A Fundamental Correspondence Between Stochastic Processes and Quantum Systems

arXiv:2302.10778 / arXiv:2309.03085 / arXiv:2402.16935
philsci:22501 / philsci:22502 / philsci:23151

Jacob Barandes
Harvard University

Mathematical Physics Seminar
Regensburg University
January 31, 2025

Outline of this Talk

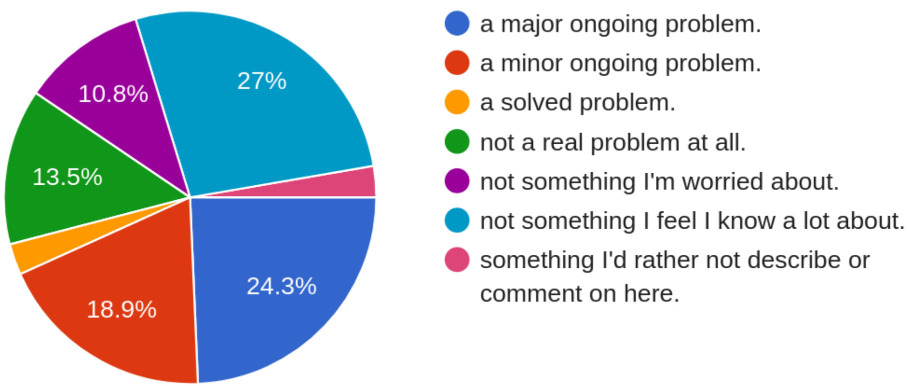
1. Introduction
2. Indivisible Stochastic Processes
3. The Stochastic-Quantum Correspondence
4. Causal Locality
5. Concluding Remarks

1. Introduction

2024 Survey of New Harvard Physics PhD Students

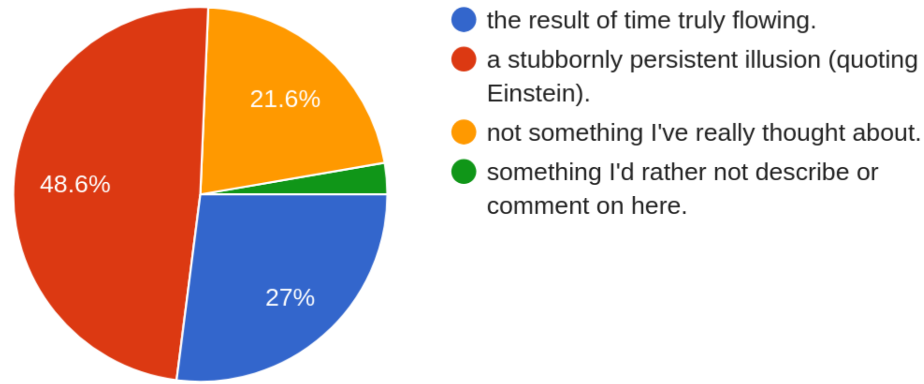
The measurement problem in quantum theory is ...

37 responses



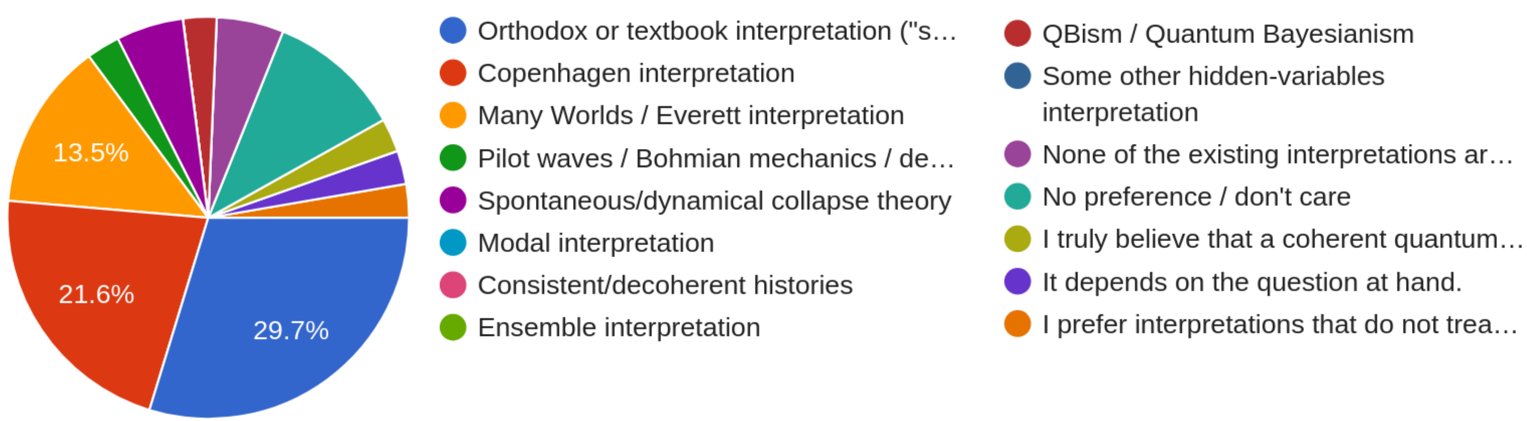
The distinction between the past, present and future is ...

37 responses



What's your preferred interpretation of quantum theory?

37 responses



Why Another Interpretation/Formulation of QM?

Don't we have too many as it is?

We don't have any that meet all of the following minimal consisteny requirements:

• Empirical adequacy (Bohmiam mechanics doesn't generalize, Everett/many worlds cannot produce empirical probabilities)

• Unambiguous predictions for macroscopic systems when treated quantum-mechanically (Dirac-von Neumann)

• Able to account at least schematically for the emergence of macroscopic systems and the classical limit (Dirac-von Neumann, Copenhagen)

• Avoidance of too many extra-empirical assumptions and speculative metaphysical hypotheses (SMHs) (Everett/many worlds)

The Wave-Function Paradigm

In most textbook treatments of quantum theory, one takes the basic object to be a wave function $|\Psi(t)\rangle$ in a Hilbert space $\mathcal{H}$ over the complex numbers $\mathbb{C}$, axiomatically evolving according to the Schrödinger equation for some self-adjoint Hamiltonian $H(t)$:

$$i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = H(t)|\Psi(t)\rangle$$

(Sometimes one instead takes the basic object to be a density matrix $\rho(t)$ on the Hilbert space, or a linear map ω on a C*-algebra)

A host of other axioms provide an instrumentalist algorithm for using $|\Psi(t)\rangle$ to calculate probabilities of measurement outcomes

Much of the interpretative debate lies within this wave-function paradigm, and asks about the nature of $|\Psi(t)\rangle$ (epistemic? ontological? nomological? something else?), and whether it should be supplemented with additional (or 'hidden') variables

Historical Origins of the Wave-Function Paradigm

The 1925 matrix mechanics of Heisenberg, Born, and Jordan took the ontology out of quantum theory

Schrödinger's 1926 undulatory mechanics replaced that missing ontology with his wave functions

Copenhagenists, QBists, and (perhaps) C*-algebraists don't agree that wave functions are part of the ontology

---

But most people operate within the wave-function paradigm – the notion that quantum theory is *about* wave functions (or density matrices) – and build everything on top of that

So now one hears talk of whether the wave function 'is complete', or if there are hidden variables in addition to wave functions, or if one is a 'psi-epistemicist' or a 'psi-ontologist' (in the language of Harrigan, Spekkens, 2010)

Goals of this Talk

The main goal of this talk is to argue that every quantum system can be understood as an 'indivisible' stochastic process in disguise, with no fundamental role for wave functions or density matrices

---

Additional goals:

• Define what an indivisible stochastic process is

• Give a correspondence between indivisible stochastic processes and quantum theory (applications for stochastic modeling?)

