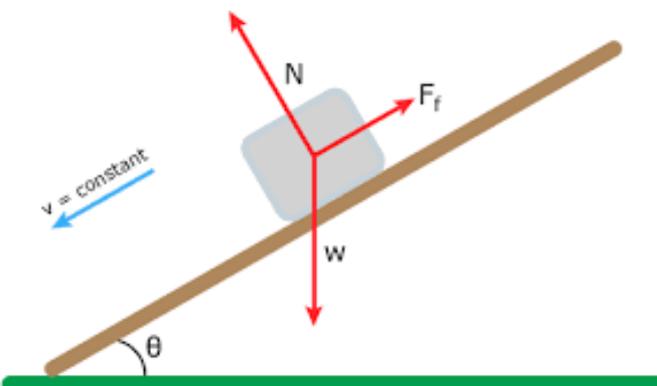
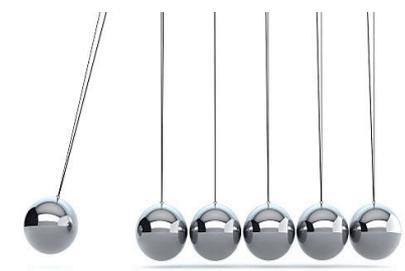
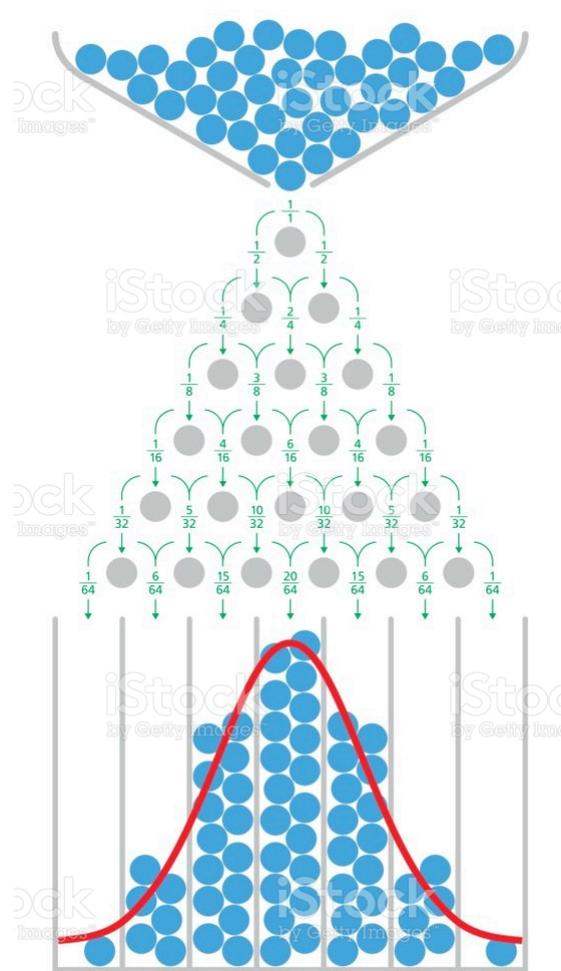
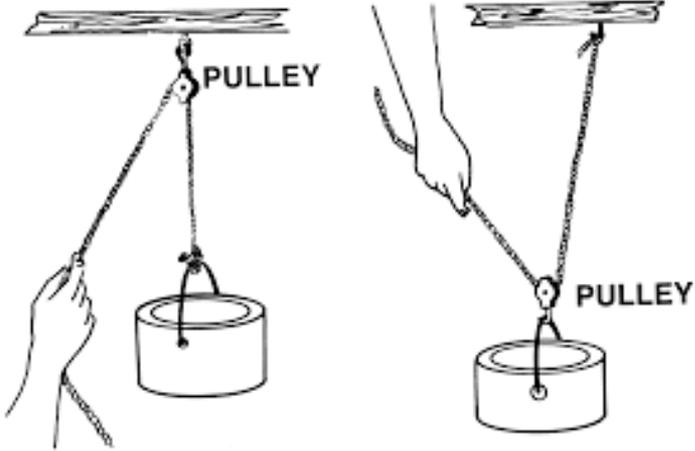


Quantum Theory and Twenty-first Century Physics

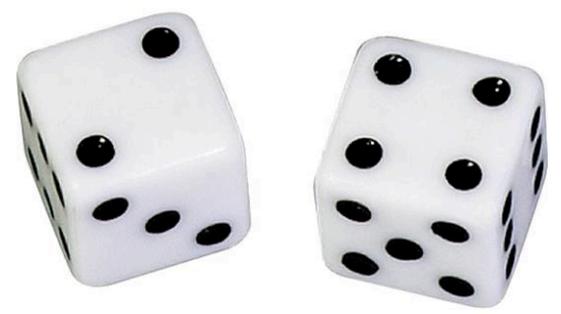
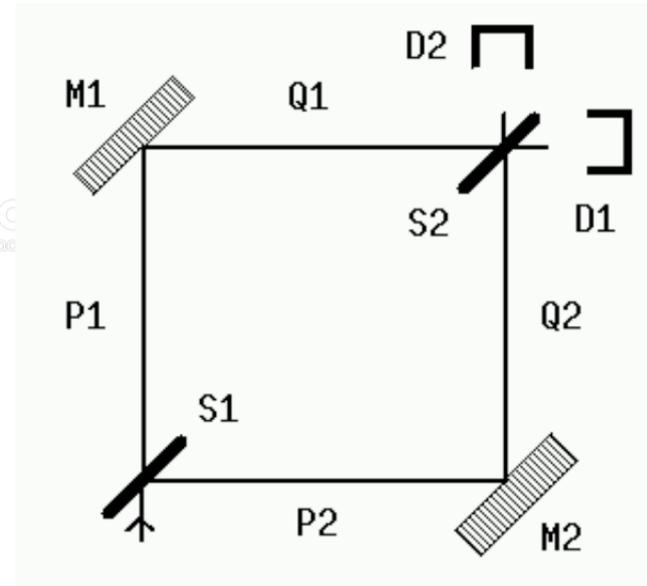
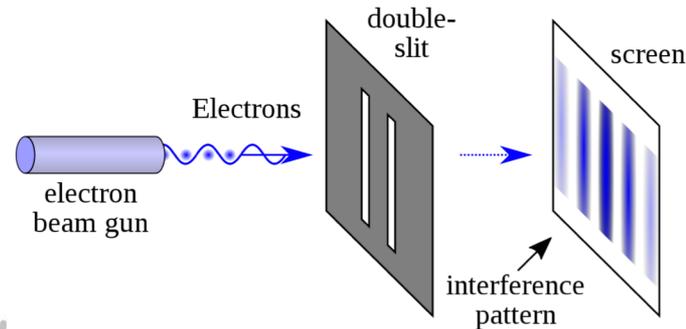
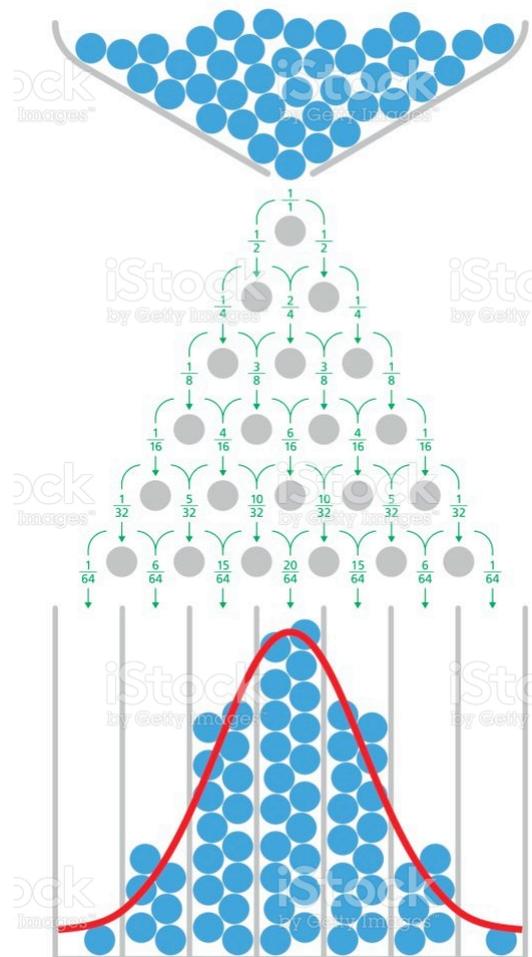
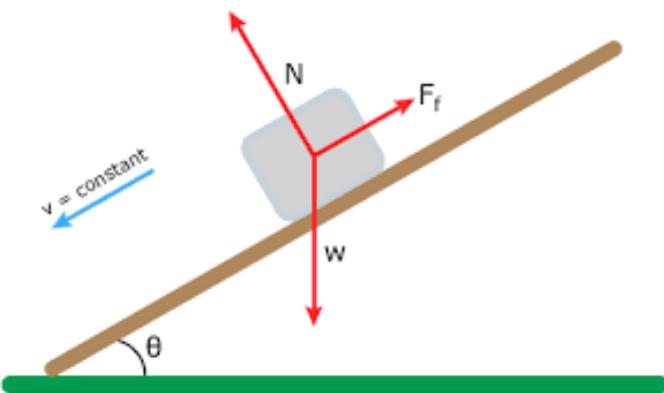
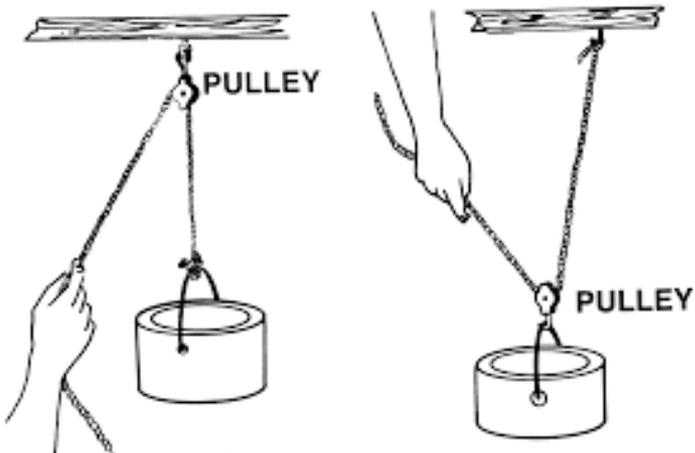
Saul Youssef

