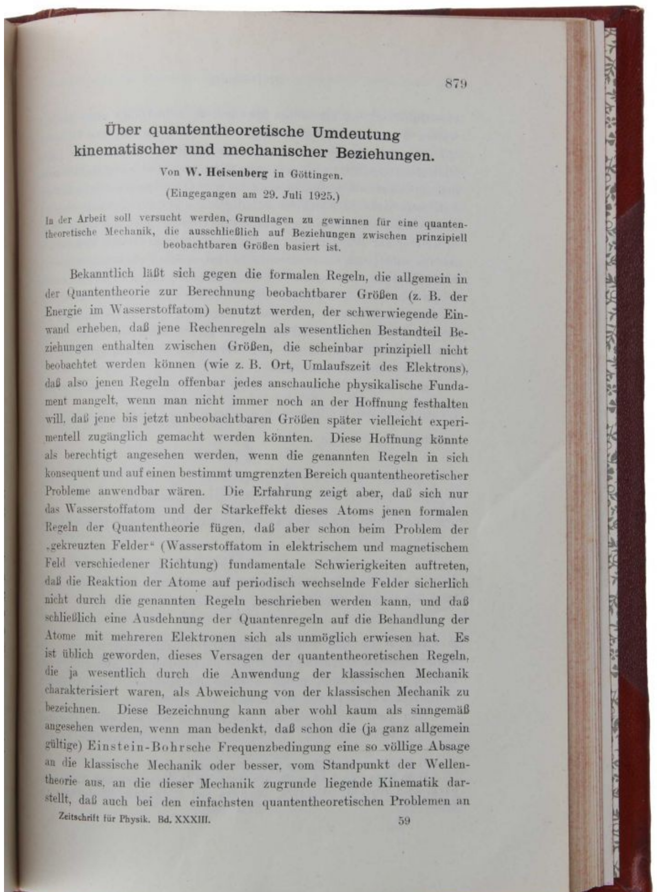
Is Quantum Mechanics Exact?

Anton Kapustin

Simons Center for Geometry and Physics

Stony Brook University

This year Quantum Theory will celebrate its 90th birthday.


879

Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen.

Von W. Heisenberg in Göttingen.

(Eingegangen am 29. Juli 1925.)

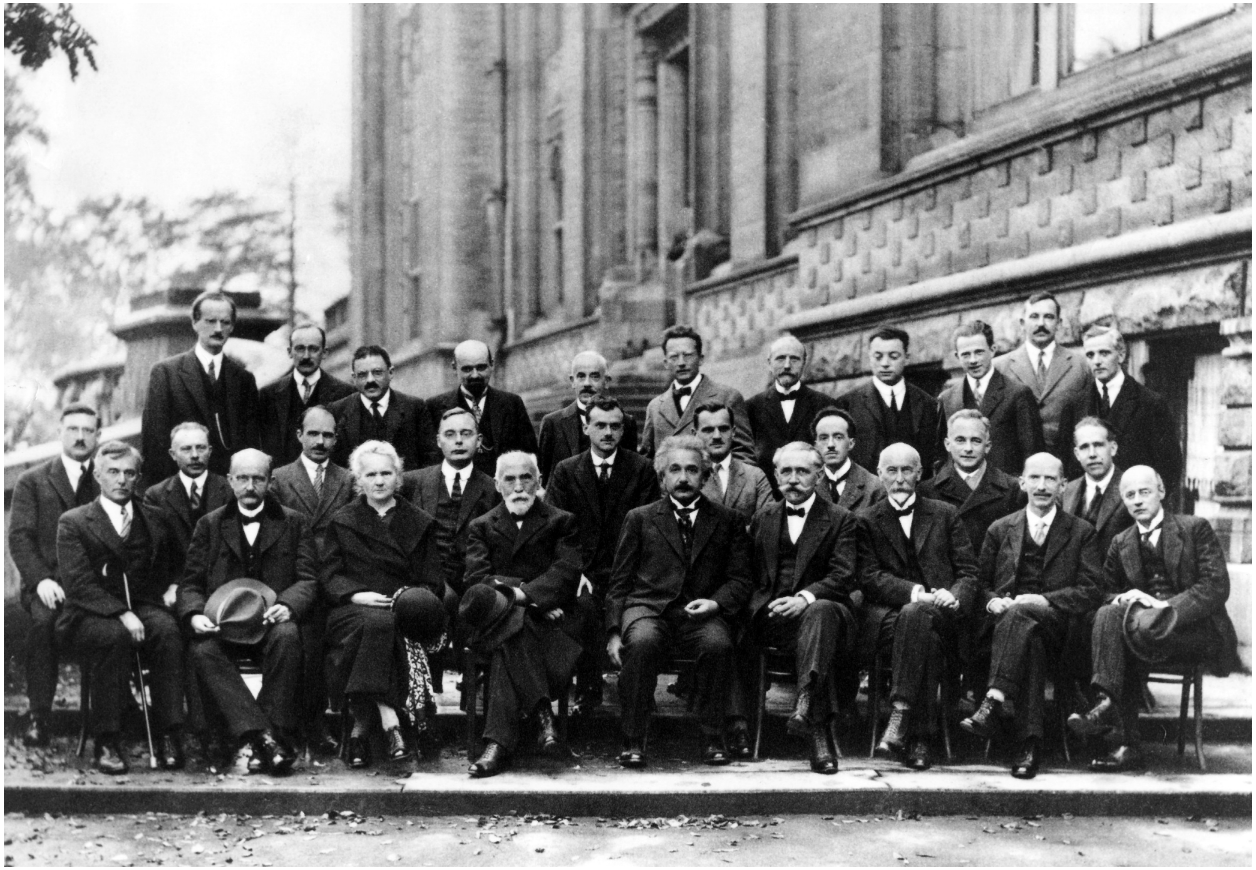
In der Arbeit soll versucht werden, Grundlagen zu gewinnen für eine quantentheoretische Mechanik, die ausschließlich auf Beziehungen zwischen prinzipiell beobachtbaren Größen basiert ist.