• Describe some concrete examples, including classical analogue models of quantum systems that are hopefully realizable in the lab or in simulations

• Demystify and deflate the exotic features of quantum systems (interference, entanglement, decoherence, measurements, etc.)

• Introduce a new nomological (lawlike) definition of causal influences for quantum systems that's local

The Textbook Axioms of Quantum Theory

(1) Each state of a quantum system is represented by a unit-norm state vector/wave function $|\Psi\rangle$, or a positive-semidefinite and unit-trace density operator/density matrix ρ, in a Hilbert space $\mathcal{H}$

(2) A closed quantum system evolves according to a unitary time-evolution operator $U(t) = U^{-1\dagger}(t)$, or the Schrödinger equation for some self-adjoint Hamiltonian $H(t) = H^{\dagger}(t)$

The measurement axioms

(3) Each observable is represented by a self-adjoint operator $A(t) = A^{\dagger}(t)$ whose eigenvalues are the possible numerical meaurement outcomes (more generally: POVMs)

(4) The Born rule $\langle\Psi|A|\Psi\rangle$ or $\mathrm{tr}(A\rho)$ gives the statistical average of measurement outcomes over measurement-outcome probabilities

(5) Immediately after a measurement, the quantum state collapses to single out a unique measurement outcome

Note: decoherence doesn't single out a unique measurement outcome!

The Wigner's Friend Paradox

- $\mathcal{W}$ ("Wigner") remains outside a perfectly sealed box
- $\mathcal{F}$ ("Wigner's Friend") is inside the box and does a measurement on a quantum system also inside the box

Question: Do we activate the collapse axiom (5) or not?

Measurements are a very narrow category – how do we account for phenomena happening more generally? ("category problem")

We have essentially four options for resolving this ambiguity:

(A) Invoke the collapse axiom (5), but then we need a rigorous definition of a measurement (the "measurement problem")

(B) Avoid invoking the collapse axiom (5) for $\mathcal{W}$ while assuming a unique outcome, but then the quantum state is manifestly incomplete ($\mathcal{F}$'s outcome is then literally a "hidden variable")

(C) Replace the collapse axiom (5) with something else (e.g., spontaneous dynamical collapse) (need new parameters!)

Focus of this talk

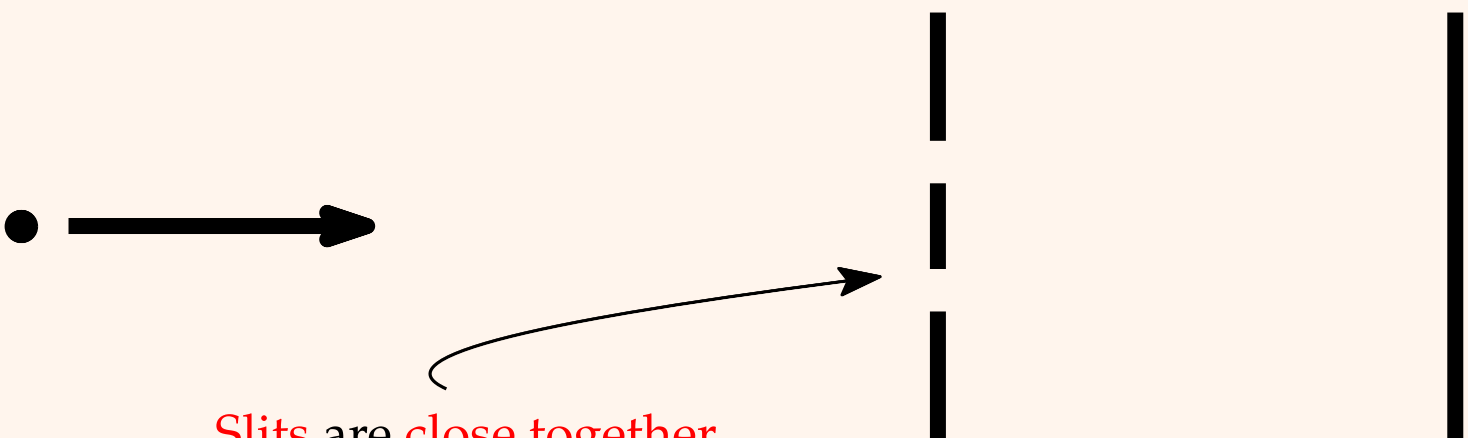
(D) Don't assume a unique outcome ("Everett/many-worlds interpretation"), but then what does probability mean (etc.)?

Revisiting the Double-Slit Experiment

This new formulation isn't limited to any specific kind of quantum system, but a well-known single-particle thought experiment will provide a good example

In the double-slit experiment, one imagines sending particles, one at a time, toward a wall with two slits, and then observing where the particle arrives on a screen

slyly assumed so that we have a 3D *configuration* space that resembles *physical* 3D space—otherwise we'd need a $3N$-dimensional space! (Where are the two slits supposed to be? And how is this intuitive?)



Slits are close together

A = starting conditions B = which slit C = where it lands

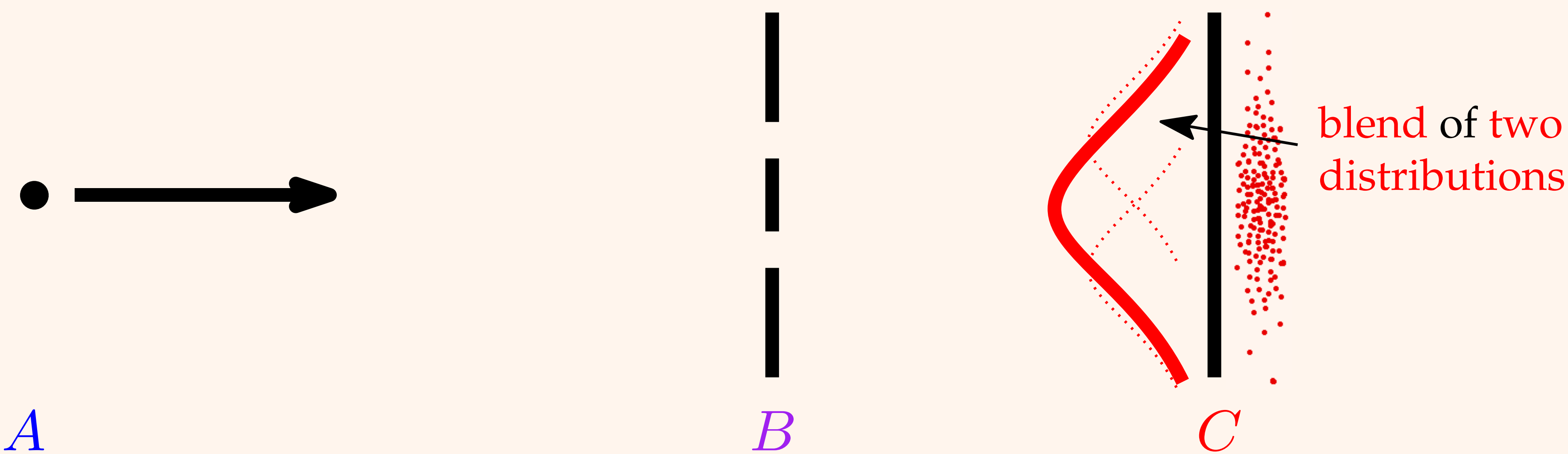
Double-Slit Experiment (cont.)

Usual classical assumption: $p(C|A) = \sum_B p(C|B)p(B|A)$

note: doesn't follow from general marginalization rules!