Boston University
November 2, 2021

Physics Lab



Physics Lab



691503838

What does it mean to modify probability theory?

$$(a \rightarrow b) \in P$$

Property of L	Expected property of P
\wedge is associative	$*$ is associative
\vee is associative	$+$ is associative
\wedge is commutative	—
\vee is commutative	$+$ is commutative
\wedge distributes over \vee	$*$ distributes both ways over $+$
\vee distributes over \wedge	—
0 is the minimum	P has an additive identity “0”
1 is the maximum	P has a two-sided multiplicative identity “1”

- Probabilities can be any real associative algebra with unit.

R.T. Cox, *Am. J. Phys.* 15, 1 (1946).

S. Youssef, *A reformulation of quantum mechanics*, *Mod. Phys. Lett*, A6, 225-236 (1991).

S. Youssef, *Quantum Mechanics as Complex Probability Theory*, *Mod. Phys.Lett.* A9, 2571 (1994).

S. Youssef, *Physics with Exotic Probability Theory*, arXiv:hep-th/0110253, (2001).

But what does probability “ $3 + 5i$ ” mean?

First, consider how this works in standard probability theory...

1. Consider N copies of the predicted situation with probability p .
2. The probability of n/N successful predictions peaks at $p = n/N$.
3. With the additional assumption which is, roughly, “probability zero events don’t actually happen,” we predict frequency $n/N = p$.

As pointed out by Ed Jaynes, there is nothing intrinsic about probability and frequency being equal. Probability theory would work just as well if ≥ 1 would also work if the predicted frequency was $1/p$.

S. Youssef, Physics with Exotic Probability Theory, arXiv:hep-th/0110253, (2001).

Everything works if probabilities have the additional property of being a square norm...

With this, standard probability returns for subsets of X .
 "wave functions" are functions from X to P

In order to get real non-negative numbers from probabilities, we take P to have a square norm $\| \cdot \|: P \rightarrow \mathbf{R}^{0,+}$ satisfying $\| p q \| = \| p \| \| q \|$ for $p, q \in P$. Given this, we will show that, under certain conditions,

$$\text{Prob}_t(b|a) = \frac{\int_X \| \Psi_{a \rightarrow b}^t \|}{\int_X \| \Psi_{a \rightarrow 1}^t \|} \quad (7)$$

is a probability in the ordinary sense. When it doesn't cause confusion, we will suppress the function name inside integrals as a notational convenience. We may, for example, write

$$\text{Prob}_t(b|a) = \frac{\int_X \| a \rightarrow b \wedge x_t \|}{\int_X \| a \rightarrow 1 \wedge x_t \|}. \quad (8)$$

Note that probabilities like $(a \rightarrow b \wedge c \wedge x_t)$ are typically zero and, of course, $(a \rightarrow x_t)$ isn't equal to $\Psi_a^t(x)$.

To derive properties of Prob_t , note that

$$\text{Prob}_t(b \wedge c|a) = \frac{\int_X \| a \rightarrow b \wedge c \wedge x_t \|}{\int_X \| a \rightarrow x_t \|} \quad (9)$$

is equal to

$$\frac{\int_X \| a \rightarrow b \| \| a \wedge b \rightarrow c \wedge x_t \|}{\int_X \| a \rightarrow x_t \|} * \frac{\int_X \| a \wedge b \rightarrow x_t \|}{\int_X \| a \wedge b \rightarrow x_t \|} \quad (10)$$

and, rearranging and using $\| a \rightarrow b \| \| a \wedge b \rightarrow x_t \| = \| a \rightarrow b \wedge x_t \|$, we have

$$\text{Prob}_t(b \wedge c|a) = \text{Prob}_t(b|a) \text{Prob}_t(c|a \wedge b) \quad (11)$$

as desired. If we also knew that for $b \wedge c = 0$,

$$\text{Prob}_t(b \vee c|a) = \text{Prob}_t(b|a) + \text{Prob}_t(c|a) \quad (12)$$

Thus, probabilities must be an associative algebra with unit, with a square norm.

...but there are exactly three such algebras...

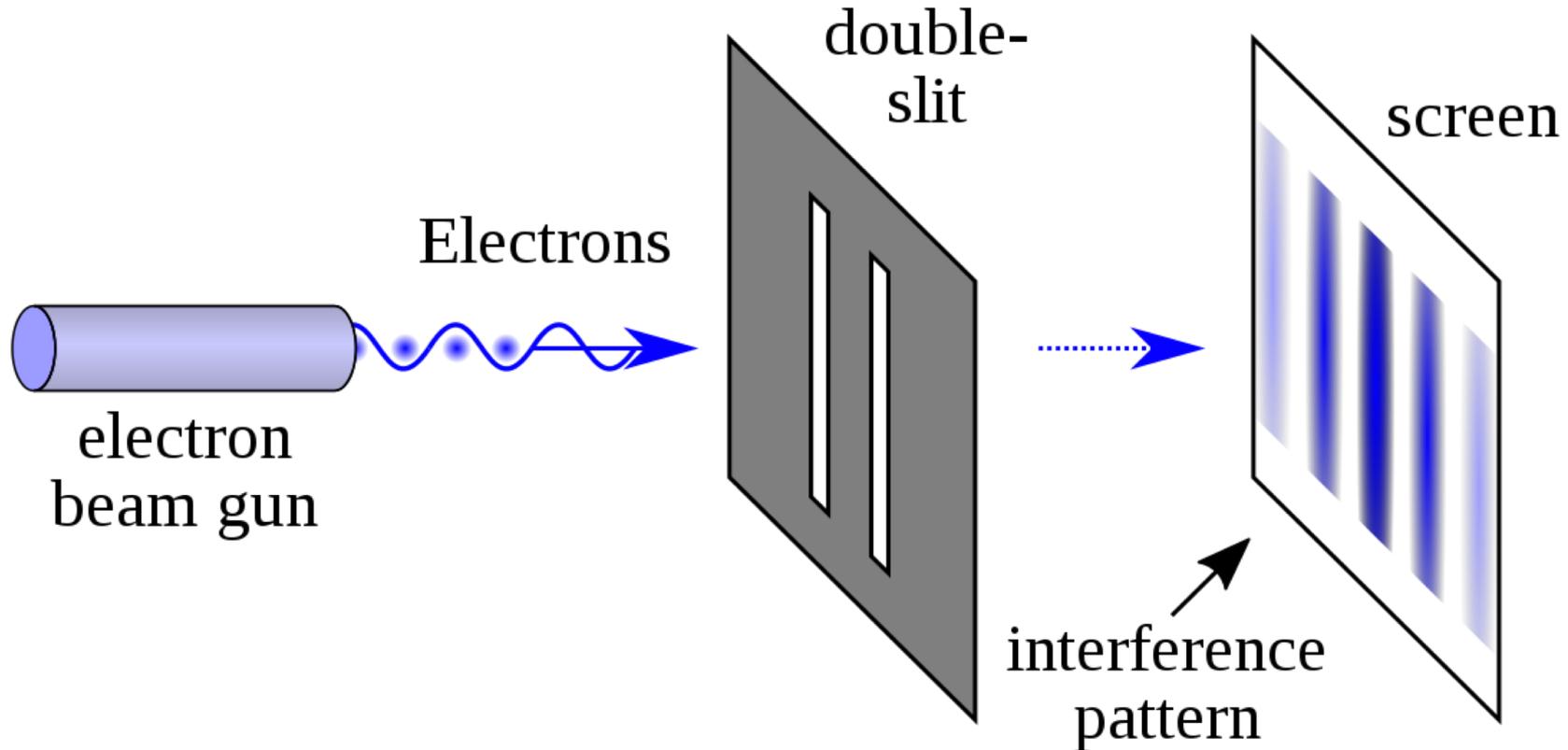
- The real numbers
- The complex numbers
- The quaternions