Bekanntlich läßt sich gegen die formalen Regeln, die allgemein in der Quantentheorie zur Berechnung beobachtbarer Größen (z. B. der Energie im Wasserstoffatom) benutzt werden, der schwerwiegende Einwand erheben, daß jene Rechenregeln als wesentlichen Bestandteil Beziehungen enthalten zwischen Größen, die scheinbar prinzipiell nicht beobachtet werden können (wie z. B. Ort, Umlaufszeit des Elektrons), daß also jenen Regeln offenbar jedes anschauliche physikalische Fundament mangelt, wenn man nicht immer noch an der Hoffnung festhalten will, daß jene bis jetzt unbeobachtbaren Größen später vielleicht experimentell zugänglich gemacht werden könnten. Diese Hoffnung könnte als berechtigt angesehen werden, wenn die genannten Regeln in sich konsequent und auf einen bestimmt umgrenzten Bereich quantentheoretischer Probleme anwendbar wären. Die Erfahrung zeigt aber, daß sich nur das Wasserstoffatom und der Starkeffekt dieses Atoms jenen formalen Regeln der Quantentheorie fügen, daß aber schon beim Problem der „gekreuzten Felder" (Wasserstoffatom in elektrischem und magnetischem Feld verschiedener Richtung) fundamentale Schwierigkeiten auftreten, daß die Reaktion der Atome auf periodisch wechselnde Felder sicherlich nicht durch die genannten Regeln beschrieben werden kann, und daß schließlich eine Ausdehnung der Quantenregeln auf die Behandlung der Atome mit mehreren Elektronen sich als unmöglich erwiesen hat. Es ist üblich geworden, dieses Versagen der quantentheoretischen Regeln, die ja wesentlich durch die Anwendung der klassischen Mechanik charakterisiert waren, als Abweichung von der klassischen Mechanik zu bezeichnen. Diese Bezeichnung kann aber wohl kaum als sinngemäß angesehen werden, wenn man bedenkt, daß schon die (ja ganz allgemein gültige) Einstein-Bohrsche Frequenzbedingung eine so völlige Absage an die klassische Mechanik oder besser, vom Standpunkt der Wellentheorie aus, an die dieser Mechanik zugrunde liegende Kinematik darstellt, daß auch bei den einfachsten quantentheoretischen Problemen an

Zeitschrift für Physik. Bd. XXXIII. 59

Werner Heisenberg's paper "Quantum theoretic reinterpretation of kinematic and mechanical relations" was published in September 1925.

