Is Quantum Mechanics An Island In Theoryspace?

SCOTT AARONSON
University of California, Berkeley CA USA
aaronson@cs.berkeley.edu

This paper investigates what happens if we change quantum mechanics in several ways. The main results are as follows. First, if we replace the 2-norm by some other p-norm, then there are no nontrivial norm-preserving linear maps. Second, if we relax the demand that norm be preserved, we end up with a theory that allows rapid solution of hard computational problems known as PP-complete problems (as well as superluminal signalling). And third, if we restrict amplitudes to be real, we run into a difficulty much simpler than the usual one based on parameter-counting of mixed states.

1 Introduction

> "It is striking that it has so far not been possible to find a logically consistent theory that is close to quantum mechanics, other than quantum mechanics itself." —Steven Weinberg, *Dreams of a Final Theory* [13]

The title of this paper should be self-explanatory, but if not: "theoryspace" is the space of logically conceivable physical theories, with two theories close to each other if they differ in few respects. An "island" in theoryspace is a natural and interesting theory, whose neighbors are all somehow perverse or degenerate.[a] The Standard Model isn't an island, because we don't know any compelling (non-anthropic) reason why the masses and coupling constants should have the values they do.[b] Likewise, general relativity is probably not an island, because of alternatives such as the Brans-Dicke theory.

To many physicists, however, quantum mechanics *does* seem like an island: change any one aspect, and the whole structure collapses. This view is buttressed by three types of results:

(1) **"Derivations" of the $|\psi|^2$ probability rule.** Gleason's Theorem [9] shows that, in a Hilbert space of dimension 3 or higher, the usual quantum probability rule is the only one consistent with a requirement of noncontextuality. Deutsch [7] and Zurek [14] derived the rule from other assumptions.

(2) **Arguments for complex amplitudes.** If $f(n)$ is the number of real parameters needed to specify an n-dimensional mixed state, then only when amplitudes are complex numbers does $f(n_A n_B) = f(n_A) f(n_B)$ (since $f(n) = n^2$). With real amplitudes, $f(n) = n(n+1)/2$ and thus $f(n_A n_B) > f(n_A) f(n_B)$. With quaternionic amplitudes, $f(n) = 2n^2 - n$ and thus $f(n_A n_B) < f(n_A) f(n_B)$. Caves, Fuchs, and Schack [6] exploited this observation to show that a "quantum de Finetti Theorem" (which justifies Bayesian reasoning) works only if amplitudes are complex. Hardy [10] also made essential use of the observation in his derivation of quantum mechanics from "five simple axioms."

(3) **"Perverse" consequences of nonlinearity.** After Weinberg [12] proposed nonlinear variants of the Schrödinger equation, Gisin [8] and Polchinski [11] independently observed that almost all such variants would allow superluminal signalling. Later Abrams and Lloyd [1] argued that a "nonlinear quantum computer" could solve NP-complete and even #P-complete problems in polynomial time.[c]

This paper won't attempt another axiomatic derivation like that of Hardy [10]—its more modest goal is just to stroll through quantum mechanics' neighborhood of theoryspace. All mathematical results in this paper

[a] A bit of pedantry: a physicist might call the neighbors of quantum mechanics I'll discuss "inconsistent," since they contradict auxiliary assumptions that the physicist considers obvious. I'll stick to milder epithets like "perverse."

[b] More pedantry: whether a theory is an island is therefore a function of our knowledge, not just of the theory itself.

[c] Abrams and Lloyd claimed furthermore that their nonlinear algorithms are robust against small errors. This claim does not withstand detailed scrutiny; whether nonlinear quantum computers can solve NP- and #P-complete problems robustly therefore remains an intriguing open problem. On the other hand, if arbitrary 1-qubit nonlinear gates can be implemented without error, then even PSPACE-complete problems can be solved in polynomial time. This result is tight, since nonlinear quantum computers can also be simulated in PSPACE. These claims will be proved in another paper.

are trivial. So why write it then? Apart from the fact that triviality never stopped a quantum philosopher before, I hope to make a point: that if you change quantum mechanics in the most obvious ways, you'll run into problems that have nothing to do with the subtleties of contextuality, locality, or entanglement. Even in "quantum mechanics lite"—where there are no mixed states, no tensor products, and no intermediate measurements, just vectors representing probabilities that get mapped to other vectors—you'll need to worry about conservation of probability, and about closure properties of the allowed vector maps.