$\Longrightarrow$ requires a Markov or divisibility assumption at B

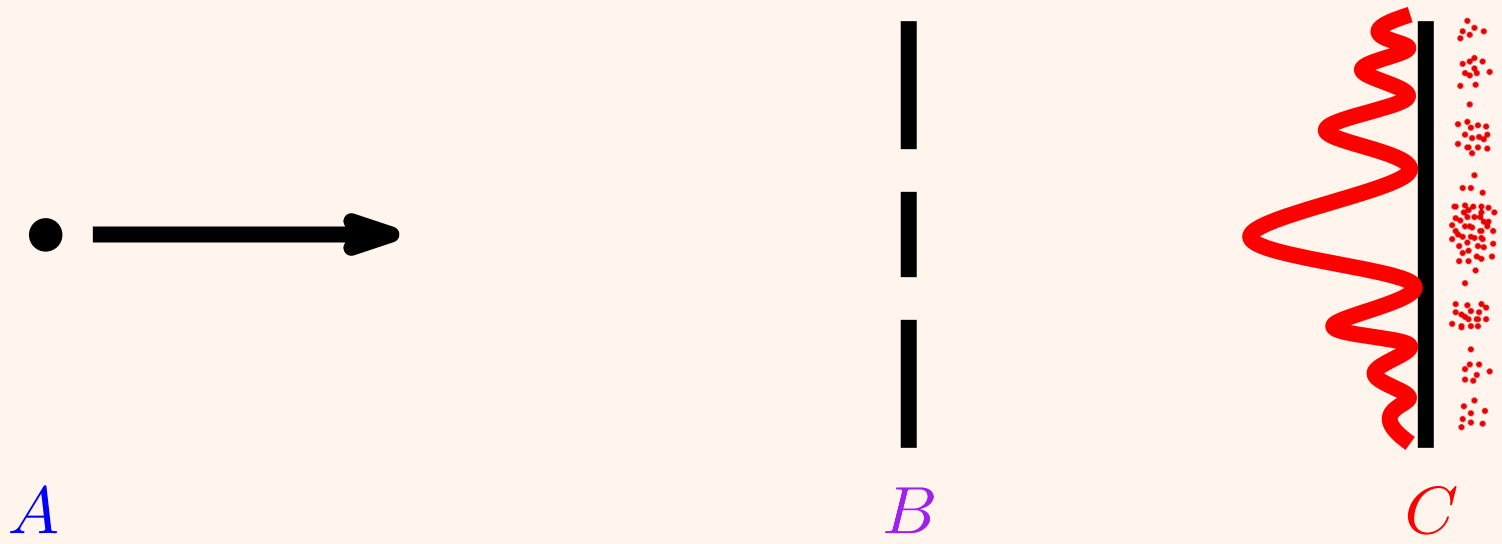
This predicts the following pattern over many repetitions:



This matches observations for macroscopic particles, like stones

Double-Slit Experiment (cont.)

For electrons over many repetitions, one instead observes what looks like an interference pattern:



Does this mean that each particle is really a 'Schrödinger wave,' or that each particle somehow 'goes through both slits'? *No!*

We'll show that one can account for this pattern merely by allowing the dynamics to be non-Markovian or indivisible:

$$p(C|A) \neq \sum_B p(C|B)p(B|A)$$

2. Indivisible Stochastic Processes

Indivisible Stochastic Processes

An indivisible stochastic process is a simple generalization of a non-Markovian stochastic process

The new axioms (simpler!)

Fixed ingredients (i.e., the same in every run):

division events

• Kinematics: Configuration space $\mathcal{C}$ (i.e., the sample space)

• Dynamics: For times t, t_0 in respective index sets $\mathcal{T}, \mathcal{T}_0$, a 1st-order conditional-probability map $\Gamma(t \leftarrow t_0) : \mathcal{C} \times \mathcal{C} \to [0, 1]$

Contingent ingredients (i.e., can differ between runs)

• For times $t \in \mathcal{T}$, a probability distribution $p(t) : \mathcal{C} \to [0, 1]$

Linear marginalization rule: $\quad p_i(t) = \sum_j \Gamma_{ij}(t \leftarrow t_0) p_j(t_0)$

important for later!

Indivisibility

Indivisibility is a simple and remarkably new idea: 'failure of iterativeness'

Originated in the theory of quantum channels [Wolf, Cirac, 2008]

First applied to classical stochastic processes only a few years ago! [Milz, Modi, 2021]

$$\text{For } t > t' > t_0: \qquad \Gamma(t \leftarrow t_0) \neq \Gamma(t \leftarrow t')\Gamma(t' \leftarrow t_0)$$

More precisely: No such $\Gamma(t \leftarrow t')$ generically exists

---

Unlike for a textbook non-Markovian stochastic process, no higher-order conditional probabilities are specified by the model

Generically non-Markovian of "infinite order"

Ex: Interpolation of a Discrete Determinstic Process

Consider the simplest kind of discrete-time deterministic process

Configurations $1, ..., N$ that deterministically transition in time steps δt by a permutation (not assumed time-reversal invariant)

E.g., $1 \overset{\delta t}{\mapsto} 7 \mapsto 3 \mapsto 2 \mapsto 13 \mapsto 5 \mapsto \cdots \mapsto 4 \mapsto 1 \mapsto 7 \mapsto 3 \mapsto \cdots$

Can use an N-dimensional vector space and represent each configuration $1, \ldots, N$ with a member of the standard basis:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_N = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Then represent the dynamical law as a permutation matrix Σ:

$$e_1 \overset{\delta t}{\mapsto} \Sigma e_1 = e_7 \mapsto \Sigma^2 e_1 = e_3 \mapsto \Sigma^3 e_1 = e_2 \mapsto \Sigma^4 e_1 = e_{13} \mapsto \cdots$$

Unitary and Unistochastic Representations

Σ is an $N \times N$ permutation matrix

$\implies \Sigma$ is an $N \times N$ unitary matrix

($\Sigma^N = 1$ implies that the eigenvalues are the Nth roots of unity)

$\implies$ For t any smoothly variable time, $U(t \leftarrow 0) \equiv \Sigma^{t/\delta t}$ exists and is still unitary

$\implies \Gamma_{ij}(t \leftarrow 0) \equiv \left|\Sigma_{ij}^{t/\delta t}\right|^2$ gives a unistochastic matrix and defines an indivisible stochastic process

Analytically interpolates the original discrete-time deterministic process to a smooth-in-time indivisible stochastic process, with

$$\Gamma(n\,\delta t \leftarrow 0) = \Sigma^n \text{ for } n \in \mathbb{Z}$$

Divisibility and Indivisibility

Using the unitary matrix $U(t \leftarrow 0)$, can define for any pair of times t, t':

$$U(t \leftarrow t') \equiv U(t \leftarrow 0)U^{\dagger}(t' \leftarrow 0)$$

Convenient composition law: $\quad U(t \leftarrow 0) = U(t \leftarrow t')U(t' \leftarrow 0)$

But the indivisible stochastic process will fail to have such a composition law for most times:

$$\Gamma(t \leftarrow 0) \neq \Gamma(t \leftarrow t')\Gamma(t' \leftarrow 0)$$

If one *tries* to define $\Gamma_{ii'}(t \leftarrow t') \equiv |U_{ii'}(t \leftarrow t')|^2$ anyway, then:

$$\Gamma(t \leftarrow 0) - \Gamma(t \leftarrow t')\Gamma(t' \leftarrow 0) = (\text{interference terms!}) \neq 0$$

However, the stochastic process *will* divide at the special integer-step times $t' = n\,\delta t \implies$ division events

An Emergent Quantum Theory

Notice: $U(t \leftarrow 0)$ is a smooth function of $t \implies$ we can define a Hamiltonian:

$$H(t) \equiv i\hbar \frac{\partial U(t \leftarrow 0)}{\partial t} U^\dagger(t \leftarrow 0) = H^\dagger(t)$$