Quantum Mechanics was created by Heisenberg, Born, Pauli, Schroedinger and Dirac in 1925-1926 and has remained essentially unchanged since then.



Mathematical foundations of Quantum Mechanics have been clarified by von Neumann (1932). Although from a physical viewpoint he did not add anything new, von Neumann's axiomatization of QM turned out quite useful.

Another axiomatization of QM was given by Mackey (1957).

Kinematical structure of QM

- States are vectors in Hilbert space (up to a scalar multiple), or more generally "density matrices" (positive Hermitian operators with unit trace).
- Observables are arbitrary Hermitian operators. Possible values of A are its eigenvalues a_i.
- Born rule:

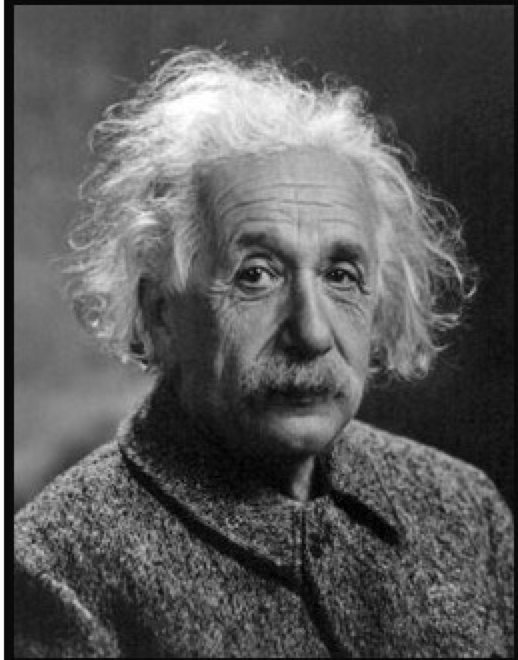
$$p(a_i) = \mathrm{Tr}\,\rho P_i$$

Combining Quantum Mechanics with relativistic invariance turned out to be a difficult problem, but it was solved in the end, on the physical level of rigor. Both QM and Relativity Theory emerged unscathed.

A mathematically satisfactory formulation of Relativistic Quantum Mechanics, i.e. Quantum Field Theory, is still incomplete, but everybody is sure it can be accomplished without modifying the rules of QM.

Is Quantum Theory an exact theory, or only an approximation to some deeper theory?

The answer is not obvious. Many reasonably smart people thought about it, without coming to any definite conclusion.



Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the old one. I, at any rate, am convinced that He does not throw dice.

(Albert Einstein)

Hidden variables?

Many people proposed that QT is incomplete, and that probabilities only arise because it does not take into account all relevant variables (Einstein, Bohm, etc.).

But J.S. Bell (1964) showed that any theory with local hidden variables will badly violate some prediction of QT.

But since QT works quite well, this means that hidden variables would violate causality. Not very appealing.

J. S. Bell himself was not satisfied:

“So for me, it is a pity that Einstein's idea doesn't work. The reasonable thing just doesn't work.”

More recently, Gerard ‘t Hooft was the leading proponent of the idea that QT is not fundamental, but is only an approximation.

Arguments against exactness of QT

- No physical theory is exact, everything is an approximation.
- Where do probabilities come from?
- How does one explain the collapse of the wave-function?
- There are persistent problems in combining QT and Einstein's gravity theory (although most people blame gravity, not QT).

Actually, there are some things which we know to be exactly true.

Say, all electrons are identical and indistinguishable. Otherwise Pauli statistics would not even make sense.

Alternatively, since electrons are excitations of a quantum Dirac field, it is logically necessary for them to be indistinguishable.

Collapse of the wave function

The Copenhagen interpretation of QM:
measurement induces a collapse of the wave function

$$|\psi\rangle \mapsto \sum_i P_i |\psi\rangle\langle\psi| P_i$$

$$\rho \mapsto \sum_i P_i \, \rho \, P_i$$

What is measurement, and why is it different from ordinary evolution?

Possible responses

- No collapse ever happens (most common viewpoint), apparent collapse can be explained in terms of decoherence.
- Collapse is a real physical phenomenon which violates unitary evolution of the wave function (Penrose: perhaps this is a gravity effect).