I *won't* make this point regarding nonlinear quantum mechanics, for the simple reason that there it seems false. Contrary to what I originally thought, one can define a large, natural class of discrete norm-preserving nonlinear gates. This class includes "Weinberg gates" such as

$$W\left(\begin{array}{c} x \\ y \end{array}\right) = \left(\begin{array}{c} x \\ e^{iy}y \end{array}\right),$$

as well as "polynomial gates" such as

$$G\left(\begin{array}{c} x \\ y \end{array}\right) = \left(\begin{array}{c} x^2 - (y^*)^2 \\ 2\,\mathrm{Re}\, xy \end{array}\right).$$

Since $\|G(v)\|_2 = \|v\|_2^2$, the gate G preserves the 2-norm of v provided $\|v\|_2 = 1$. As far as I can tell, any argument for the implausibility of W or G needs to be based on *physical* effects, such as superluminal signalling or efficient solubility of NP-complete problems.

The paper is organized as follows. Section 2 shows that when $p \neq 2$, the only p-norm-preserving linear transformations are permutations of diagonal matrices. In other words, if you want to base quantum mechanics on a p-norm other than the 2-norm, then you'll need to include some sort of "manual normalization." However, manual normalization brings with it most of the hazards of nonlinearity: superluminal signalling, distinguishability of non-orthogonal states, and polynomial-time solubility of "obviously hard" problems.[d] Section 3 addresses the last point in detail, by using the concept of *postselection* to study the computational power of alternative quantum theories. The punchline, which might be of independent interest to computer scientists, is that all the alternative theories considered have at least the power of the complexity class PP,[e] and many have *exactly* the power of PP.

Finally, Section 4 gives an argument for why amplitudes are complex rather than real, that has nothing to do with the parameter-counting arguments of Refs. [5,6,10]. Unfortunately, my argument says nothing about why amplitudes are complex rather than quaternionic.

2 Other p-Norms

> "Addition in proof: More careful considerations show that the probability is proportional to the square of the [amplitude] Φ_{nrm}." —Max Born [4], in a footnote to his 1926 paper introducing the probability interpretation (the main text says the probability is proportional to Φ_{nrm} itself)

No doubt about it: the 2-norm is special. The Pythagorean Theorem, Fermat's Last Theorem, and least-squares regression all involve properties of a sum of squares that fail for a sum of cubes or of any other power. Still, given that classical probability theory is based on the 1-norm and quantum mechanics on the 2-norm, it's natural to wonder what singles out 1 and 2. What happens if we try to base a theory on the p-norm[f] for some other p? In this section I'll explain why the 2-norm is the only p-norm that permits nontrivial norm-preserving linear maps.[g]

[d] NP-complete problems are obviously hard; factoring and graph isomorphism are not.

[e] See www.cs.berkeley.edu/~aaronson/zoo.html for definitions of over 370 complexity classes.

[f] The main reason for restricting attention to p-norms is their behavior under tensor products: disregarding zany functions that depend on the Axiom of Choice, if $f(\alpha\beta) = f(\alpha) f(\beta)$ for all α, β then $f(\alpha)$ must have the form $|\alpha|^p$. However, it might be interesting to consider theories where the probability of measuring a basis state $|x\rangle$ depends on *all* amplitudes, not just that of $|x\rangle$.

[g] When $p = 0$ all linear maps are norm-preserving, but they have no effect because all outcomes of a measurement are always equiprobable. When $p = \infty$ only generalized diagonal matrices are norm-preserving, as in the case $2 < p < \infty$. I refuse even to discuss the case $p < 0$.

It's easiest to start with real amplitudes and then generalize to complex ones. We want to know which matrices $A \in \mathbb{R}^{n \times n}$ have the property that for all vectors x, $\|Ax\|_p = \|x\|_p$, where $\|\cdot\|_p$ denotes the p-norm. We can gain some intuition by counting constraints. When $p = 1$ *and* we restrict our attention to x with nonnegative entries, we obtain the set of stochastic matrices, or nonnegative matrices that satisfy n linear constraints. When $p = 2$, we obtain the set of orthogonal matrices, or those $A = (a_{jk})$ such that

$$\sum_{j=1}^{n} a_{jk} a_{kl} = \delta_{kl} \tag{1}$$

for all k, l. Equation 1 imposes $n(n+1)/2$ quadratic constraints on A, cutting the number of parameters needed to specify A roughly in half. Continuing, when $p = 3$ we expect order n^3 cubic constraints, when $p = 4$, order n^4 quartic constraints, and so on. That the number of constraints exceeds the number of parameters for $p > 2$ makes us suspect that $p = 2$ is the "end of the line."