Define a complex-valued, time-evolving state vector or wave function:

$$|\Psi(t)\rangle \equiv U(t \leftarrow 0)e_1 = U(t \leftarrow 0)\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

known initial condition
(no initial epistemic uncertainty)

Then we have the Schrödinger equation:

$$i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = H(t)|\Psi(t)\rangle$$

And the Born rule:

$$p_i(t) = |\Psi_i(t)|^2$$

Handling Initial Epistemic Uncertainty

More generally, if we don't know the initial condition with certainty, then we can encode initial epistemic probabilities into a diagonal density matrix:

$$\rho(0) \equiv \mathrm{diag}\,(p_1(0), \ldots, p_N(0))$$

Define a time-evolving, non-diagonal density matrix:

$$\rho(t) \equiv U(t \leftarrow 0)\rho(0)U^\dagger(t \leftarrow 0)$$

Then we have the von Neumann equation:

$$i\hbar \frac{\partial \rho(t)}{\partial t} = [H(t), \rho(t)]$$

And the Born rule becomes

$$p_i(t) = \mathrm{tr}(P_i \rho(t))$$

with $P_i \equiv e_i e_i^\dagger$ an elementary projection matrix

Ex: Classical Analogue Model for a Qubit

Consider a black box containing either of two possible classical systems, $\mathcal{A}$ or $\mathcal{B}$, each with two similar-looking configurations 1 and 2

For $\mathcal{A}$: $E_1 < E_2$, whereas for $\mathcal{B}$: $E_2 < E_1$, but with the same energy difference

Thermally couple to a reservoir with gradually changing temperature $T(t)$ that we can control, with $T(0) = 0$

There are then four Boltzmann occupation probabilities $p(t) \propto \exp(-E/kT(t))$:

$$\Gamma(t \leftarrow 0) \equiv \begin{pmatrix} p(1,t|1,0) & p(1,t|2,0) \\ p(2,t|1,0) & p(2,t|2,0) \end{pmatrix}$$

By symmetry: $p(1,t|1,0) = p(2,t|2,0)$, $\quad p(1,t|2,0) = p(1,t|2,0)$

Another Emergent Quantum Theory

So $\Gamma(t \leftarrow 0)$ is a 2×2 doubly stochastic $\implies$ unistochastic

That is, there exists a 2×2 unitary matrix $U(t \leftarrow 0)$ such that $\Gamma_{ij}(t \leftarrow 0) = |U_{ij}(t \leftarrow 0)|^2 \implies$ Schrödinger equation, etc.

The original stochastic description is generically indivisible (e.g., take $p(1, t|1, 0) = \exp(-t^2/\tau^2)$ for some constant time scale $\tau > 0$)

But the unitary description has a nice composition property:

$$U(t \leftarrow 0) = U(t \leftarrow t')U(t' \leftarrow 0)$$

So working with the unitary description gives us a convenient 'divisible' formalism (*that's* what Hilbert spaces are *for*!)

Cost: density matrix $\rho(t)$ with nonzero off-diagonal entries (coherences) $\implies$ artifacts of the indivisibility/non-Markovianity

Amplitudes and Interference

Consider the following two amplitudes that share the same initial and final conditions but differ at an intermediate time t':

amplitudes

$$\mathrm{path}(1) = \langle 2|U(t \leftarrow t')|1\rangle\langle 1|U(t' \leftarrow 0)|1\rangle$$

$$\mathrm{path}(2) = \langle 2|U(t \leftarrow t')|2\rangle\langle 2|U(t' \leftarrow 0)|1\rangle$$

It is easy to show that:

$$|\mathrm{path}(1) + \mathrm{path}(2)|^2 = p(2, t|1, 0)$$

But:

$$|\mathrm{path}(1)|^2 + |\mathrm{path}(2)|^2 \neq p(2, t|1, 0)$$

This is precisely *probabilistic* interference, in a classical analogue model of a qubit (note: no hackneyed analogy here with electromagnetic interference!)

This analogue model can be generalized to two or more qubits, to demonstrate entanglement

3. The Stochastic-Quantum Correspondence

The Stochastic-to-Quantum Direction

Given any indivisible stochastic process with N configurations, introduce a (not unique) complex $N \times N$ matrix $\Theta(t \leftarrow 0)$ according to:

$$\Gamma_{ij}(t \leftarrow 0) = |\Theta_{ij}(t \leftarrow 0)|^2$$

This new matrix satisfies the sum rule

$$\sum_i |\Theta_{ij}(t \leftarrow 0)|^2 = 1$$

If $\Gamma(t \leftarrow 0)$ is unistochastic, then $\Theta(t \leftarrow 0)$ can be assumed to be a unitary matrix $U(t \leftarrow 0)$

If not, then place each column of $\Theta(t \leftarrow 0)$ into an empty $N \times N$ matrix $K_\beta(t \leftarrow 0)$, where $\beta = 1, ..., N$

These are Kraus operators $\implies$ Stinespring-dilate to a unitary!

$\implies$ Emergent quantum system in a Hilbert-space representation!

Linear marginalization rule $\implies$ *Linear* time evolution!

Unistochastic Matrices and the Complex Numbers

For $N > 2$, an $N \times N$ unistochastic matrix will not generally be orthostochastic (i.e., based on a real orthogonal matrix)

Hence, to exploit the stochastic-quantum theorem and unitary evolution, the complex numbers (or an algebraic construct isomorphic to them) will be necessary!

Hilbert spaces are fictions anyway, and the complex numbers also let us invoke the spectral theorem, symmetry generators, Hamiltonians, energy eigenvalues, stationary states, the Schrödinger equation, the uncertainty principle, spinors, etc.

Actually, one also needs the complex-conjugation operator K (needed for time-reversal transformations), which satisfies:

$$K^2 = 1, \quad Ki = -iK$$

Then i, K, and iK generate a Clifford algebra called the pseudo-quaternions [Stueckelberg, 1960]

Wave Functions and the Schrödinger Equation

If the density matrix is rank-one, then there exists an $N \times 1$ state vector $|\Psi(t)\rangle$ that gives a simple factorization:

$$\rho(t) = |\Psi(t)\rangle\langle\Psi(t)| \qquad \text{(i.e., } \Psi(t)\Psi^\dagger(t)\text{)}$$

The state vector then satisfies the Schrödinger equation:

$$i\hbar\frac{\partial|\Psi(t)\rangle}{\partial t} = H(t)|\Psi(t)\rangle \longleftarrow$$

The linearity ultimately descends from the linear maginalization rule!

So wave functions and the Schrödinger equation are secondary pieces of derived mathematics, not primary ontological furniture

We therefore see that wave functions are not purely epistemic, but encode a blend of epistemic and nomological information, and are not physical or ontological objects (like Magritte's pipe, or aether!)

And the standard unitary Hilbert-space formalism formally yields a 'divisible', 1st-order differential equation for the dynamics!

The Quantum-to-Stochastic Direction

Given any unitarily evolving quantum system with an N-dimensional Hilbert space, pick a convenient orthonormal basis, and then define an indivisible stochastic process via the unistochastic matrix

which we also do to define a path integral!

$$\Gamma_{ij}(t \leftarrow 0) \equiv |U_{ij}(t \leftarrow 0)|^2$$

Are we losing phase information here? It actually doesn't matter!

Remember: all empirical results come from measurement processes, and these should now be modeled explicitly using measuring devices

If the measuring device is properly regarded as a subsystem of the overall indivisible stochastic process, and the chosen orthonormal basis captures the device's pointer variables, then, by construction, the indivisible stochastic process will produce the correct final measurement-outcome probabilities!