From this point of view, to define a quantum theory, simply....

1. Choose a “state space” of disjoint propositions X .
2. Choose **R**, **C**, or **H**.

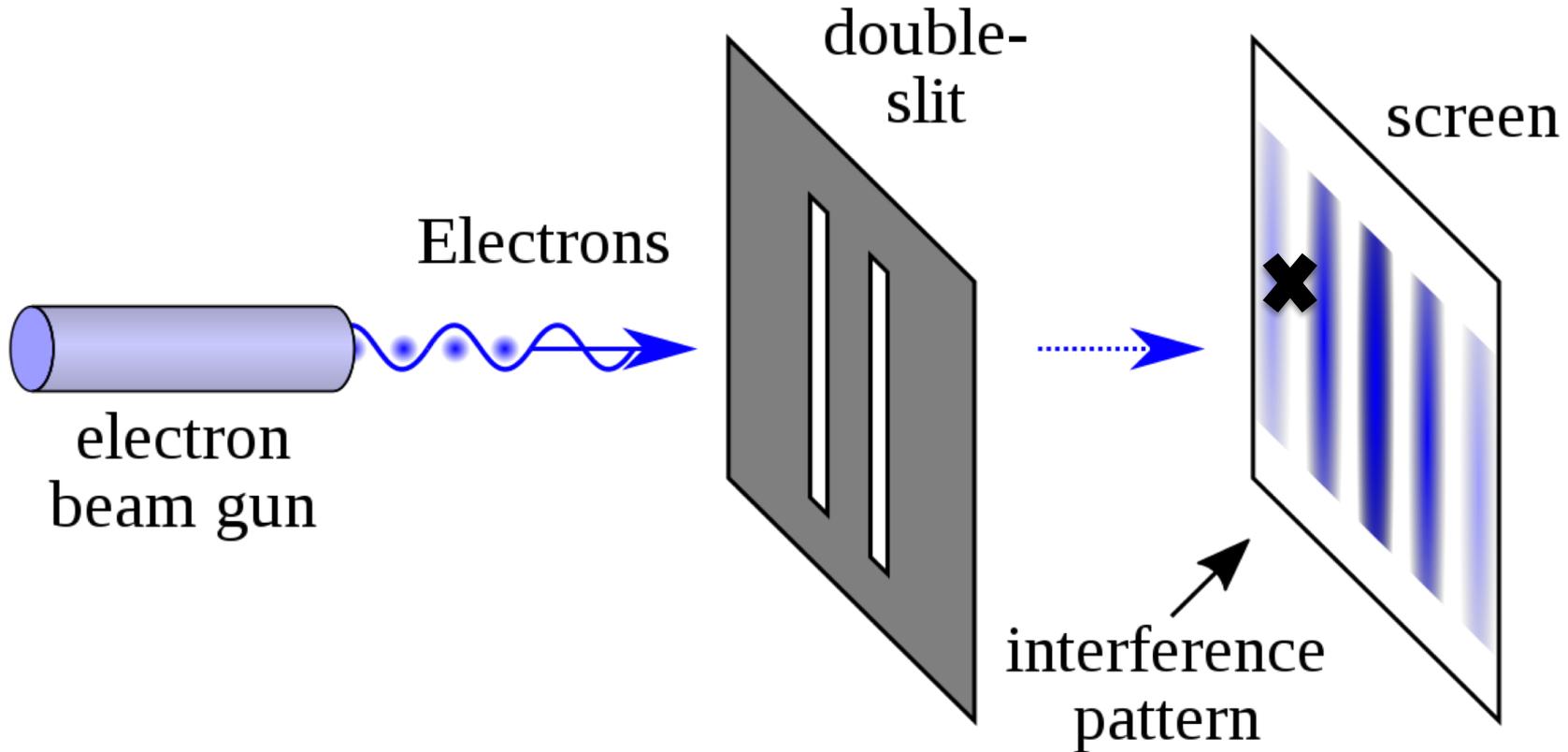
Hurwitz theorem. See, for instance *Spinors and Calibrations*, by F. Reese Harvey, Academic Press, (1990).

Example:



“It’s not true that the electron either goes through one slit or the other slit.”

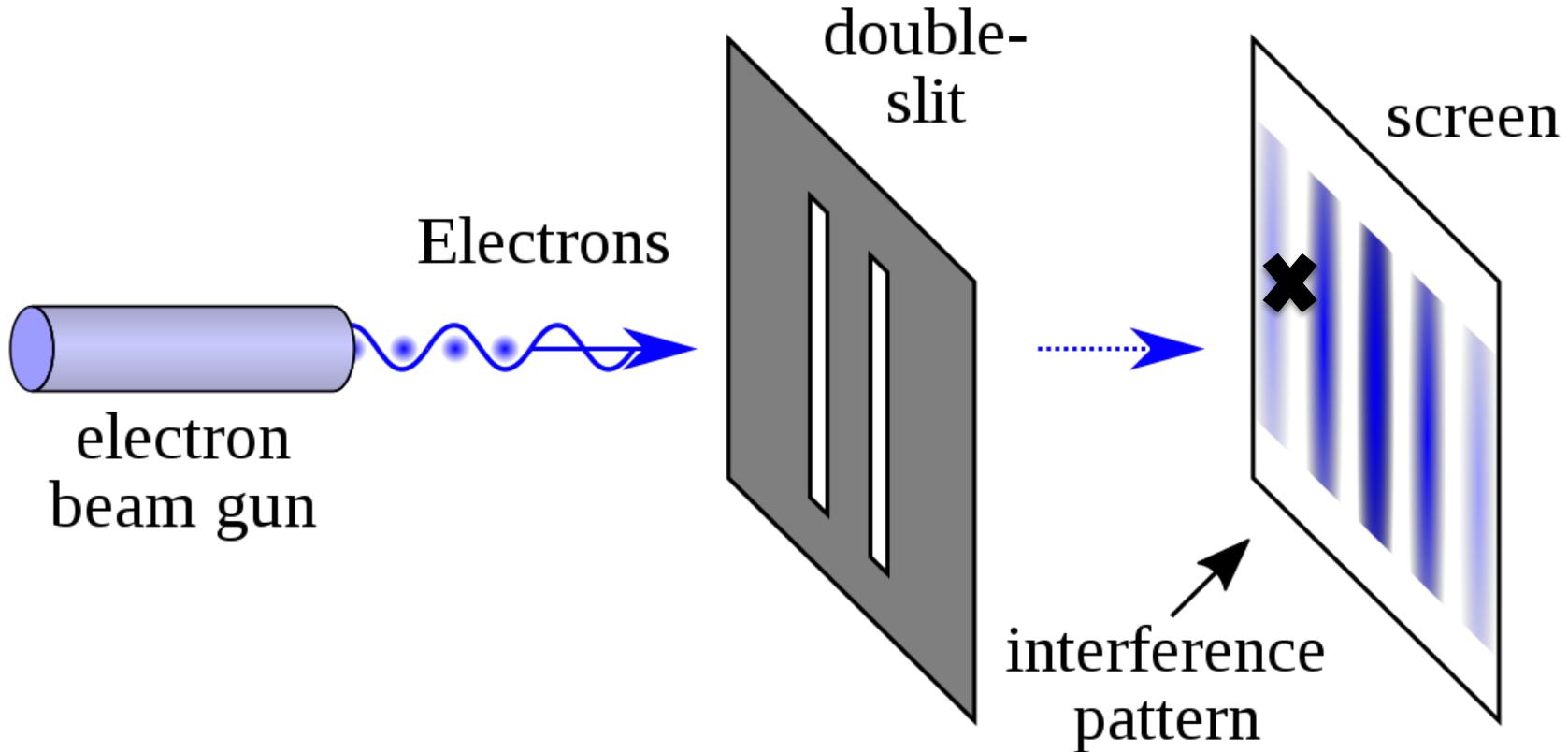
What does this mean?