But that's not a rigorous argument, because we know there are matrices that are norm-preserving for all p: the generalized diagonal matrices (that is, products of permutation matrices and diagonal matrices). To show that these are the only norm-preserving matrices, first let p be an even integer greater than 2. Then letting $x = (x_j)$, the requirement

$$\sum_{j=1}^{n} x_j^p = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{jk} x_k \right)^p \tag{2}$$

for all x implies that the left- and right-hand sides are identical as formal polynomials, and therefore (among other constraints) that

$$\sum_{j=1}^{n} a_{jk}^{p-2} a_{jl}^2 = \delta_{kl}$$

for all k, l. This in turn implies that for all j and $k \neq l$, either $a_{jk} = 0$ or $a_{jl} = 0$. But since every column must contain nonzero entries by the constraint $\sum_j a_{jk}^p = 1$, this implies that A is a generalized diagonal matrix.

Next let p be an odd positive integer. We claim that, so long as $x_1, \ldots, x_n$ are nonnegative, the entries of Ax never change sign. Clearly there exist $s_1, \ldots, s_n \in \{-1, 1\}$ such that

$$\sum_{j=1}^{n} x_j^p = \sum_{j=1}^{n} s_j y_j^p$$

as formal polynomials, where $y_j = \sum_{k=1}^{n} a_{jk} x_k$. Suppose by contradiction that, keeping all x_j's nonnegative, we could make $\operatorname{sgn}(y_j) s_j = -1$ for some j, where $\operatorname{sgn}(y_j)$ is 0 if $y_j = 0$ and $y_j / |y_j|$ otherwise. Then

$$\sum_{j=1}^{n} \operatorname{sgn}(y_j) y_j^p = \sum_{j=1}^{n} x_j^p = \sum_{j=1}^{n} s_j y_j^p$$

as formal polynomials, which implies that

$$\sum_{\operatorname{sgn}(y_j) s_j = -1} \operatorname{sgn}(y_j) y_j^p = 0.$$

Since every term in the above sum is nonnegative, we have $y_j^p = 0$ for all j such that $\operatorname{sgn}(y_j) s_j = -1$, which implies that $a_{jk} = 0$ for all j, k such that $\operatorname{sgn}(y_j) s_j = -1$, contradiction.

Since the entries of Ax never change sign when x is nonnegative, it follows that in each row of A, all entries have the same sign. So if we define a new matrix B by $b_{jk} = |a_{jk}|$, then B also has the property that $\|Bx\|_p = \|x\|_p$ for all x. But then when $p \geq 3$, the same reasoning from the case of even p implies that B is generalized diagonal, which implies that A was generalized diagonal as well. When $p = 1$, B is stochastic, and it is easily checked that the only stochastic matrices that preserve the 1-norm of all vectors (not just nonnegative ones) are permutation matrices.

Finally, let $p > 0$ be an arbitrary real that is not an integer. Let $\widetilde{x}_j = |x_j|^p$; then

$$\sum_{j=1}^{n} \widetilde{x}_j = \sum_{j=1}^{n} \left| \sum_{k=1}^{n} a_{jk} \widetilde{x}_k^{1/p} \right|^p$$

for all $\widetilde{x}_1, \ldots, \widetilde{x}_n$. It follows that there exist $s_1, \ldots, s_n \in \{-1, 1\}$ such that

$$\sum_{j=1}^{n} \widetilde{x}_j = \sum_{j=1}^{n} \left(s_j \sum_{k=1}^{n} a_{jk} \widetilde{x}_k^{1/p} \right)^p$$

as formal functions. But this implies that A is generalized diagonal, since otherwise the right-hand side could never be simplified to a linear function in the $\widetilde{x}_j$'s.

So much for real amplitudes. When we generalize to complex amplitudes $x_j \in \mathbb{C}$, there are two defensible choices: letting $x_j = \alpha_j + i\beta_j$, we could require either $\sum_{j=1}^{n} (|\alpha_j|^p + |\beta_j|^p) = 1$ or $\sum_{j=1}^{n} |x_j|^p = 1$, where $|x_j| = \sqrt{\alpha_j^2 + \beta_j^2}$ as usual. Under the first choice, we can consider $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ as a vector of $2n$ reals and A as a $2n \times 2n$ matrix; then the results from the real-amplitude case immediately imply that A is generalized diagonal. Under the second choice, we can choose an $x_l \neq 0$ and replace it by $e^{i\theta} x_l$, holding all other x_k's fixed. Then since $\sum_{j=1}^{n} |x_j|^p$ remains constant as we vary θ,

$$\sum_{j=1}^{n} |y_j|^p = \sum_{j=1}^{n} \left| a_{jl} e^{i\theta} x_l + \sum_{k \neq l} a_{jk} x_k \right|^p$$

must also remain constant. But when $p \neq 2$, this is possible only if for all j, either $a_{jl} = 0$ or $\sum_{k \neq l} a_{jk} x_k = 0$. Intuitively, once we sneak the 2-norm in "through the back door" in defining the norm of a complex number, consistency forces us to use it everywhere. We omit a proof of this fact, since it follows easily from a case analysis similar to that for real amplitudes.