Division Events

We now have a framework that allows us to explain on theoretical grounds why the well-known Markov approximation (irrelevance of past states) often works so well in applications

Consider a composite system $\mathcal{S}\mathcal{E}$ ($\mathcal{S}$ubject + $\mathcal{E}$nvironment)

Suppose that for each configuration i of the subject system, the environment has a corresponding configuration $e(i)$

Suppose that the overall transition matrix $\Gamma^{\mathcal{S}\mathcal{E}}(t \leftarrow 0)$ yields

$$p^{\mathcal{S}\mathcal{E}}_{i' e'}(t') = p^{\mathcal{S}}_{i'}(t')\delta_{e' e(i')}$$ (classical correlation)

environment configuration that depends on subject configuration

Division Events (cont.)

Then from *classical* marginalization over the environment at $t > t'$, one can show that

$$p_i^{\mathcal{S}}(t) = \sum_e p_{ie}^{\mathcal{S}\mathcal{E}}(t) = \sum_{i'} \Gamma_{ii'}^{\mathcal{S}}(t \leftarrow t') p_{i'}^{\mathcal{S}}(t')$$

Hence:

$$\Gamma^{\mathcal{S}}(t \leftarrow 0) = \Gamma^{\mathcal{S}}(t \leftarrow t')\Gamma^{\mathcal{S}}(t' \leftarrow 0)$$

That is, due to the correlating interaction with the environment, there is automatically a new 'division event' at t' playing the role of $t = 0$

Division events are ubiquitous for open systems in noisy environments, thereby explaining why the Markov approximation often works so well on macroscopic scales

Decoherence

It is easy to show that at t', the subject system's (reduced) density matrix becomes momentarily diagonal

$$\rho(t') = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \\ \rho_{21} & \rho_{22} & \cdots & \\ \vdots & \vdots & \ddots & \\ & & & \rho_{NN} \end{pmatrix} \to \begin{pmatrix} \rho_{11} & 0 & \cdots & \\ 0 & \rho_{22} & \cdots & \\ \vdots & \vdots & \ddots & \\ & & & \rho_{NN} \end{pmatrix}$$

This is decoherence!

So we learn that the off-diagonal entries in a density matrix (coherences, corresponding to superpositions in wave functions) are merely a mathematical artifact of indivisible dynamics

Coherences are the price for making the dynamics look divisible!

Meanwhile, decoherence itself is just what the prosaic leakage of correlations into the environment looks like when seen through the lens of the Hilbert-space formulation

Entanglement

To start, note that even in classical-deterministic physics, during an interaction, systems have non-factorizing dynamics

Given two subsystems $\mathcal{A}, \mathcal{B}$, if they are not interacting with each other from $t = 0$ up to just before $t' > 0$, then the composite system's transition matrix tensor-factorizes:

$$\Gamma^{\mathcal{AB}}(t \leftarrow 0) = \Gamma^{\mathcal{A}}(t \leftarrow 0) \otimes \Gamma^{\mathcal{B}}(t \leftarrow 0) \quad (\text{for } t < t')$$

However, a transition matrix encodes *cumulative* statistical information, so for all $t > t'$ until a division event, the composite transition matrix fails to tensor-factorize:

$$\Gamma^{\mathcal{AB}}(t \leftarrow 0) \neq \Gamma^{\mathcal{A}}(t \leftarrow 0) \otimes \Gamma^{\mathcal{B}}(t \leftarrow 0) \quad (\text{for } t > t')$$

That is, the theory just doesn't contain or supply laws for the subsystems $\mathcal{A}$ and $\mathcal{B}$ separately

Entanglement (cont.)

This already looks like entanglement, but seen entirely from the stochastic side of the stochastic-quantum correspondence

If there is a division event (e.g., by the environment) at some later time $t'' > t'$, then the composite system's transition matrix divides starting at t'':

$$\Gamma^{\mathcal{AB}}(t \leftarrow 0) = \Gamma^{\mathcal{AB}}(t \leftarrow t'')\Gamma^{\mathcal{AB}}(t'' \leftarrow 0) \quad \text{(for } t > t'' > t'\text{)}$$

If the subsystems are no longer interacting after t' or t'', then the relative transition matrix tensor-factorizes:

$$\Gamma^{\mathcal{AB}}(t \leftarrow t'') = \Gamma^{\mathcal{A}}(t \leftarrow t'') \otimes \Gamma^{\mathcal{B}}(t \leftarrow t'') \quad \text{(for } t > t'' > t'\text{)}$$

So decoherence causes a 'breakdown' in entanglement, as expected, and notice that we haven't used Hilbert spaces here!

$\{$Observables$\}$ = $\{$Beables$\}$ $\cup$ $\{$Emergeables$\}$

*Be*ables are just random variables on the configuration space $\mathcal{C}$

Use a diagonal matrix: $A(t) \equiv \operatorname{diag}(a_1(t), \ldots, a_N(t))$

Expectation values: $\langle A(t) \rangle = \operatorname{tr}(A(t)\rho(t))$

If $A(t) = P_i \equiv \operatorname{diag}(0, \ldots, 1, \ldots, 0)$ is an elementary projector and $\rho(t) = \Psi(t)\Psi^\dagger(t)$ is rank-one, one obtains the Born rule:

$$p_i(t) = |\Psi_i(t)|^2$$

By modeling the measurement process as an overall unistochastic process, one sees patterns in the dynamics that look just like beables to measuring devices, treated as stochastic systems as well

These "*emerge*ables" are represented by non-diagonal self-adjoint matrices, and together with the beables constitute the system's noncommutative algebra of observables, thereby completing the textbook axioms (see the papers for the detailed calculations!)

The Stochastic-Quantum Correspondence

So we arrive at a *stochastic-quantum correspondence*, according to which the Hilbert-space formalism serves as a form of 'analytical mechanics' for stochastic systems, giving rise to an effective 1st-order differential equation

---

This correspondence is many-to-one in both directions

This is like how classical mechanical systems based on 2nd-order differential equations have a many-to-one correspondence with the 1st-order Hamiltonian phase-space formalism

Like any form of analytical mechanics, the Hilbert-space formalism provides a powerful set of mathematical tools for specifying microphysical laws in a systematic manner, for studying dynamical symmetries, and for calculating predictions

4. Causal Locality

Conventional Wisdom on (Causal?) Locality

Depending on whom you ask, the conventional wisdom you may hear about (causal?) locality in quantum theory could be:

• No-go theorems have ruled out hidden variables altogether

• Hidden-variables theories are possible in principle, but they entail nonlocality or nonlocal causation, and without them quantum theory is (causally) local

• Quantum theory is *unavoidably* nonlocal and/or causally nonlocal, with or without hidden variables

---

All these mutually inconsistent parcels of conventional wisdom are widely disputed!

Bell's theorem, according to Bell himself, actually only asserts the last of the three! (And I'll be challenging that assertion in this talk)

Nonlocality and Forces

In Newtonian mechanics, there is a perfectly clear way to identify nonlocality in the dynamical laws of a system

A Newtonian system's dynamical laws exhibit nonlocality precisely if they include an action-at-a-distance force or potential (e.g., in Newtonian gravity)

The trouble is that if we leave forces and potentials behind, as in stochastic processes or in quantum theory, then this simple definition of nonlocality is no longer available!

What is nonlocality rigorously supposed to *mean* now?

Can we look to causal influences to determine whether a system's dynamical laws are nonlocal?