The black hole evaporation problem

Stephen Hawking showed in 1974 that black holes produce thermal radiation. As a result, they gradually evaporate.

Hawking's result uses both Quantum Theory and Einstein's gravity.

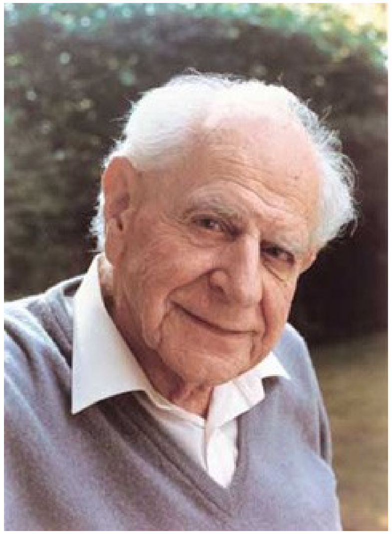
But the result is very hard to reconcile with causality. So one of the three (Einstein's gravity, causality, or QT) must be modified.

Arguments for exactness of QM

- It works great! All of atomic theory, solid state physics, and elementary particle theory is based on QM, and so far not serious disagreements have been found.
- Nobody managed to propose a sensible modification or generalization of QM.

Testing Quantum Mechanics

How do we test QM? This is harder than it looks.



K. Popper: to test a theory, just test any of its predictions. If a prediction is falsified, the theory is wrong.

But if a theory is already quite successful, we need to know where to look for deviations from it, or else we might easily miss them.

In other words, to test QM, we need a plausible alternative to QM.

Ideally, we would like to have a theory with a small parameter λ which reduces to QT if we set $\lambda=0$.

Then we make predictions using this new theory, find those which depend on λ, and try to compare with experiment.

If no deviations from QT are found, we can put an upper bound on λ. Say, $\lambda<10^{-10}$.

Modifications of Quantum Mechanics

There are not many proposals of this sort.

1. Ghirardi, Rimini, Weber (1985): QM plus objective collapse of the wave function.

2. Weinberg (1987): nonlinear QM.

GRW proposal

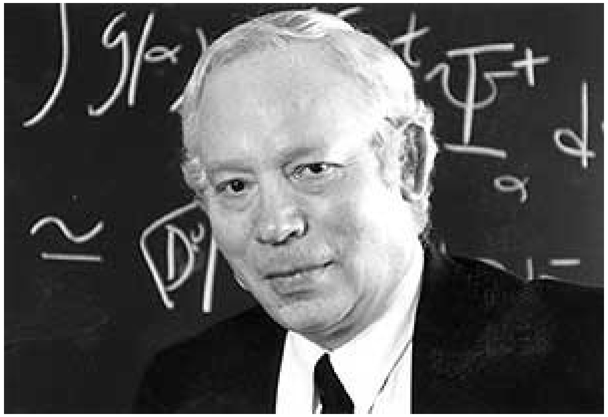
GRW postulate that the Hilbert space of a multi-particle system has a preferred basis of states: states where the particles have definite positions.

GRW propose that general wave-function spontaneously collapse towards such states. This happens randomly and has nothing to do with measurements.

Problems with the GRW scheme

- Makes sense only for spin-less particles.
- Makes sense only for non-relativistic QM.
- Violates energy conservation.

Weinberg’s nonlinear QM



Weinberg proposed a modification of QM which avoids some of these pitfalls.

Main idea: states are still vectors in Hilbert space, but observables are not simply Hermitian operators.

Hermitian operators are still observables, but not the most general ones.

Fix a basis ψ_i in Hilbert space. Then Hermitian operators are represented by Hermitian matrices:$_j$

$$\hat{A} \mapsto A^i_j = \langle \psi_i | \hat{A} | \psi_j \rangle$$

Weinberg proposed to generalize this to a sum of tensors of all possible ranks:

$$(A^i_j, A^{ij}_{kl}, A^{ijk}_{lmn}, \ldots)$$

The usual observables are a special case:

$$(A^i_j, 0, 0, \ldots)$$

Features of Weinberg's proposal

Noether theorem holds: every symmetry corresponds to a conserved quantity.

In particular, energy is conserved if the Hamiltonian is time-independent.

Problem: the notion of an eigenvector and eigenvalue is not defined. Theory makes no prediction about probabilities of measuring a particular value of an observable A.