Stepping back, what can we say about why the 2-norm is special? The standard answer—that the 2-norm is special because it's preserved under rotations—merely pushes the question from quantum mechanics back to the Pythagorean Theorem. The latter might be thought a good enough place to stop. However, although the Pythagorean Theorem dates back some 3800 years, I confess to having never *understood* it at a gut level. So if pressed, I'd instead answer the question as follows: values of p other than positive even integers are almost nonstarters, since we want $|x|^p$ to be defined and smooth at $x = 0$. But when $p = 4, 6, 8, \ldots$, Equation 2 involves terms of the form $(a_{jk} x_k)^q (a_{jl} x_l)^{p-q}$ where q and $p - q$ are both positive even integers, and that immediately forces A to be generalized diagonal. So all that's left is $p = 2$.

If you still want to define quantum mechanics using a p-norm where $p \neq 2$, the only option seems to be *manual normalization*. This means that when a state $|\psi\rangle = \sum_x \alpha_x |x\rangle$ is measured in the standard basis, the probability of outcome $|x\rangle$ is $|\alpha_x|^p / \sum_y |\alpha_y|^p$. Since keeping $|\psi\rangle$ normalized is no longer imperative, three options present themselves for how $|\psi\rangle$ evolves:

(i) As usual, $|\psi\rangle$ can be mapped to $U|\psi\rangle$ where U is any unitary matrix.

(ii) $|\psi\rangle$ can be mapped to $A|\psi\rangle$ where A is any invertible matrix.

(iii) $|\psi\rangle$ can be mapped to $A|\psi\rangle$, but then *local normalization* is performed on the subsystem acted on by A.

To illustrate option (iii), suppose the nonunitary gate

$$\begin{pmatrix} q & r \\ s & t \end{pmatrix}$$

is applied to the second qubit of the normalized state $\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$. Then the unnormalized result is

$$(q\alpha + r\beta)|00\rangle + (s\alpha + t\beta)|01\rangle + (q\gamma + r\delta)|10\rangle + (s\gamma + t\delta)|11\rangle,$$

so the locally normalized result is

$$\frac{\sqrt{\alpha^2+\beta^2}}{\sqrt{(q\alpha+r\beta)^2+(s\alpha+t\beta)^2}}\left[(q\alpha+r\beta)\left|00\right\rangle+(s\alpha+t\beta)\left|01\right\rangle\right]+$$
$$\frac{\sqrt{\gamma^2+\delta^2}}{\sqrt{(q\gamma+r\delta)^2+(s\gamma+t\delta)^2}}\left[(q\gamma+r\delta)\left|10\right\rangle+(s\gamma+t\delta)\left|11\right\rangle\right],$$

in contrast to the globally normalized result of

$$\frac{(q\alpha+r\beta)\left|00\right\rangle+(s\alpha+t\beta)\left|01\right\rangle+(q\gamma+r\delta)\left|10\right\rangle+(s\gamma+t\delta)\left|11\right\rangle}{\sqrt{(q\alpha+r\beta)^2+(s\alpha+t\beta)^2+(q\gamma+r\delta)^2+(s\gamma+t\delta)^2}}.$$

So, what's wrong with these prescriptions? Nothing, as long as you can stomach the following:

(1) Distinguishability of non-orthogonal states. Here's how to distinguish $d=\Omega\left(\sqrt{p}\right)$ states of a single qubit with constant probability of error, under option (i) (and therefore under (ii) and (iii) as well). Let the j^{th} state be $\left|\psi_j\right\rangle=\cos\left(\pi j/d\right)\left|0\right\rangle+\sin\left(\pi j/d\right)\left|1\right\rangle$ where $j\in\{0,\ldots,d-1\}$. Apply a $d\times d$ unitary matrix to $\left|\psi_j\right\rangle$ whose first two columns are

$$\left(\begin{array}{c}\cos\left(\pi 0/d\right)/\sqrt{d}\\ \vdots\\ \cos\left(\pi\left(d-1\right)/d\right)/\sqrt{d}\end{array}\right),\left(\begin{array}{c}\sin\left(\pi 0/d\right)/\sqrt{d}\\ \vdots\\ \sin\left(\pi\left(d-1\right)/d\right)/\sqrt{d}\end{array}\right).$$