“It’s not true that the electron either goes through one slit or the other slit.”

$$\text{Prob}(\text{arrive at } \times) \geq \text{Prob}(\text{arrive at } \times \mid \text{slit 1})$$

What does this mean?



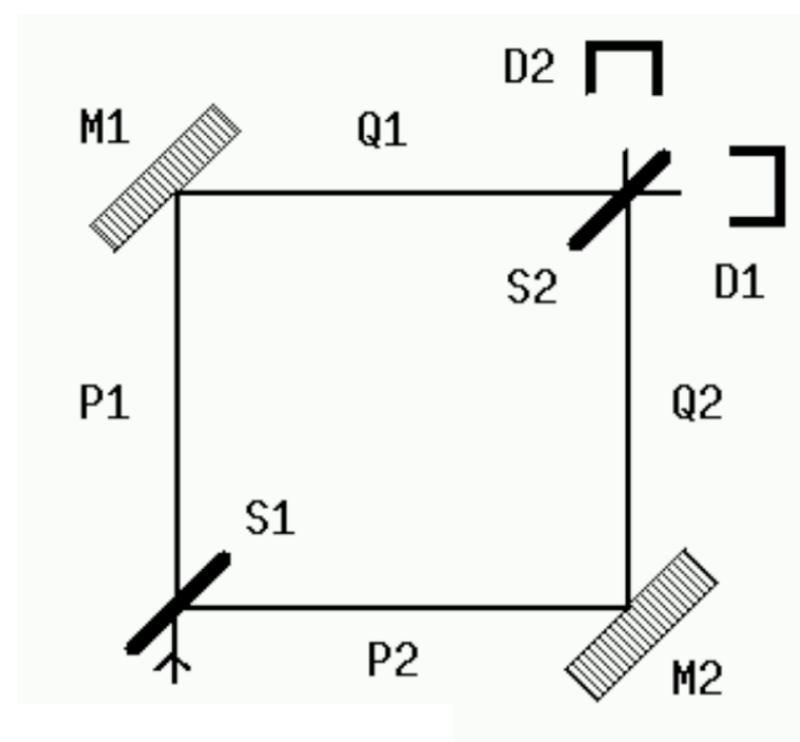
“It’s not true that the electron either goes through one slit or the other slit.”

← This is actually wrong!

$$\text{Prob}(\text{arrive at } \times) \geq \text{Prob}(\text{arrive at } \times \mid \text{slit 1})$$

State space: \mathbf{R}^3

Probability: \mathbf{C}



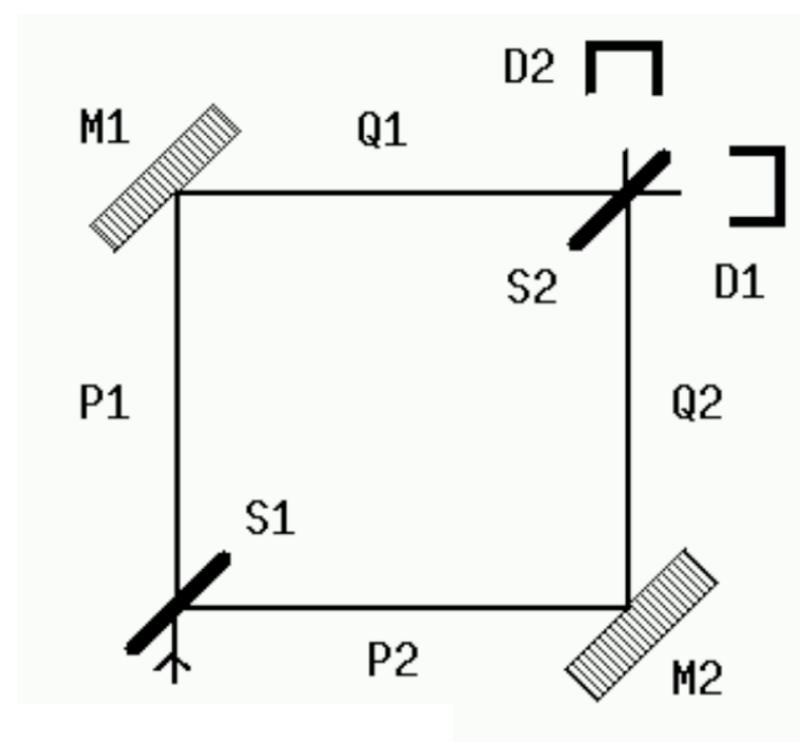
$$(e \rightarrow D_j) = \sum_{n,m=1}^2 (e \rightarrow P_n)(e \wedge P_n \rightarrow Q_m)(e \wedge P_n \wedge Q_m \rightarrow D_j). \quad (25)$$

Since P_1 is equivalent to a point in X , previous knowledge is irrelevant and we have $(e \wedge P_n \rightarrow Q_m) = (P_n \rightarrow Q_m)$. We also clearly want to assume that the particle can't hop the rails, in other words we assume that $(P_n \rightarrow Q_m)$ is zero unless $n = m$. This causes one of the sums to disappear giving

$$(e \rightarrow D_j) = \sum_{n=1}^2 (e \rightarrow P_n)(P_n \rightarrow Q_n)(Q_n \rightarrow D_j) \quad (26)$$

But what about which path arguments?...

“If you can tell which path is taken, the interference is lost.”