Perhaps we should look for modifications of QM more systematically.

Or perhaps we can prove that no sensible modification is possible (a no-go theorem).

Related question: can one have a consistent theory where some degrees of freedom are classical and some are quantum? That is, is $\hbar$ universal?

Superselection sectors

A simple way to generalize QM is to allow "superselection sectors: the Hilbert space is written as a sum

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 + \ldots$$

and one does not allow states which are nontrivial linear combinations of states from different sectors.

One also does not allow operators which mix states from different sectors.

Typical example: each superselection sector contains states with a fixed electric charge, or baryon number.

In a sense, the charge labeling superselection sectors is a classical observable, but it is not a dynamical observable.

Is there a way to mix classical and quantum dynamics?

In the absence of good physical ideas, one can try to use

MATH!

Deforming theories

One can try to classify all possible deformations of an existing theory within the space of all sensible theories.

If no deformations exist, one says that the theory is rigid.

Inventing Special Relativity

Einstein discovered Special Relativity using physical considerations. But one could also discover it using the deformation approach (this was first noticed by F. Dyson).

Newton's mechanics is invariant under Galilean transformations:

$$\vec{x} \mapsto \Lambda \vec{x} - \vec{v} t, \quad t \mapsto t.$$

Here Λ is an orthogonal 3 x 3 matrix, and v is velocity.

The Galilean group is a Lie group of dimension 3+3=6.

Can it be deformed within the class of six-dimensional Lie groups?

Yes, and here is an easy way to see it.

The Galilean group can be described as the group of all linear transformations which leave invariant a contra-variant "metric"

$$G^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^i}$$

If we deform this degenerate metric to a non-degenerate one, we will also deform the group.

Therefore consider a contra-variant metric

$$G^{-1} = \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^i} - \frac{1}{c^2}\frac{\partial}{\partial t} \otimes \frac{\partial}{\partial t}$$

It has a real parameter $1/c^2$. A priori, it can be either positive or negative. This metric has an inverse:

$$G = dx^i \otimes dx^i - c^2 dt^2$$

If $c^2>0$, the invariance group is the Lorentz group SO(3,1).

If $c^2<0$, the invariance group is the 4d rotation group SO(4).
This case is unphysical.

The Galilean group is obtained in the limit $c \to \infty$.

The Lorenz group SO(3,1) is rigid:
no further deformations are possible.

Pitfalls

If the class of objects considered is too narrow, we may miss some interesting deformations.

For example, why did we not consider translations?

Combining Lorenz transformations and translations we get Poincare transformations:

$$x^\mu \mapsto \Lambda^\mu_\nu x^\nu + a^\mu, \quad \Lambda^\mu_\nu G_{\mu\rho} \Lambda^\rho_\tau = G_{\nu\tau}.$$

The Poincare group is not rigid and can be deformed into either SO(4,1) or SO(3,2).

SO(4,1) is the symmetry of the de Sitter space (vacuum solution of Einstein equations with a positive cosmological constant).

SO(3,2) is the symmetry of the anti-de Sitter space (vacuum solution of Einstein equations with a negative cosmological constant).

Poincare group is recovered in the limit of zero cosmological constant.

IS QUANTUM MECHANICS RIGID?

To answer this question, we first need to describe the set of sensible physical theories.

Minimal requirements:

(1) Both Classical Mechanics and Quantum Mechanics are elements of this set.

(2) Quantum Mechanics is a deformation of the Classical one.

Classical Mechanics

- Basic object: phase space (M,ω) where M is a manifold, ω is a symplectic structure:

$$\omega = dp_i dq^i$$

- Observables are real-valued functions on M.
- States are points on M (more generally, probability measures on M).
- Evolution is Hamiltonian:

$$\frac{dO}{dt} = \{H, O\}$$

Quantum Mechanics

- Basic object: Hilbert space V (a complex vector space with a positive-definite Hermitian scalar product).
- Observables are Hermitian operators on V.
- States are rays in V (or more generally, density “matrices” on V).
- Evolution of observables is given by

$$i\frac{d\hat{O}}{dt} = \left[\hat{H}, \hat{O}\right]$$

CM and QM look very different; can we really think of them as particular examples of a more general structure?

Similarities: Lie bracket on observables (Poisson bracket in CM, commutator in QM).