Then measure in the standard basis. Suppose without loss of generality that $j=0$ and that d is odd; then the probability of any outcome other than 0 being measured is $q/\left(q+1\right)$ where

$$\begin{aligned}q&=2\sum_{k=1}^{(d-1)/2}\left|\cos\left(\frac{\pi k}{d}\right)\right|^p\\&\leq 2\sum_{k=1}^{(d-1)/2}\left(1-\frac{\left(\pi k/d\right)^2}{2}+\frac{\left(\pi k/d\right)^4}{24}\right)^p\\&\leq 2\sum_{k=1}^{(d-1)/2}\left(1-\frac{\pi^2k^2}{4d^2}\right)^{2p}\\&\leq 2\sum_{k=1}^{(d-1)/2}\exp\left(-\frac{\pi^2k^2p}{2d^2}\right)\end{aligned}$$

which is bounded away from 1 so long as $p\geq cd^2$ for some constant c. It would be interesting to obtain bounds on how many states can be reliably distinguished in higher-dimensional Hilbert spaces.

(2) Superluminal signalling. Under option (ii), given an EPR pair, Alice can communicate a bit to Bob by mapping $\left|00\right\rangle+\left|11\right\rangle$ to either $\left|00\right\rangle+\varepsilon\left|11\right\rangle$ or $\varepsilon\left|00\right\rangle+\left|11\right\rangle$. Indeed, using the ideas from part (1), she can communicate $\Omega\left(\sqrt{p}\right)$ bits to Bob using a single EPR pair! I conjecture this is tight. Under options (i) and (iii), Alice can communicate a bit to Bob given enough EPR pairs, by taking advantage of Bob's ability to distinguish nonorthogonal states. Note that under options (ii) and (iii), superluminal signalling is possible even when $p=2$.

(3) Efficient solubility of NP-complete and even harder problems. Suppose you're given a Boolean function $f:\{0,1\}^n\rightarrow\{0,1\}$. Under option (ii), first prepare $\sum_x\left|x\right\rangle\left|f\left(x\right)\right\rangle$, then apply the nonunitary gate

$$G=\left(\begin{array}{cc}2^{-2n}&0\\0&1\end{array}\right) \tag{3}$$

to the f register and measure to learn whether there exists an x such that $f\left(x\right)=1$. Indeed, Section 3 shows that under options (i), (ii), and (iii), you could solve even PP-complete problems in polynomial time, which are believed to be harder than NP-complete problems.

(4) Singularity. Under options (ii) and (iii), the matrix A could be arbitrarily close to a non-invertible matrix, which can map nonzero states to the zero state.

3 Quantum Computing With Postselection

This section can be skipped by physicists with no interest in computational complexity.[h] Its goal is to show that, if you change quantum mechanics in any of three ways, then the class of problems efficiently solvable on a quantum computer expands drastically, from BQP (Bounded-Error Quantum Polynomial-Time) to PP (Probabilistic Polynomial-Time). Here PP is a well-studied classical complexity class, consisting of all decision problems for which there exists a probabilistic polynomial-time Turing machine that accepts with probability at least $1/2$ if the answer is "yes," and with probability less than $1/2$ if the answer is "no." The three changes that would give quantum computers the power of PP are: replacing the 2-norm by the p-norm for any $p \neq 2$,[i] allowing arbitrary invertible matrices instead of just unitary matrices, or allowing postselection on measurement outcomes. Any *combination* of these changes would also yield PP. Note, however, that I always assume global normalization (corresponding to options (i) and (ii) in Section 2).

It will be convenient to define a new complexity class:

Definition 1 PostBQP *(or* BQP *with postselection) is the class of languages L for which there exists a uniform family of polynomial-size quantum circuits such that for all inputs x,*

(i) At the end of the computation, the first qubit has a nonzero probability of being measured to be $|1\rangle$.

(ii) If $x \in L$, then conditioned on the first qubit being $|1\rangle$, the second qubit is $|1\rangle$ with probability at least $2/3$.

(iii) If $x \notin L$, then conditioned on the first qubit being $|1\rangle$, the second qubit is $|1\rangle$ with probability at most $1/3$.

Intuitively, postselection means that at some point in the computation, you can measure a qubit that has a nonzero probability of being $|1\rangle$, and *assume* that the outcome will be $|1\rangle$ (or equivalently, discard all runs where the outcome is $|0\rangle$). Just as Bernstein and Vazirani [3] showed that intermediate measurements don't increase the power of ordinary quantum computers, so it's easy to show that intermediate postselection steps don't increase the power of PostBQP (since these steps can all be deferred to the end). On the other hand, if operations can be performed conditioned on measurement outcomes, then mixing postselection and measurement steps *could* increase the power of PostBQP.