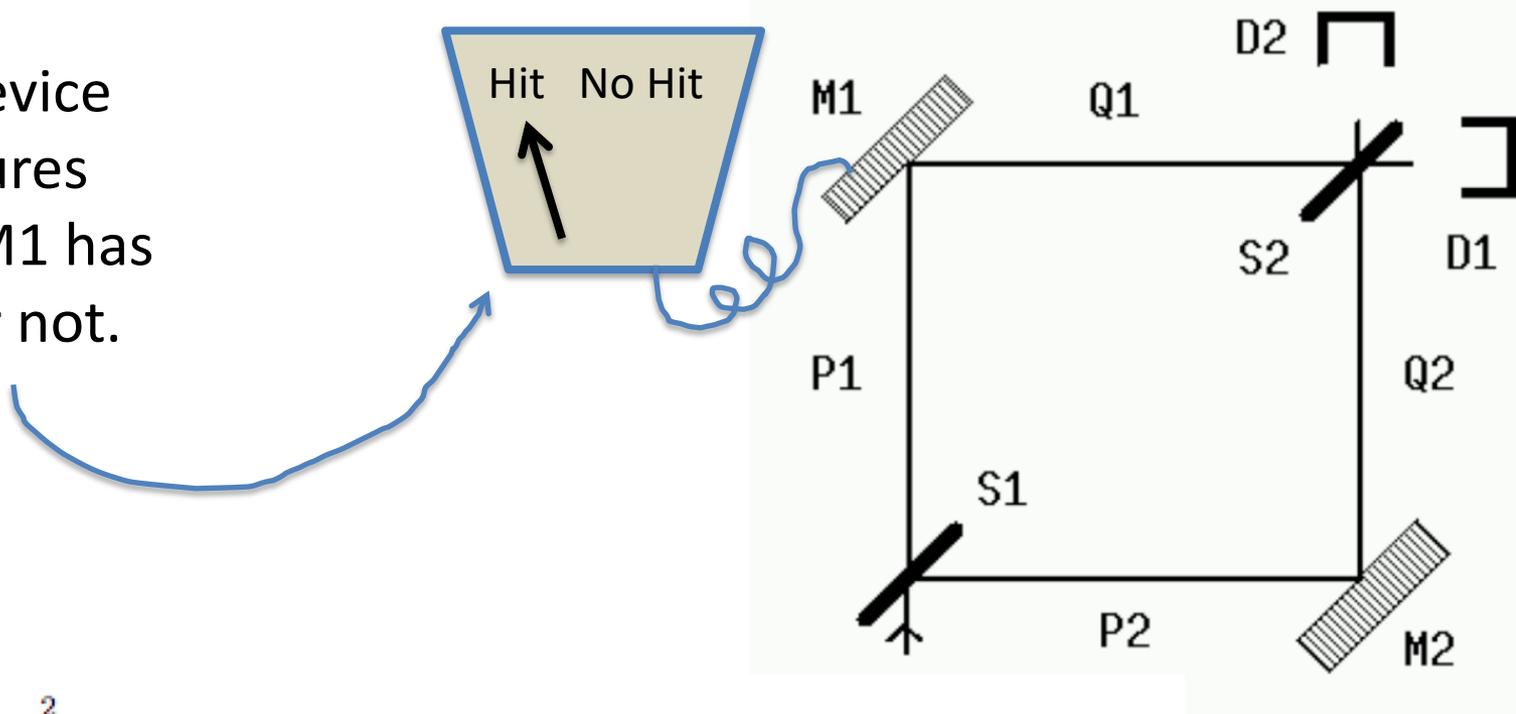


$$(e \rightarrow D_j) = \sum_{n,m=1}^2 (e \rightarrow P_n)(e \wedge P_n \rightarrow Q_m)(e \wedge P_n \wedge Q_m \rightarrow D_j). \quad (25)$$

Since P_1 is equivalent to a point in X , previous knowledge is irrelevant and we have $(e \wedge P_n \rightarrow Q_m) = (P_n \rightarrow Q_m)$. We also clearly want to assume that the particle can't hop the rails, in other words we assume that $(P_n \rightarrow Q_m)$ is zero unless $n = m$. This causes one of the sums to disappear giving

$$(e \rightarrow D_j) = \sum_{n=1}^2 (e \rightarrow P_n)(P_n \rightarrow Q_n)(Q_n \rightarrow D_j) \quad (26)$$

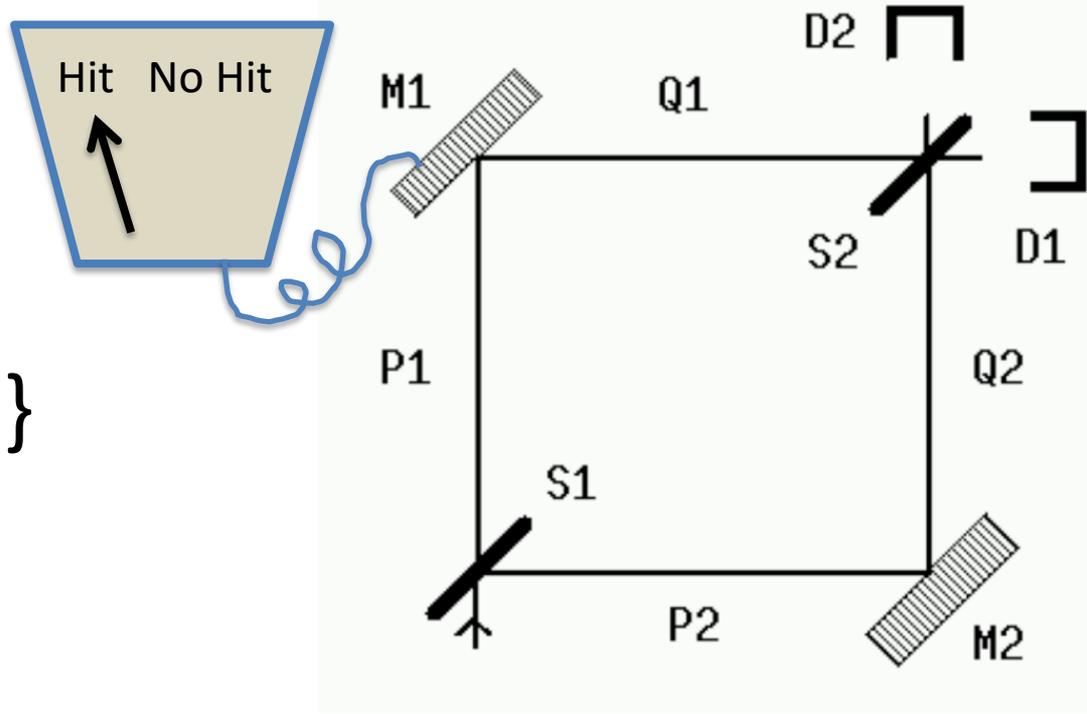
Attach a device
that measures
Whether M1 has
been hit or not.



$$(e \rightarrow D_j) = \sum_{n,m=1}^2 (e \rightarrow P_n)(e \wedge P_n \rightarrow Q_m)(e \wedge P_n \wedge Q_m \rightarrow D_j). \quad (25)$$

Since P_1 is equivalent to a point in X , previous knowledge is irrelevant and we have $(e \wedge P_n \rightarrow Q_m) = (P_n \rightarrow Q_m)$. We also clearly want to assume that the particle can't hop the rails, in other words we assume that $(P_n \rightarrow Q_m)$ is zero unless $n = m$. This causes one of the sums to disappear giving

$$(e \rightarrow D_j) = \sum_{n=1}^2 (e \rightarrow P_n)(P_n \rightarrow Q_n)(Q_n \rightarrow D_j) \quad (26)$$



State space:

$$\mathbf{R}^3 \times \{\text{Hit, No Hit}\}$$

Probability: \mathbf{C}

Sure enough, if you plug this in, the interference disappears. Also, if the device works so poorly that {Hit, No hit} is independent of whether P1 or P2 is taken, then the interference is restored.

But what about Hamiltonians and the Schrodinger equation etc.?