There is a good reason for Lie bracket: Noether theorem which states that conserved quantities are related to symmetries. Holds both in CM and QM

Equivalently, we want to be able to identify observables with infinitesimal symmetries of the system. Infinitesimal symmetries of anything form a Lie algebra.

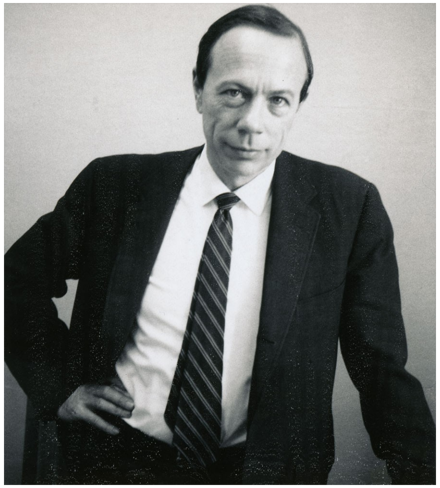
But:

Why do complex numbers appear in Quantum Mechanics but not in Classical Mechanics?

Why are observables in QM represented by Hermitian operators in Hilbert space?

The Born supremacy

I did not ask "Why are states represented by density matrices?", or equivalently "Why the Born rule?" because there is already a good answer.



A. M. Gleason (1957): the only sensible way to define probability of measuring a particular value of an observable is via the Born rule.

Elementary outcomes

Elementary outcomes are "yes/no" observables. They correspond to projectors in V, $O^2=O$.

Projectors to orthogonal subspaces correspond to mutually exclusive outcomes.

A probability measure P(O) should satisfy P(O)≥0. If the projectors O_i satisfy

$$\sum_i O_i = \mathbf{1}$$

then we must have

$$\sum_i P(O_i) = 1.$$

Gleason's theorem

If V has dimension 3 or greater, any probability measure on projectors has the form

$$P(O) = \mathrm{Tr}\,\rho\,O,$$

where ρ is a density matrix on V.

Therefore let’s forget about states and focus on observables.

Observables in CM form a real vector space. They can also be multiplied, and this operation is associative and commutative.

In mathematical terms, observables in CM form a commutative algebra over reals.

There is also a Lie bracket, which satisfies the Leibniz rule:

$$[FG, H] = G[F, H] + F[G, H]$$

But Hermitian operators do not form an associative algebra: if A and B are Hermitian operators, AB is not Hermitian, if [A,B] is nonzero.

Put differently, multiplication of complex operators does not have a physical meaning, because complex operators on V do not have a physical meaning.

This has bothered the Founding Fathers of QM, but a satisfactory answer never materialized.

Jordan algebras

Pascual Jordan (1933) proposed to consider the anti-commutator as the basic operation:

$$A \circ B = \frac{1}{2}(AB + BA).$$

This operation is commutative, but not associative. Still:

$$(A \circ B) \circ (A \circ A) = A \circ (B \circ (A \circ A))$$

(the Jordan identity).

A real vector space with a commutative "multiplication" satisfying the Jordan identity is called a Jordan algebra.

But what is the physical meaning of $A \circ B$, and why should it satisfy this weird identity?

Jordan's idea: first, it is sufficient to consider $A^2 = A \circ A$, because:

$$A \circ B = \frac{1}{2}\left((A+B)^2 - A^2 - B^2\right).$$

Higher powers of A can be defined recursively:

$$A^n = A \circ A^{n-1}.$$

The Jordan identity then implies power-associativity:

$$A^n \circ A^m = A^{n+m}$$

Thus, if we are given a squaring operation, and if the corresponding bilinear operation $A \circ B$ satisfies the Jordan identity, we can consistently define f(A) for any real polynomial function f(x).

The ability to define real functions of observables is a very desirable property.

Jordan identity is sufficient for this, but is it necessary?

Jordan (1933): it is necessary if the squaring operation is "formally real":

$$\sum_{i=1}^{N} A_i^2 = 0 \;\Rightarrow\; A_i^2 = 0 \;\forall i$$

P. Jordan, "Über Verallgemeinerungsmöglichkeiten des Formalismus der Quantenmechanik".

Jordan's theorem

Suppose the set of observables $\mathcal{A}$ is a real vector space such that

- f(A) can be consistently defined for any polynomial function f(x). That is if h(x)=f(g(x)), then h(A)=f(g(A)).
- The squaring operation is formally real.