In the remainder of the section, I'll first show that PostBQP = PP (Theorem 2), and then use that result to prove that the other changes also give quantum computers the power of PP.

Theorem 2 PostBQP = PP.

Proof. The inclusion PostBQP $\subseteq$ PP follows easily from the techniques used by Adleman, DeMarrais, and Huang [2] to show that BQP $\subseteq$ PP.

For the other direction, let $f : \{0,1\}^n \rightarrow \{0,1\}$ be a Boolean function and let $s = |\{x : f(x) = 1\}|$. Then we need to decide in PostBQP whether $s < 2^{n-1}$ or $s > 2^{n-1}$. (As a technicality, we can guarantee using padding that $s > 0$ and $s \neq 2^{n-1}$.) The algorithm is as follows: first prepare $2^{-n/2} \sum_{x \in \{0,1\}^n} |x\rangle |f(x)\rangle$. Then following Abrams and Lloyd [1], apply Hadamard gates to all n qubits in the first register and postselect[j] on that register being $|0\rangle^{\otimes n}$, to obtain $|0\rangle^{\otimes n} |\psi_s\rangle$ where

$$|\psi_s\rangle = \frac{(2^n - s)|0\rangle + s|1\rangle}{\sqrt{(2^n - s)^2 + s^2}}.$$

[h] The rest of paper can be skipped by computational complexity theorists with no interest in physics.

[i] If p is not a positive even integer, then the power increases *at least* to PP and possibly further.

[j] Actually postselection is overkill here, since the first register has at least $1/4$ probability of being $|0\rangle^{\otimes n}$.

Next, for some positive real α, β to be specified later, prepare $\alpha |0\rangle |\psi_s\rangle + \beta |1\rangle |\phi_s\rangle$ where

$$|\phi_s\rangle = \frac{2^n |0\rangle + (2^n - 2s) |1\rangle}{\sqrt{2 (2^n - s)^2 + 2s^2}}$$

is the result of applying a Hadamard to $|\psi_s\rangle$. Postselecting on the second qubit being $|1\rangle$ then yields the state

$$\left|\varphi_{s,\beta/\alpha}\right\rangle = \frac{\alpha s |0\rangle + \sqrt{1/2}\beta (2^n - 2s) |1\rangle}{\sqrt{\alpha^2 s^2 + \beta^2 (2^n - 2s)^2 / 2}}$$

in the first qubit. A simple calculation now reveals that if $s < 2^{n-1}$, then there exists an integer i in the range $[-n, n]$ such that

$$\left|\langle +|\varphi_{s,2^i}\rangle\right| \geq \frac{1+\sqrt{2}}{\sqrt{6}} \approx 0.986$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. If $s > 2^{n-1}$, on the other hand, then for all such i we have $\left|\langle +|\varphi_{s,2^i}\rangle\right| \leq 1/\sqrt{2}$. So by running the whole algorithm $n(2n+1)$ times in parallel, with n invocations for each integer $i \in [-n, n]$, we can learn whether $s < 2^{n-1}$ or $s > 2^{n-1}$ with exponentially small probability of error. ■

Let $\mathsf{BQP}_{\text{nu-global}}$ be the class of problems solvable by a uniform family of polynomial-size, bounded-error quantum circuits, if the circuits can consist of arbitrary invertible gates, not just unitary gates. Option (ii) from Section 2 is used for normalization; that is, before a measurement, the amplitude α_x of each basis state $|x\rangle$ is divided by $\sqrt{\sum_y |\alpha_y|^2}$.

Proposition 3 $\mathsf{BQP}_{\textit{nu-global}} = \mathsf{PP}$.

Proof. The inclusion $\mathsf{BQP}_{\text{nu-global}} \subseteq \mathsf{PP}$ follows easily from Ref. [2]. For the other direction, by Theorem 2 it suffices to observe that $\mathsf{PostBQP} \subseteq \mathsf{BQP}_{\text{nu-global}}$. To postselect on a qubit being $|1\rangle$, simply apply the nonunitary gate G from Equation 3. ■

Define $\mathsf{BQP}_{\text{nu-local}}$ similarly to $\mathsf{BQP}_{\text{nu-global}}$, except that after every gate G, option (iii) (local normalization) is applied to the qubits acted on by G. Assume that arbitrary 1- and 2-qubit gates are available to polynomially many bits of precision.