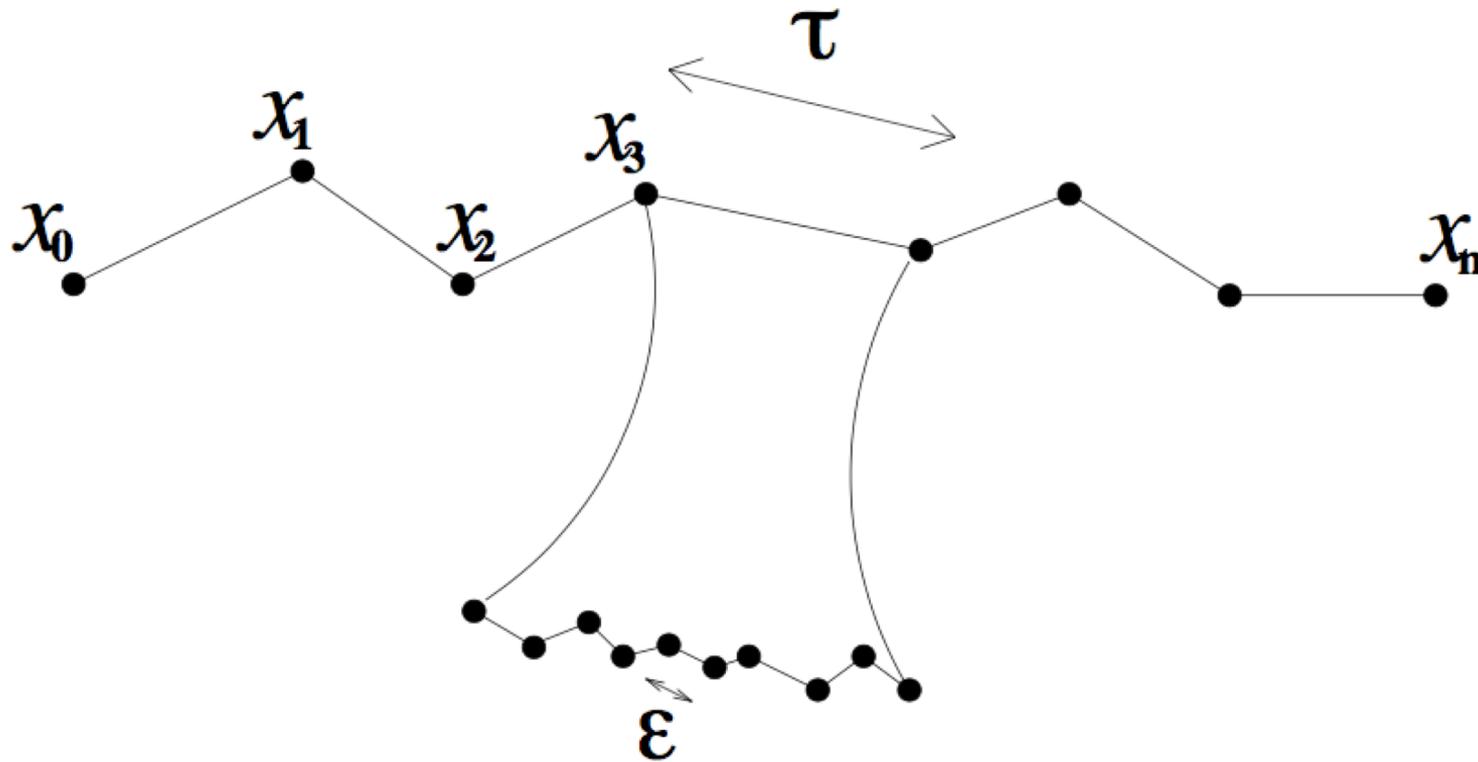


Figure 3. For $x_0, x_n \in U$, the complex probability $(x_0 \rightarrow x_n)$ is equal to a “path integral” over x_1, x_2, \dots, x_{n-1} with time step τ . The argument can be repeated making n sub-path integrals with time step ϵ .

S. Youssef, *Quantum Mechanics as Exotic Probability Theory*, proceedings of the Fifteenth International Workshop on Maximum Entropy and Bayesian Methods, K.M.Hanson and R.N. Silver, Santa Fe, 1995.

9 Time evolution

Given some initial knowledge such as A_t with $A \subset X$, the exotic probability to arrive at some $B \subset X$ at some later time t'' is given by

$$(A_t \rightarrow B_{t''}) = \int_{x \in X} (A_t \rightarrow x_{t'}) (x_{t'} \rightarrow B_{t''}) \quad (31)$$

for any time t' with $t \leq t' \leq t''$. This is called the Chapman–Kolmogorov equation in the probability literature. In the complex case with state space \mathbf{R}^d , one can either follow reference 4 or Risken[31] to conclude that for small $\tau \in \mathbf{R}$ and small $z \in X$, $(x_t \rightarrow (x+z)_{t+\tau})$ is given by

$$\frac{1}{(2\pi\tau)^{d/2} \sqrt{\det(\nu)}} \exp(-\tau [\frac{1}{2} (\frac{z_j}{\tau} - \nu_j) \nu_{jk}^{-1} (\frac{z_k}{\tau} - \nu_k) + \nu_o]) \quad (32)$$

where ν_o , ν_j and ν_{jk} are moments of the time derivative of $\omega(x, z, \tau) \equiv (x_t \rightarrow (x+z)_{t+\tau})$ defined by complex functions $\nu_o(x) \equiv \int_X \omega_\tau(x, z, 0)$, $\nu_j(x) \equiv \int_X \omega_\tau(x, z, 0) z_j$, $\nu_{jk}(x) \equiv \int_X \omega_\tau(x, z, 0) z_j z_k$. This is a central–limit–theorem–like phenomena where the details of the unknown function $(x_t \rightarrow (x+z)_{t+\tau})$ are smoothed over and only a dependence on it's lowest moments survives. Identifying z_j/τ as the velocity, equation 35 is equivalent, for example, to the Schrodinger equation in \mathbf{R}^3 identifying $\nu_o = -ieA_o$, $\nu_j = \frac{e}{m}A_j$ and $\nu_{jk} = (i/m)\delta_{jk}$. Similarly, quaternion probabilities in result in the Dirac equation[6, 7]. These arguments need to be made into proofs, but there is also a mystery as to why only parts of the available moments seem to be used by nature. Why, for instance, must ν_j be purely real in \mathbf{R}^3 ?

You don't have to choose the Hamiltonian in this approach. A Hamiltonian is effectively given to you depending on X , in the \mathbf{R}^4 case, complete with mass, \mathbf{A}_μ and $\mathbf{g}_{\mu\nu}$...

Can you guess what happens with probability \mathbf{H} instead of \mathbf{C} ?

....Yes, you get the Dirac equation.

S.K.Srinivasan, Quantum Mechanics via Extended Measures, J.Phys.A (23) 8297, (1997). See also http://physics.bu.edu/~youssef/quantum/quantum_refs.html for more.

Standard Quantum Theory

1. To define a theory, you must define a Hilbert space and a complete set of mutually commuting self-adjoint operators to serve as observables.
2. The state of a system is a ray in Hilbert space.
3. In addition, one must choose a Hamiltonian and label states by the irreducible representations of the Hamiltonian's symmetry group.
4. Time evolution is a one-parameter semigroup given by the Hamiltonian operator.

Modified Probability

1. To define a theory, choose a set X and choose probability \mathbf{R} , \mathbf{C} or \mathbf{H} .
2. The state of the system is a point in X .
3. Under suitable assumptions, dynamics is determined by the probabilities to go from x to y in X in a short time interval t .
4. Dynamics depends only on the moments of 3) in a manner similar to the central limit theorem. In the case of \mathbf{R}^4 , moments can be identified as particle mass, \mathbf{A}_μ and $\mathbf{g}_{\mu\nu}$

Our modified probabilities have the same status as probabilities do in Bayesian inference.

If you know **a**, then

$$(a \longrightarrow x)$$

is the modified probability that **x** is true given that **a** is known. Different people know different things and can have different wave functions.

If you happen to know more about the system (say, **M**), you just calculate

$$(a \wedge \mathbf{M} \longrightarrow x)$$

If you know more about a system, you get better results.
That makes sense, right?



Wave functions are not “the state of the system”

Standard Quantum Theory

1. To define a theory, you must define a Hilbert space and a complete set of mutually commuting self-adjoint operators to serve as observables.
2. The state of a system is a ray in Hilbert space.
3. In addition, one must choose a Hamiltonian and label states by the irreducible representations of the Hamiltonian's symmetry group.
4. Time evolution is a one-parameter semigroup given by the Hamiltonian operator.
5. If "mixed states" occur instead of "pure states," they must be described by density matrices.

Modified Probability

1. To define a theory, choose a set X and choose probability \mathbf{R} , \mathbf{C} or \mathbf{H} .
2. The state of the system is a point in X .
3. Under suitable assumptions, dynamics is determined by the probabilities to go from x to y in X in a short time interval t .
4. Dynamics depends only on the moments of 3) in a manner similar to the central limit theorem. In the case of \mathbf{R}^4 , moments can be identified as particle mass, \mathbf{A}_μ and $\mathbf{g}_{\mu\nu}$

What about “pure states” vs “mixed states”?

From our point of view, there is no such thing as “the system is in a pure state” or “the system is in a mixed state.” It just depends upon what you happen to know. Does this really work in detail?

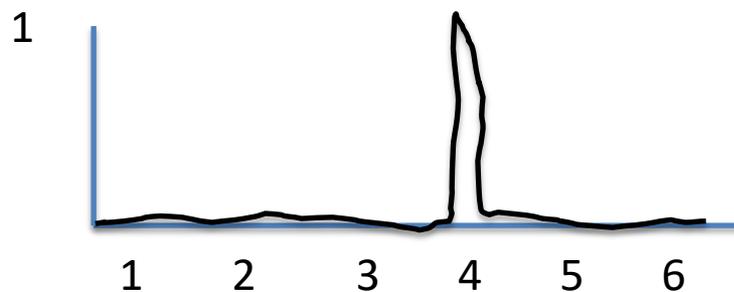
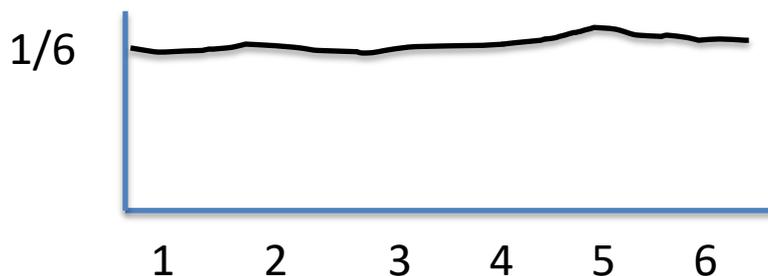
In conventional quantum mechanics, a sharp distinction is made between pure states, which can be described by a single wavefunction and statistical mixtures, which must, in general, be described by a density matrix. Since probability theory itself is no longer available to us, these “statistical mixtures” must be described entirely within complex probability theory. To investigate this issue, consider several situations which require density matrices in conventional theory. First, consider a system with initial knowledge e_o which is known to be well described by one of the wavefunctions ψ_1, ψ_2, \dots which may or may not be orthogonal. This would normally be represented as a mixture. As before, we have $(e_o \rightarrow x_o) = \sum_j (e_o \rightarrow b_j) \psi_j(x)$ where $b_j =$ “The system at $t = 0$ is best described by ψ_j .” Thus, in a quantum theory, not knowing which ψ_j best describes a system is no different from a pure superposition of ψ_j . To put it another way, all such expansions can be considered as mixtures with, in general, complex probabilities as coefficients and where a “statistical” mixture is only a special case. Density matrices are also needed in the case of “open systems”

S. Youssef, *Quantum Mechanics as Complex Probability Theory*, *Mod. Phys.Lett. A9*, 2571 (1994).

What about “wave function” collapse? The Observer problem, etc.?