The Situation in Quantum Theory

Quantum theory involves probabilistic rather than deterministic relationships between observations, so there isn't a tight linkage between purported cause-and-effect pairings

Moreover, the no-communication theorem ensures that one cannot use quantum systems to send faster-than-light messages, but that doesn't necessarily prohibit nonlocal causation from going on behind the scenes

So it's not immediately obvious whether quantum theory involves any nonlocal causation

Causal Locality, Defined

There is a case to be made that causal talk is just “folk science” [Norton, 2003], and not physically fundamental

In that case, asking whether quantum theory is fundamentally causally local is arguably either unimportant or meaningless

But let’s address those who take causal locality seriously, and, along the way, show how to make quantum theory a hospitable domain for talk of causal influences

---

Let’s start with a simple attempt at a definition of causal locality:

Causal influences cannot propagate faster than light.

Notice that this is a condition on any causal influences that *happen to occur*, not an assertion that there *must exist* particular causal influences [Myrvold, 2024]!

Einstein, Podolsky, and Rosen

In a 1935 paper, "Can [the] Quantum-Mechanical Description of Physical Reality Be Considered Complete?", Einstein, Podolsky, and Rosen used an early version of quantum steering

In simplified form, if two particles are prepared in an entangled wave function, and then separated in space, a measurement of one particle in a chosen basis can mean that the other particle's wave function collapses to a corresponding basis

So the first observer can *seemingly* "steer" the other particle to a chosen collapse-basis, in language introduced shortly thereafter by Schrödinger

However, the no-communication theorem prohibits the first observer from controlling the *specific* wave function for the other particle in that collapse-basis

The EPR Authors’ Interpretation

Einstein, Podolsky, and Rosen took for granted that the first observer couldn’t *actually* have any influence on the other particle

Some quantum-steered wave-functions are eigenstates of certain observables, so EPR argued that the other particle must *already* have predetermined values of those observables, a fact not captured by the original two-particle wave function

Hence, the authors’ conclusion that quantum theory is incomplete

Contestable Implications for Nonlocal Causation

One could attempt to read the EPR argument instead as implying that the first observer's measurement intervention nonlocally causes the other particle to collapse to its final wave function

This would be a concrete manifestation of what Einstein in 1947 called "spooky action at a distance" ("*spukhafte Fernwirkung*")

But this reading is contestable because it relies on questionable notions:

• Wave-function collapse

• An interventionist conception of causation, and interventions (in this case measurement settings and measurement outcomes) are not thought to be physically fundamental things

Bell's 1964 Theorem

In 1964, Bell was inspired by the EPR argument and an existing nonlocal hidden-variables theory (de Broglie-Bohm pilot-wave theory, or Bohmian mechanics) to write a paper "On the Einstein-Podolsky-Rosen Paradox" attempting to tackle the question of nonlocal causation head-on

Bell viewed the EPR argument as creating a logical fork: *either* accept causal nonlocality, *or* provide hidden variables that uniquely predetermine specific measurement outcomes

To that end, Bell considered measurement-deterministic hidden-variables theories in which the hidden variables dictated specific measurement outcomes (as in Bohmian mechanics)

Bell's goal was to show that the second prong of the fork could not ultimately save causal locality

Set-Up for Bell's 1964 Theorem

Ingredients:

- λ = the hidden variables
- $A, B = \pm 1$ = far-separated measurement outcomes
- $\mathbf{a}, \mathbf{b}$ = local measurement settings

Assumptions for Bell's notion of "local causation":

- $A = A(\mathbf{a}, \lambda), \quad B = B(\mathbf{b}, \lambda)$

And there is also an implicit assumption of an interventionist conception of causation

Crucial Assumption about Expectation Values

To Bell, these assumptions implied the following expression for the statistical average or expectation value of pairwise products AB of measurement outcomes:

$$P(\mathbf{a}, \mathbf{b}) = \int d\lambda\, \rho(\lambda) \underbrace{A(\mathbf{a}, \lambda) B(\mathbf{b}, \lambda)}_{\text{crucial factorization!}}$$

Here $\rho(\lambda)$ is the standalone probability distribution for the hidden variables λ (a big assumption itself!)

It's ironic that Bell's results hinge on assumptions about expectation values

Earlier, Bell had identified a flaw in a 1932 anti-hidden-variables theorem of von Neumann that likewise came down to unjustified assumptions about expectation values!

(Grete Hermann actually got there first, in the 1930s)

The Bell Inequality

Using this formula for expectation values, Bell was able to prove his famous Bell inequality, which should then be satisfied by all measurement-deterministic hidden-variables theories satisfying Bell's local-causality assumptions:

$$1 + P(\mathbf{b}, \mathbf{c}) \geq |P(\mathbf{a}, \mathbf{b}) - P(\mathbf{a}, \mathbf{c})|$$

Quantum theory predicts violations of the Bell inequality

Indeed, the 2022 Nobel Prize in Physics was awarded to Aspect, Clauser, and Zeilinger for experimentally confirming those violations (Press release: "This means that quantum mechanics cannot be replaced by a theory that uses hidden variables." (!?!))

Bell's 1964 paper therefore appears to rule out locally causal measurement-deterministic hidden-variables theories

Implications of Bell's 1964 Argument

Provided one doesn't take the EPR argument to be definitive, Bell's 1964 argument leaves open several possibilities:

- Nonlocally causal measurement-deterministic hidden-variables theories (like Bohmian mechanics)
- Measurement-stochastic hidden-variables theories
- Formulations of quantum theory that attempt to eschew hidden variables completely (includes the textbook theory!)

Bell's 1964 argument certainly doesn't rule out hidden variables altogether!

In 1975, Bell attempted to generalize his 1964 theorem to encompass the second and third possibilities (again, including the textbook theory!), and also avoided relying on interventionism

Bell's 1975 Theorem

Bell's 1975 argument applied to all theories with stochastic measurement outcomes, with or without hidden variables, so that includes the textbook theory

Bell's goal in the 1975 paper was to show that all empirically adequate such theories involve nonlocal causation

---

One big problem was how to establish causation when measurement outcomes are stochastic (so no tight linkages)

Another was how to compute the needed expectation value using a more general probability distribution $\rho(A, B|\mathbf{a}, \mathbf{b}, \lambda)$:

$$P(\mathbf{a}, \mathbf{b}) = \int d\lambda\, \rho(\lambda) \sum_{A,B} \underbrace{\rho(A, B|\mathbf{a}, \mathbf{b}, \lambda) AB}_{\text{no factorization anymore?}}$$

Bell's First Attempted Principle

Recounting his 1975 theorem in that 1990 lecture, Bell starts with a first attempt at a principle of local causality:

> "The direct causes (and effects) of events are near by, and even the indirect causes (and effects) are no further away than permitted by the velocity of light."

This is very similar to the definition of causal locality from the beginning of this talk – but then Bell goes on to say:

> "The above principle of local causality is not yet sufficiently sharp and clean for mathematics."

Just before he states his second attempted principle of local causality, he includes the following very important warning:

> "Now it is precisely in cleaning up intuitive ideas for mathematics that one is likely to throw out the baby with the bathwater. So the next step should be viewed with the utmost suspicion."

Set-Up for Bell's 1975 Theorem

Ingredients:

• λ = "*be*ables" (could even include the wave function itself!)

• A, B = far-separated measurement outcomes, $|A|, |B| \leq 1$

• $\mathbf{a}$, $\mathbf{b}$ = local measurement settings

can now be regarded as non-interventionist beables themselves!