Proposition 4 $\mathsf{PP} \subseteq \mathsf{BQP}_{\textit{nu-local}} \subseteq \mathsf{PSPACE}$.

Proof. For $\mathsf{PP} \subseteq \mathsf{BQP}_{\text{nu-local}}$, observe that in the proof of Theorem 2, the only essential postselection steps are applied to 2-qubit pure states unentangled with anything else. For 2-qubit gates acting on these states, local normalization is the same as global normalization.

For $\mathsf{BQP}_{\text{nu-local}} \subseteq \mathsf{PSPACE}$, let $\alpha_x^{(t)}$ be the amplitude of basis state $|x\rangle$ at time t. Then for all x, t we can write $\alpha_x^{(t)}$ as a function of $\alpha_{y_1}^{(t-1)}, \ldots, \alpha_{y_k}^{(t-1)}$ for some constant k and basis states $|y_1\rangle, \ldots, |y_k\rangle$. This immediately implies a depth-first recursive algorithm (using a polynomial amount of memory) for approximating any amplitude $\alpha_x^{(t)}$ to polynomially many bits of precision. ■

Finally, for any nonnegative real number p, define BQP_p similarly to BQP, except that the probability of measuring a basis state $|x\rangle$ equals $|\alpha_x|^p / \sum_y |\alpha_y|^p$. (Thus $\mathsf{BQP}_2 = \mathsf{BQP}$.) All gates are unitary.

Proposition 5 $\mathsf{PP} \subseteq \mathsf{BQP}_p \subseteq \mathsf{P}^{\#\mathsf{P}}$ *for all constants* $p \neq 2$, *and* $\mathsf{BQP}_p = \mathsf{PP}$ *provided* p *is an even integer greater than* 2.

Proof. The inclusion $\mathsf{BQP}_p \subseteq \mathsf{P}^{\#\mathsf{P}}$ is obvious. To simulate BQP_p in PP when p is a positive even integer, use the techniques of Ref. [2] (which handle the $p = 2$ case), but evaluate polynomials of degree p instead of quadratic polynomials. To simulate PP in BQP_p when $p \neq 2$, run the algorithm of Theorem 2, having initialized $O\left(n^3 / |p-2|\right)$ ancilla qubits to $|0\rangle$. To postselect on the b^{th} qubit being $|1\rangle$: if $p < 2$, then apply Hadamards to $10pn/(2-p)$ ancilla qubits conditioned on the b^{th} qubit being $|1\rangle$. If $p > 2$, then apply Hadamards to $10pn/(p-2)$ ancilla qubits conditioned on the b^{th} qubit being $|0\rangle$. ■

4 Real Amplitudes

> "C'mon, they're algebraically closed!" —A math graduate student, when asked why God would resort to complex numbers in creating quantum mechanics

To a beginner, perhaps the most unexpected fact about quantum mechanics is that amplitudes are complex. As the term 'imaginary' suggests, we tend to think of complex numbers as (useful) human inventions; it's unsettling if the source code of the Universe is best written in a language like Fortran with a complex-number data type. Also, in contrast to what we saw in Section 2, restricting amplitudes to be real doesn't lead to a theory obviously very different from quantum mechanics. All the greatest hits are still there: interference, entanglement, Bell inequality violations, noncommuting observables, non-unique decompositions of mixed states, universal quantum computing, the Zeno effect, the Gleason and Kochen-Specker theorems.

Nevertheless, Section 1 recalled a subtle difference between complex and real (or for that matter complex and quaternionic) amplitudes, based on counting the number of parameters of a mixed state. This section gives a completely different argument for why amplitudes aren't real. The advantage of this argument is that it's elementary and intuitive; the disadvantage is that it says nothing about why amplitudes are complex rather than quaternionic.

Let $\mathcal{S}$ be a set of states, and let $\mathcal{U}$ be a set of transformations from $\mathcal{S}$ to itself. Say $\mathcal{T}$ has the *square root property* if for all $U \in \mathcal{U}$, there exists another transformation $V \in \mathcal{U}$ such that $V(V(S)) = U(S)$ for all $S \in \mathcal{S}$. If time is continuous, then the importance of the square root property is obvious: without it there are transformations that can't be interpreted as the result of applying a fixed Hamiltonian for some interval of time. Even if time is discrete, the square root property is desirable, because it allows any U that acts over k time steps to be approximated by V^k for some V that acts over a single time step.[k] Clearly quantum mechanics has the square root property: given a unitary U, let $|\psi_1\rangle, \ldots, |\psi_n\rangle$ be the eigenvectors of U and let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues; then there exists a unitary V with eigenvectors $|\psi_1\rangle, \ldots, |\psi_n\rangle$ and eigenvalues $\mu_1, \ldots, \mu_n$ such that $\mu_j^2 = \lambda_j$, which therefore satisfies $V^2 = U$. Since every quaternion has a square root,[l] the same argument shows that quaternionic quantum mechanics has the square root property.