In our theory, the “wave function” is just a function from X to P representing what you happen to know about a system.

Just as in Bayesian probability theory, it makes no sense to say that such a function is “the state of the system”. The state of the system is simply some unknown point in X .



Exponential Decay



Suppose that we have some system that can decay irreversibly to something else.

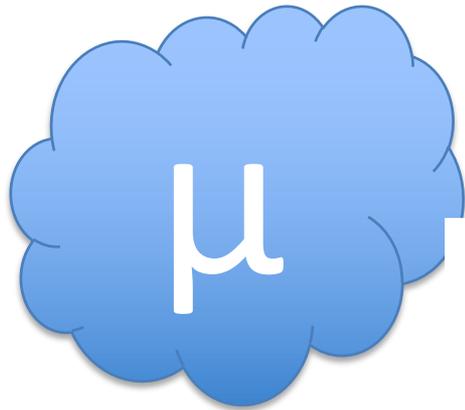
Suppose that the probability that this decay happens is independent of the past.

$$1 - e^{-\lambda t}$$

However, this does not follow in Quantum Theory. (See Sakurai's book, e.g.)

This is a lesser known paradox of quantum theory. Since quantum theory disagrees with this prediction in general, that means that such systems **aren't** independent of their past. Right? Or are they? Is a muon more likely to decay if it's old?

Exponential Decay



We avoid this paradox because the standard probability argument does not follow in our theory.

$\langle \mu_0 \rightarrow \mu_t \rangle = 1$, $\langle \alpha_0 \rightarrow \mu_t \rangle = a(1 - e^{-\lambda t})$, and $\langle \mu_0 \rightarrow \alpha_t \rangle = 0$ for some $\lambda \in P$ and $a \in \mathbf{R}$. Although the exotic probabilities are simple exponentials, this isn't preserved in the predicted frequencies. The ordinary probability to remain free for time t is

$$\text{Prob}(\alpha_t|\alpha_0) = \frac{\int_{\alpha} \|\alpha_0 \rightarrow x_t\|}{\int_{\alpha} \|\alpha_0 \rightarrow x_t\| + \int_{\beta} \|\alpha_0 \rightarrow x_t\|} \quad (27)$$

and, using $\int_{\alpha} \|\alpha_0 \rightarrow x_t\| = \|\alpha_0 \rightarrow \alpha_t\| \int_{\alpha} \|\alpha_t \rightarrow x_t\|$ and $\int_{\beta} \|\alpha_0 \rightarrow x_t\| = \|\alpha_0 \rightarrow \beta_t\| \int_{\beta} \|\alpha_0 \wedge \beta_t \rightarrow x_t\|$, we have

$$\text{Prob}(\alpha_t|\alpha_0) = \frac{1}{1 + k(t) \|e^{-\lambda t} - 1\|} \quad (28)$$

where

$$k(t) = a^2 \frac{\int_{\beta} \|\alpha_0 \wedge \beta_t \rightarrow x_t\|}{\int_{\alpha} \|\alpha_t \rightarrow x_t\|}. \quad (29)$$

For small t and assuming that λ is real and negative, $\text{Prob}(\alpha_t|\alpha_0)$ will decrease more slowly than $1 - 2\lambda t$. If we also know that α_0 and $x_t \in \beta_t$ can be taken to be independent for sufficiently large t , then we say that the system is “forgetful.” In this case, $k(t)$ is asymptotically constant and $\text{Prob}(\alpha_t|\alpha_0)$ will be exponential for large times. Such deviations from exponential decay have only recently been observed experimentally[30].

S. Youssef, *Quantum Mechanics as Complex Probability Theory, Mod. Phys.Lett. A9, 2571 (1994).*

But isn't this all impossible because of Bell's theorem?

In 1964, Bell analyzed a version of the Einstein-Podolsky-Rosen experiment.

Bell's results are almost always summarized this way:

No local realistic theory can give the predictions of Quantum Theory.

In the 1990s there were many, many, confirmations of this result.

-
- [6] J.S.Bell, *Physics*, 1 (1964) 195; J.S.Bell, *Rev.Mod.Phys.* 38 (1966) 447. See J.Bub, *Found.Phys.* 3 (1972) 29 for comments and N.D.Mermin, *Rev.Mod.Phys.* 65 (1993) 803 for a recent review.
 - [7] N.D.Mermin, *Am.J.Phys.* 49 (1981) 10.
 - [8] D.M.Greenberger, M.A.Horne, A.Shimony and A.Zeilinger, *Am.J.Phys.* 58 (1990) 1131.
 - [9] L.Hardy, *Phys.Rev.Lett.* 20 (1992) 2981.
 - [10] E.P.Wigner, *Am.J.Phys.* 38 (1969) 1005; J.F.Clauser and M.A.Horne, *Phys.Rev. D* 10 (1974) 526; P.H.Eberhard, *Nuovo Cimento* 38 B (1977) 75; J.D.Franson, *Phys.Rev.Lett.* 62 (1989) 2205; N.D.Mermin, *Phys.Rev.Lett.* 65 (1990) 3373; A.Peres, *Phys.Lett. A* 151 (1990) 107; L.Hardy, *Phys.Lett. A* 161 (1991) 21; I.Pitowsky, *Phys.Lett. A* 156 (1991) 137; A.C.Elitzur, S.Popescu and D.Rohrlich, *Phys.Lett. A* 162 (1992) 25; L.Hardy, *Phys.Lett. A* 167 (1992) 17; L.Hardy and E.J.Squires, *Phys.Lett. A* 168 (1992) 169; A.Mann, K.Nakamura and M.Revzen, *J.Phys. A* 25 (1992) L851; C.Pagonis and R.Clifton, *Phys.Lett. A* 168 (1992) 100; M.Vinduska, *Found.Phys.* 22 (1992) 343; A.Elby and M.R.Jones, *Phys.Lett. A* 171 (1992) 11; M.Ardehali, *Phys.Lett. A* 181 (1993) 187; M.Ardehali, *Phys.Rev. A* 47 (1993) 1633; H.J.Bernstein, D.M.Greenberger, M.A.Horne and A.Zeilinger, *Phys.Rev. A* 47 (1993) 78; P.Busch, P.Kienzler, P.Lahti and P.Mittelstaedt, *Phys.Rev. A* 47 (1993) 4627; D.N.Klyshko, *Phys.Lett. A* 172 (1993) 399; H.P.Stapp, *Phys.Rev. A* 47 (1993) 847; B.Yurke and D.Stoler, *Phys.Rev. A* 47 (1993) 1704; M.Czachor, *Phys.Rev. A* 49 (1994) 2231; L.Hardy, *Phys.Rev.Lett.* 73 (1994) 2279.



All conclude the same thing: Local realistic theories are impossible.

But wait...

Bell's result is an argument in standard probability theory.

It does not follow for us.

In Bell's analysis[6], two spin $\frac{1}{2}$ particles in a singlet state are emitted towards two distant Stern–Gerlach magnets. Let e_t define the known orientations of the two magnets and the description of the initial singlet state and let $M_{t''}$ be a description of one of the possible results of the final measurements. Let t be the time when the singlet state is created, t'' be the time of the final measurement and let $t < t' < t''$. Bell's theorem is an argument in probability theory beginning with an expansion in “hidden variable” λ in state space U : $P(e_t, M_{t''}) = P(e_t, U_{t'} \wedge M_{t''})$ and so

$$P(e_t, M_{t''}) = \int_{\lambda \in U} P(e_t, \lambda_{t'} \wedge M_{t''}) = \int_{\lambda \in U} P(e_t, \lambda_{t'}) P(e_t \wedge \lambda_{t'}, M_{t''}). \quad (6)$$

However, since ordinary probability theory has been abandoned, equation (6) must be justified within complex probability theory. But if the definition of *Prob* is extended to mixed times

Bell has exactly shown that if you want a local realistic theory, you have to abandon probability theory.

