase wavefunctions as well, if they are to be generalizing classical knowledge. This is our prediction for multi-observer quantum physics, that every observer has a unique local wavefunction.

4.2   Prediction and Experiment

We predict that multi-observer quantum mechanics will be required for a proper generalization of classical knowledge. We propose an experiment where we inject an "observer" into a quantum eraser[6]. The injected observer will have to be a small apparatus that is

- able to record a measurement ψ of the eraser particle path.
- able to forget the measurement ψ.
- able to demonstrate that a measurement was made with a record ρ.

Let ϕ be the lab technician's wavefunction. In ϕ the apparatus is entangled with the subject particles in the eraser experiment. We need to find out if there is another wavefunction in play that is not just part of ϕ. The key here is that the apparatus is able to

---

[4]This change in knowledge is tangible, it always occurs as a transfer of matter/energy from the environment to the observer [4].

[5]Note that current quantum theory is a theory of one observer, usually the experimenter in a lab. An immediate subject is quantum key distribution(QKD), which requires at least three observers, Alice, Bob, and Eve; the security of QKD depends on multi-observer quantum ontology.

[6]Here we probe the boundaries of what constitutes an observer and a measurement. We claim that all versions of observer and measurement will be detectable in this experiment whenever it can be done.

demonstrate that it "knew" something, or in other words that another wavefunction ψ existed, using ρ.
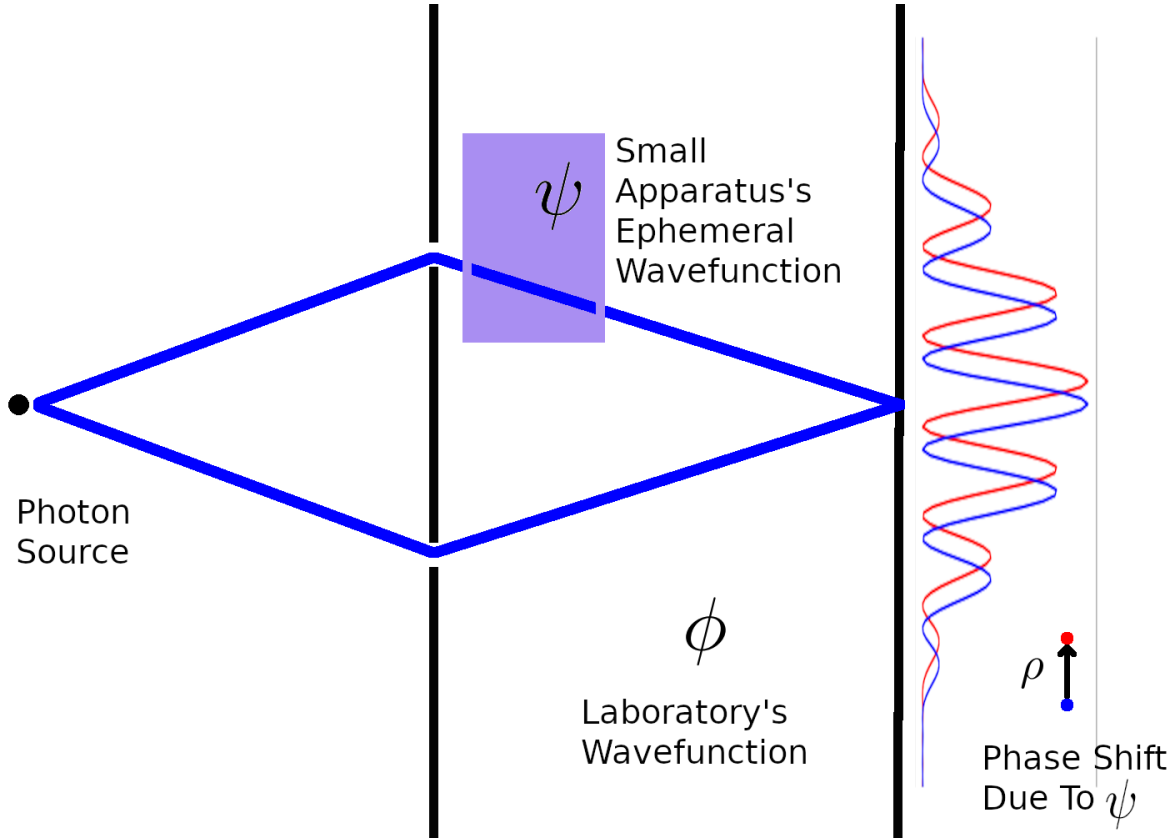

Figure 4: Generic experimental set up. The laboratory has a wave function ϕ that includes the experimental apparatus, the technician, and the photon. The apparatus interacts with the photon, making measurements which are encoded into an ephemeral local wave function ψ. From the laboratory's point of view, the photon and the apparatus are entangled. The temporary existence of ψ manifests as a phase shift on the top path. This would be evident in the interference pattern existing and shifting, which we call ρ.

The apparatus can not keep ψ for proper erasure, but must "forget" it. Observed erasure, via an interference pattern, proves that ψ is not a part of ϕ. Finally, there should be a record ρ in the apparatus that it did at one point in time record a measurement ψ. Receipt of the report ρ and eraser interference pattern together show that ψ and ϕ were necessarily part of a two-observer system.

4.3   A Overly Simple Example

An overly simplified case is letting the "apparatus" be a piece of glass. The atoms in the glass interact with the photon and cause a slow down inside of the material. This results in a phase shift and a shifting interference pattern. An interacting electron temporarily has an increase in energy, for a short time. This increase in energy records the fact that the photon in on the top path. We call this knowledge ψ. The electron eventually passes the energy back to the electromagnetic field, forgetting ψ, with phase shift evidence ρ.

5   Acknowledgments

References

[1] E. Verlinde, "On the origin of gravity and the laws of newton," *Journal of High Energy Physics*, vol. 2011, no. 4, pp. 1–27, 2011.

[2] J. D. Bekenstein, "Black holes and entropy," *Phys. Rev. D*, vol. 7, pp. 2333–2346, Apr 1973.

[3] K. Player, "Information theory and the physics of "matter channels"," 2023.

[4] K. Player, "The unruh effect as holographic thrust," 2023.

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition.* Cambridge University Press, 2011.