Then $\mathcal{A}$ is a formally real Jordan algebra.

Such algebras (in the finite-dimensional case) have been classified by Jordan, von Neumann and Wigner (1934).

There are four infinite families, and one isolated exceptional case.

Three of the infinite families arise from real, complex and quaternionic matrices.

One more infinite family is related to the algebra of Dirac's gamma-matrices.

The exceptional one is related to octonions.

Not a very satisfactory outcome.

- Why does Nature only use one of five cases?

- Why should the squaring operation be formally real?

Indeed, if A and B do not commute, then the set-up for measuring A^2+B^2 has nothing to do with measuring A or B.

But the idea that the basic physical operations are squaring and the Lie bracket is sound.

They are defined both for QM and CM, and have a clear physical meaning.

We just need to understand why for CM the squaring gives rise to an associative multiplication, but for QM it does not.

Composability

Niels Bohr proposed that nontrivial constraints could come from the requirement of composability:

Given two physical systems with their sets of observables, one should be able to form a composite system. Observables for the composite should include products of observables of subsystems.

This idea was taken up by Bohr's assistant Aage Petersen and Emil Grgin.

Theorem (Grgin and Petersen, 1972).

Suppose the set of observables of every physical system is a real vector space with two bilinear operations $A \circ B$ and $[A,B]$, where $A \circ B$ is symmetric and $[A,B]$ is a Lie bracket.

Suppose further that when forming a composite system the spaces of observables are tensored, and the bilinear operations are linear combinations of products of bilinear operations for the subsystems.

Then there is a real number λ, universal for all systems, such that

$$(A \circ B) \circ C - A \circ (B \circ C) = \lambda[[A, C], B]$$

The case $\lambda=0$ is the case of Classical Mechanics.

The case $\lambda>0$ is the case of Quantum Mechanics. If we set $\lambda=\hbar^2$ then the Grgin-Petersen identity says that $A \circ B$ and $[A,B]$ are expressed through an associative product:

$$A \circ B = \frac{1}{2}(AB + BA)$$
$$[A, B] = \frac{1}{2i\hbar}(AB - BA)$$

The case $\lambda<0$ is similar to QM, but the “Planck constant” $\hbar$ is purely imaginary.

Remarks

- The Grgin-Petersen theorem “explains” where complex numbers in QM come from: they arise when we rewrite the properties of real operations $A \circ B$ and $[A,B]$ in terms of a more familiar associative product.
- It also explains why $\hbar$ is the same for all systems.
- Still need to explain why observables are matrices (rather than elements of more general associative algebras).
- Also need to explain why imaginary $\hbar$ is not OK.
- The conditions of the theorem are not physically motivated.

Synthesis

- Observables of any physical system form a real Lie algebra.
- There is also a squaring operation.
- When forming composite system, there should be a way to multiply observables of subsystems.
- Observables of two different subsystems commute.

Theorem (A.K., 2013).

With these assumptions, the conclusions of the Grgin-Petersen Theorem hold.

Corollary.

(Finite-dimensional) Quantum Mechanics is rigid.

Indeed, since all the data are given by an associative multiplication, every deformation of QM should arise from a deformation of the algebra. But algebras of matrices are known to be rigid.

This is a no-go theorem we have been looking for.

It shows that finite-dimensional QM cannot be modified without violating some basic physical requirements, like Noether theorem or the ability to form composite systems.

Adding a couple more reasonable requirements, we can even prove that for finite-dimensional systems the QM is the only sensible theory.

Namely, the Planck constant must be real and nonzero, and the algebra of observables must be a sum of matrix algebras over complex numbers.

These extra requirements are:

(1) Every observable has a nonempty set of possible values.

(2) If there is only one possible value, the observable must be trivial (i.e. it is a c-number).

Conclusions

- QM is likely to be an exact theory of Nature for finite-dimensional systems.
- Complex numbers are needed for a good reason. $\hbar$ is universal.
- Any modification of QM is likely to violate the Noether theorem, or destroy a probabilistic interpretation.
- One possible loophole: infinite-dimensional systems (perhaps more general algebras than the algebras of bounded operators in Hilbert space are needed).