What about those 25 confirmations of Bell's results?

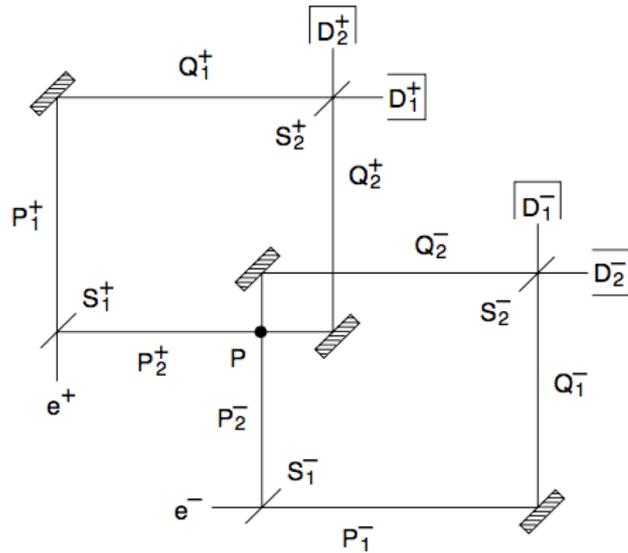


Figure 2. Two overlapping interferometers as described in reference 11. Electrons and positrons first encounter a beam splitter (S_1^\pm), then one of two mirrors, a second beam splitter (S_2^\pm) and are then detected in one of D_j^\pm . The electron and positron are assumed to annihilate with probability one if they meet at point P .

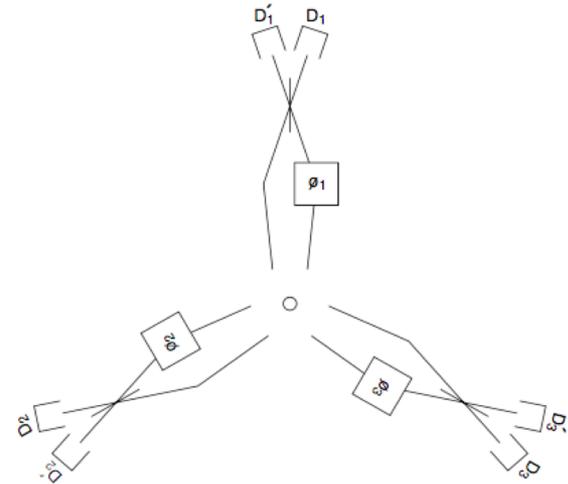


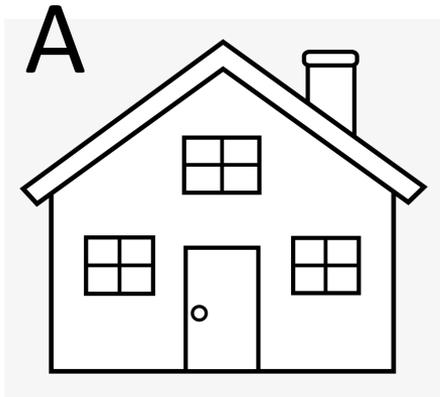
Figure 1. The three armed interferometer proposed in reference 8 consists of a central source which emits three particles in the plane of the paper with zero total momentum. Each particle is brought to a beam splitter and is detected in either a primed or unprimed detector (D_j' and D_j respectively). For each arm of the interferometer, one of its paths induces an adjustable phase shift ϕ_1, ϕ_2 or ϕ_3 as indicated.

These also fail to rule out local realistic theories. They do not rule out modified probabilities as we have been discussing.

S. Youssef, Is Complex Probability Theory Consistent with Bell's Theorem?, Phys. Lett. A 204, 181 (1995).

What about non-local effects?

Doesn't Bell show that Quantum Theory is non-local?



Cut a penny so that there is a heads $\frac{1}{2}$ and a tails $\frac{1}{2}$. Secretly mail one half to house A and the other half to house B. When we open the envelope in house A, does this cause a non-local effect at house B?

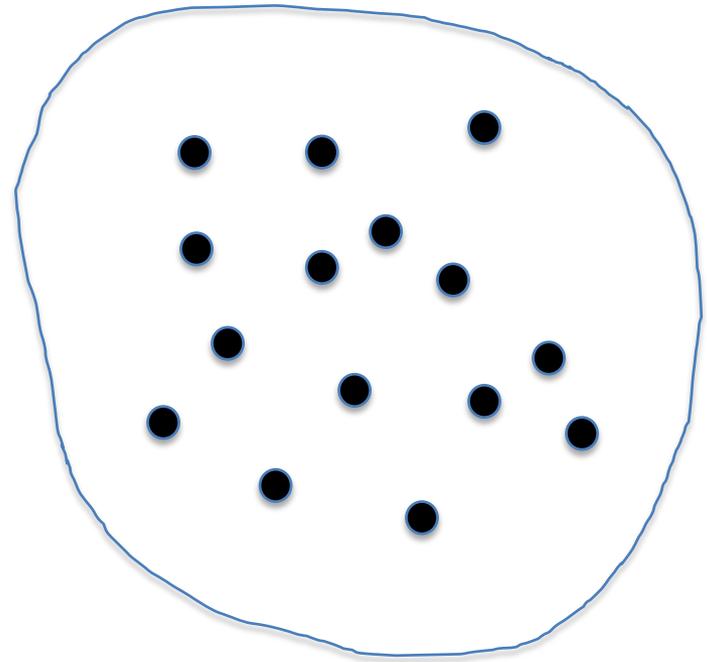
Answer: **No**. What's called "non-local effects" in quantum theory are just correlations in modified probability theory.

S. Youssef, Quantum Mechanics as Complex Probability Theory, Mod. Phys.Lett. A9, 2571 (1994).

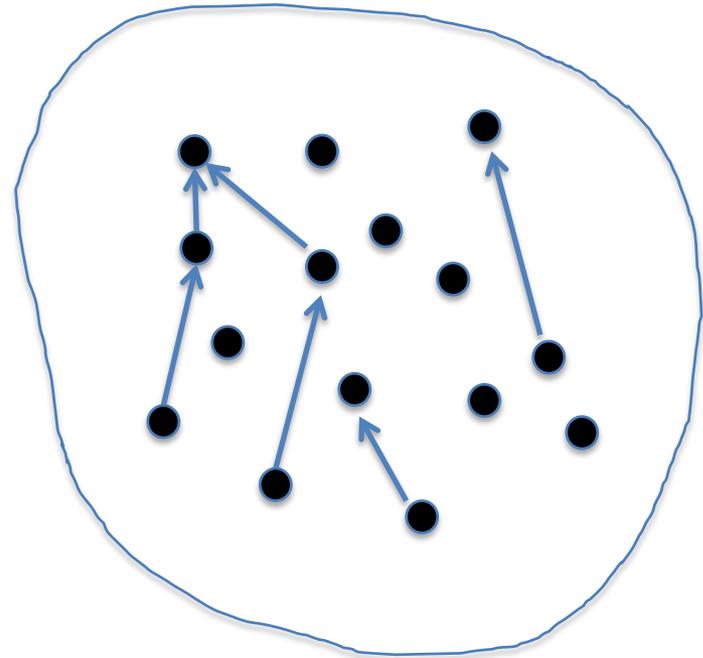
What does this all mean?



What could the ultimate answer be?



What could the ultimate answer be?



In the case of a die, it would be silly to say that $\{1,2,3,4,5,6\}$ is the true ultimate state space of a real physical die, but that is only because we know that dies are made of ivory, have physical dots, move in gravity, etc. On the other hand, if we propose an ultimate state space, we can no longer ask any more questions and can never know what those points “really are.”

In this scenario, you just have propositions and their (necessarily) simple relationship to each other. Because we have assumed that this is all that can be known, these simple things must, therefore, determine the Standard Model, Gravity and all the vast, detailed, seemingly highly specific world that we know.