However, real quantum mechanics doesn't have the square root property. This is immediate since orthogonal matrices such as

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

with determinant -1 can't be written as squares of matrices with real determinants. If we want to restore the square root property, then we have two choices. The first choice is to restrict to the group $\mathrm{SO}(n)$—that is, to real orthogonal matrices with determinant 1. It's not hard to see that for every $U \in \mathrm{SO}(n)$, there exists a $V \in \mathrm{SO}(n)$ such that $V^2 = U$. On the other hand, natural 1-qubit operations such as the above two can only be implemented by using ancillia qubits. The second choice is to allow the "square root" of U to have larger dimension than U. For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

contains the 1-qubit phase flip as a 2×2 submatrix. This is an instance of a well-known geometrical fact, that a mirror reversal in n dimensions can be accomplished by a rotation in $n+1$ dimensions. Indeed, *any* $n \times n$ orthogonal matrix U has a real square root of dimension $(n+1) \times (n+1)$, since there exists an element of $\mathrm{SO}(n+1)$ that contains U as a submatrix. With either choice, the price we pay is that our n-dimensional theory can be fully described only in $n+1$ dimensions. But the $(n+1)$-dimensional theory requires $n+2$ dimensions to describe, and so on ad infinitum—unless we declare that the $(n+1)^{st}$ dimension is physically different from dimensions 1 to n.

[k] To write U exactly as V^k we'd need "a k^{th} root property," which also holds for quantum mechanics but fails for real quantum mechanics.

[l] Indeed some, such as -1, have infinitely many square roots.

Acknowledgments

I thank Andrei Khrennikov for including this paper in the Växjö proceedings despite its having no relation to anything I talked about at the conference; and Chris Fuchs, without whom this "saucy paper" (his words) wouldn't have been written. Supported by an NSF Graduate Fellowship, by NSF ITR Grant CCR-0121555, and by the Defense Advanced Research Projects Agency (DARPA).

References

1. D. S. Abrams and S. Lloyd. Nonlinear quantum mechanics implies polynomial-time solution for NP-complete and #P problems, *Phys. Rev. Lett.* 81:3992–3995, 1998. quant-ph/9801041.
2. L. Adleman, J. DeMarrais, and M.-D. Huang. Quantum computability, *SIAM J. Comp.* 26(5):1524–1540, 1997.
3. E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing* 26(5):1411–1473, 1997.
4. M. Born. Zur quantenmechanik der stoßvorgänge, *Zeitschrift für Physik* 37:863–867, 1926. English translation in *Quantum Theory and Measurement* (J. A. Wheeler and W. H. Zurek, eds.), Princeton, 1983, pp. 52–55.
5. C. M. Caves, C. A. Fuchs, and P. Rungta. Entanglement of formation of an arbitrary state of two rebits, *Foundations of Phys. Lett.* 14:199–212, 2001. quant-ph/0009063.
6. C. M. Caves, C. A. Fuchs, and R. Schack. Unknown quantum states: the quantum de Finetti representation, *J. Math. Phys.* 45:4537, 2002. quant-ph/0104088.
7. D. Deutsch. Quantum theory of probability and decisions, *Proc. Royal Soc.* A455:3129–3137, 1999. quant-ph/9906015.
8. N. Gisin. Weinberg's non-linear quantum mechanics and superluminal communications, *Phys. Lett. A* 143:1–2, 1990.
9. A. M. Gleason. Measures on the closed subspaces of a Hilbert space, *J. Math. Mech.* 6:885–893, 1957.
10. L. Hardy. Quantum theory from five reasonable axioms, manuscript, 2003. quant-ph/0101012.
11. J. Polchinski. Weinberg's nonlinear quantum mechanics and the Einstein-Podolsky-Rosen paradox, *Phys. Rev. Lett.* 66:397–400, 1991.
12. S. Weinberg. Precision tests of quantum mechanics, *Phys. Rev. Lett.* 62:485, 1989.
13. S. Weinberg. *Dreams of a Final Theory*, Vintage, 1994.
14. W. H. Zurek. Environment-assisted invariance, causality, and probabilities in quantum physics, *Phys. Rev. Lett.* 90:120404, 2003. quant-ph/0211037.