Assumptions for Bell's new principle of local causality:

• There is a sufficiently rich collection of local beables λ in the overlap of the past light cones of A, B (treated as beables) that

$$\rho(A|\mathbf{a}, \mathbf{b}, B, \lambda) = \rho(A|\mathbf{a}, \lambda), \quad \rho(B|\mathbf{a}, \mathbf{b}, A, \lambda) = \rho(B|\mathbf{b}, \lambda)$$

In general: $\rho(A, B|\mathbf{a}, \mathbf{b}, \lambda) = \rho(A|\mathbf{a}, \mathbf{b}, B, \lambda)\rho(B|\mathbf{a}, \mathbf{b}, \lambda)$

So equivalently: $\rho(A, B|\mathbf{a}, \mathbf{b}, \lambda) = \underbrace{\rho(A|\mathbf{a}, \lambda)\rho(B|\mathbf{b}, \lambda)}_{\text{factorization!}}$

Expectation Values and Inequalities Revisited

The expectation value for the product AB from before is now:

$$P(\mathbf{a},\mathbf{b}) = \int d\lambda\, \rho(\lambda) \underbrace{\left(\sum_A \rho(A|\mathbf{a},\lambda)A\right)\left(\sum_B \rho(B|\mathbf{b},\lambda)B\right)}_{\text{factorization!}}$$

Bell was then able to prove a generalization of his inequality, first written down by Clauser, Horne, Shimony, Holt in 1969

So any measurement-stochastic theory that satisfies Bell's local-causality assumptions is ruled out by experiment

Bell concluded that all other theories (including the textbook theory) therefore involve nonlocal causation

Reichenbach's Principle of Common Causes

But Bell's new principle of local causality tucked in a new implicit assumption, as is manifest in its factorization form:

Reichenbach's principle of common causes (1956)

In modern form, Reichenbach's principle states that if A, B are variables that are correlated

$$P(A, B) \neq P(A)P(B),$$

but A, B do not causally influence each other, then there is the assertion that there must exist some other variable C such that

$$P(A, B|C) = P(A|C)P(B|C)$$

In words, there is the assertion that there must exist a "common cause" variable that "screens off" the correlation

Bell's Principle of Local Causality and Reichenbach

Recall Bell's principle of local causality:

$$\rho(A|\mathbf{a},\mathbf{b},B,\lambda) = \rho(A|\mathbf{a},\lambda), \quad \rho(B|\mathbf{a},\mathbf{b},A,\lambda) = \rho(B|\mathbf{b},\lambda)$$

And in its equivalent factorization form:

$$\rho(A,B|\mathbf{a},\mathbf{b},\lambda) = \rho(A|\mathbf{a},\lambda)\rho(B|\mathbf{b},\lambda)$$

These are clearly an application of Reichenbach's principle of common causes, and actively assert the existence of a particular causal influences

Here the role of the asserted "common cause" C is played by the local beables λ in the overlap of the past light cones of A, B

Essentially, lacking a concrete microphysical theory of causal influences, Bell *assumed* that any causally local theory should entail local Reichenbachian common-cause variables

Causal Locality versus Local Causality

A terminological distinction, due to Myrvold:

• "Causal locality" is the more basic condition that any causal influences that happen to occur should respect the finite speed of light

Notice that causal locality does not actively assert the existence of any particular causal influences!

• "Local causality" is a stronger condition involving the active assertion that some (local) causal influences must exist in some situations, such as the common causes in Reichenbach principle and Bell's principle

Good Reasons to Doubt Reichenbach

Reichenbach's principle of common causes may seem intuitively plausible, but that's far from an obvious reason to require it as part of a condition on causal locality!

For one thing, it assumes every "common cause" is a variable that can be conditioned on and summed/integrated over!

As Unruh wrote in 2002:

> "It is true that this common cause cannot be stated in exactly the form which for example Reichenbach set up to describe common causes for a classical statistical system. But that is not surprising. Quantum mechanics is not classical mechanics. The structure of the correlations in a quantum system differ from those in a classical system, as Bell so succinctly showed. But those correlations do not arise mysteriously somehow in the development of a widely spaced system. Those correlations do not require some mysterious non-local action to be explained. They are simply there, as are correlations in a classical system, due to the evolution from a common (quantum) cause in the past."

In short, interactions are simply not Reichenbachian variables!

Other Theorems

Other arguments purporting to demonstrate nonlocal causation (EPR, GHZ, etc.) depend on an interventionist conception of causation, or (Bong et. al.) assume the *a priori* existence of theoretical probability distributions at intermediate times without adequate justification

It is not clear how one would derive the Bell inequality or prove these other theorems when working at the level of the *atoms(!)* that make up the measuring devices

Note that advocates of Everettian QM already deny that the premises of these theorems capture the correct notion of locality

---

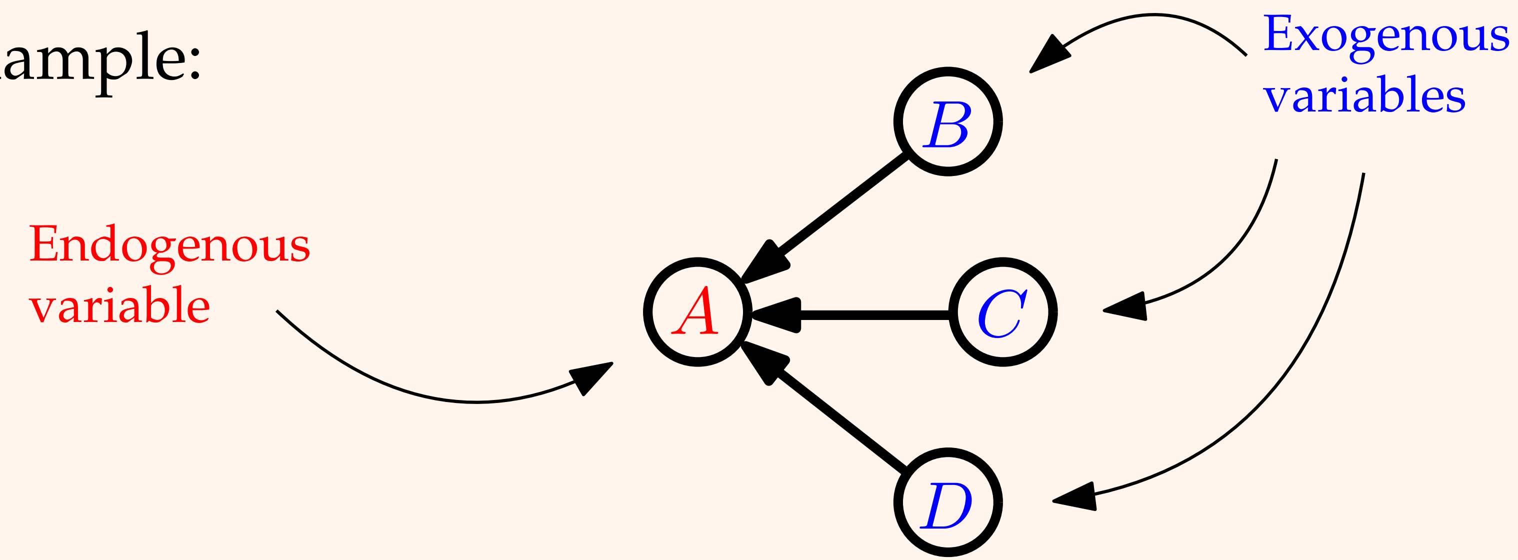
Next I'll show that the new indivisible formulation exploits these loopholes to allow for an improved criterion for causal locality, and that it's causally local according to that criterion

Bayesian Networks and Causation

The first step is to note a key connection with another theory

In the theory of Bayesian networks, one considers a set of random variables related by a collection of basic, directed conditional probabilities

For example:



Here we have a simple Bayesian network with four random variables A, B, C, D that supplies a basic, directed conditional probability distribution $p(A = a | B = b, C = c, D = d)$

Basic versus Derived Conditional Probabilities

In a concrete instantiation of this Bayesian network, one has a contingent joint probability distribution $p(b, c, d)$

This induces a contingent standalone probability distribution $p(a)$ according to the following multilinear relation:

$$p(a) = \sum_{b,c,d} p(a|b, c, d)p(b, c, d)$$

One can define other conditional probability distributions, but these will be derived rather than basic, and will depend nonlinearly on the contingent probabilities of the given instantiation

For example: $$p(b|a, c, d) = \frac{p(a|b, c, d)p(b, c, d)}{\sum_{b'} p(a|b', c, d)p(b', c, d)}$$

So the basic, directed conditional probabilities supplied by the Bayesian network have a more fundamental status as the laws of the model

A Non-Interventionist Causal Reading

Crucially, Bayesian networks admit a non-interventionist causal reading

If the model supplies a basic, directed conditional probability distribution $p(a|b, c, d)$, then any stochastic fluctuations in B, C, D dictate stochastic fluctuations in A

Taking this seriously, one can argue that by supplying $p(a|b, c, d)$ in its laws, the model is asserting that B, C, D exerts a causal influence on A

Notice how the directedness of these nomological (lawlike) conditional probabilities nicely captures the asymmetric nature of cause-and-effect relationships

Yet it does so without privileging an arrow of time, so it threads a very fine needle!

A Nomological Theory of Causation

The indivisible formulation of quantum theory in this talk is based on microphysical laws consisting of directed nomological (lawlike) conditional probabilities, just like for a Bayesian network

The indivisible formulation therefore provides a hospitable domain for causal talk

One can arguably apply a causal reading to those nomological conditional probabilities, motivating a new theory of causation:

> On a theory X, to say that the variables $B, C, \ldots$ have nomological causal influences on a variable A is just to say that the theory X specifies, in its basic, fixed laws, a conditional probability of the form $p(A|B, C, \ldots)$, read as "the nomological conditional probability of A, given $B, C, \ldots$."

In that sense, one can regard quantum theory as a nomological theory of causation *par excellence*

Causal Independence

Consider an overall unistochastic system consisting of two subsystems $\mathcal{Q}, \mathcal{R}$

The overall system's nomological conditional probabilities then take the following form:

$$p((q_t, r_t), t|(q_0, r_0), 0)$$

Definition: $\mathcal{Q}$ is free of causal influences from $\mathcal{R}$ over the time interval from 0 to t if after marginalizing over r_t, the resulting conditional probability distribution has no dependence on r_0:

$$p(q_t, t|(q_0, r_0), 0) = p(q_t, t|q_0, 0)$$

An Improved Principle of Causal Locality

One can now state an improved principle of causal locality:

> A theory with microphysical directed conditional probabilities is *causally local* if any pair of localized systems $\mathcal{Q}$ and $\mathcal{R}$ that remain at spacelike separation in the given situation never exert causal influences on each other, in the sense that the directed conditional probabilities for $\mathcal{Q}$ are independent of $\mathcal{R}$, and vice versa.

It's a straightforward calculation to show that this principle is satisfied by the indivisible formulation presented in this talk

standard rule

Spacelike separation $\implies U^{\mathcal{Q}\mathcal{R}}(t \leftarrow 0) = U^{\mathcal{Q}}(t \leftarrow 0) \otimes U^{\mathcal{R}}(t \leftarrow 0)$

$$\implies \Gamma^{\mathcal{Q}\mathcal{R}}(t \leftarrow 0) = \Gamma^{\mathcal{Q}}(t \leftarrow 0) \otimes \Gamma^{\mathcal{R}}(t \leftarrow 0)$$

mod-square commutes with tensor products

$$\implies p((q_t, r_t), t|(q_0, r_0), 0) = p(q_t, t|q_0, 0)\, p(r_t, t|r_0, 0)$$

$$\implies p(q_t, t|(q_0, r_0), 0) = p(q_t, t|q_0, 0) \qquad \text{QED}$$

Interactions and Entanglement

By contrast, suppose that $\mathcal{Q}, \mathcal{R}$ are not kept at spacelike separation and *do* interact – at some interaction time t'

Then: $U^{\mathcal{Q}\mathcal{R}}(t' \leftarrow 0) \neq U^{\mathcal{Q}}(t' \leftarrow 0) \otimes U^{\mathcal{R}}(t' \leftarrow 0)$

$$\implies \Gamma^{\mathcal{Q}\mathcal{R}}(t \leftarrow 0) \neq \Gamma^{\mathcal{Q}}(t \leftarrow 0) \otimes \Gamma^{\mathcal{R}}(t \leftarrow 0) \text{ for all } t \geq t'$$

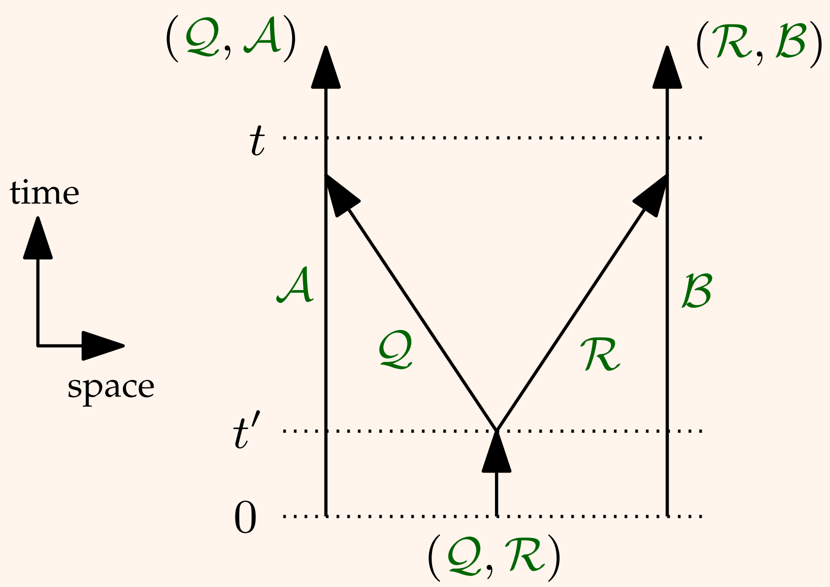
because $\Gamma^{\mathcal{Q}\mathcal{R}}(t)$ encodes cumulative statistical effects starting at time 0, at least until the next division event

This breakdown in factorization of $\Gamma^{\mathcal{Q}\mathcal{R}}(t)$ starting at t' is precisely entanglement as manifested at the level of the underlying indivisible unistochastic process

The two subsystems $\mathcal{Q}, \mathcal{R}$ are exerting causal influences on each other, stemming from their local interaction at t' – i.e., the common cause – but notice that it's not a Reichenbachian variable!

Revisiting the EPR Argument

Adding observer-subsystems $\mathcal{A}$ ("Alice") and $\mathcal{B}$ ("Bob") to model the EPR argument doesn't change these basic facts



One can show by a straightforward calculation that $\mathcal{B}$ does not exert causal influences on $\mathcal{A}$ in the required sense:

$$p(a_t, t|(q_0, r_0, a_0, b_0), 0) = p(a_t, t|(q_0, r_0, a_0), 0)$$

The only causal influences on $\mathcal{A}$ come from $\mathcal{Q}, \mathcal{R}$, which are both in its past light cone, as expected

5. Concluding Remarks

Conclusion

In short, consider a stochastic process over a fixed orthonormal basis, without imposing all the intricate nomological structure of a textbook stochastic process, and you get quantum theory

One then seems to have a causally local hidden/physical-variables theory, based on simpler axioms than textbook quantum theory, arguably without a measurement problem, and deflating a lot of the exotic talk about quantum phenomena

---

There are many prospects for future research directions:

- Applications to dynamical systems and stochastic processes?
- New algorithms for quantum simulations?
- New ways to think about quantum causal models?
- Implications for old problems in statistical mechanics?
- Ramifications for algebra-first approaches?
- Generalizations of quantum theory?
- New possibilities for quantum gravity?

Thank you!