Church's thesis meets quantum mechanics

Warren D. Smith
WDSmith@fastmail.fm

Feb. 1999; revised April 2002

Abstract — **"Church's thesis" is the notion that any "reasonable physical system" may be "simulated" by a Turing machine (TM). The "strong" Church's thesis adds "...with at most polynomial slowdown." The (new) "intermediate" Church's thesis instead says "...with at most polynomial amplification of memory-space requirements." (All the terms in quotes need to be defined.) With any particular set of physical laws: is Church's thesis true?**

Church's theses are central issues for both physics and computer science. With the precise specification of some set of laws of physics, and a precise definition of "simulation" and "polynomial," Church's theses become susceptible to mathematical proof or disproof. A previous report essentially settled the question for classical mechanics. Recent theoretical investigations of "quantum computers" make it look likely that the strong Church thesis is *false* in linear quantum mechanics, and indeed even in some models of *open* quantum systems.

In the present paper, I show (at least, if one adopts my assumptions and definitions – and there are a large number of them) that the weak Church's thesis is *true* for nonrelativistic quantum mechanics with well-enough behaved interparticle potentials, e.g. Coulombic. I.e., *quantum mechanics is simulable.* We give a simulation algorithm. If the simulation is performed by a quantum computer rather than a conventional one, then the slowdown is only polynomial. In other words, even Church's strong thesis becomes true if "TM" is replaced by "quantum TM." With a conventional TM, we then automatically get the intermediate thesis; and if the initial quantum state is represented non-sparsely, (i.e. in a format in which exponentially many complex amplitudes are specified) then the simulation of quantum time evolution actually runs in quasipolynomial time with respect to that input length. On the other hand, QM is *not* algorithmic, nor even self-consistent, in the presence of point *magnetic* dipoles.

The proof strategy involves (1) defining what "simulation" and "reasonable physical system" should be. (2) showing that "regularizing" the potential introduces acceptably small error. (For Coulomb potentials, the most natural regularization procedure is to replace "point" charges by uniform distributions of charge within small balls centered at the point.) (3) Showing how a quantum computer can approximately evaluate Feynman path integrals with phase factor integrands corresponding to regularized potentials. (4) Obtaining effectively computable error bounds for this approximation. (5) Finally, the quantum computer is simulated by a conventional computer.

A different method, based on Rayleigh-Ritz approximate eigenfunctions, also seems to yield Church's intermediate thesis. (Treated in an appendix.) Although this method is conceptually simpler, it apparently does not lead to efficient algorithms for a quantum computer. But it does yield a proof that the spectral energies of quantum bound systems form a computable real sequence.

Keywords — Church's thesis, algorithmicity of quantum mechanics, quantum N-body problem, rigorous physics, uncertainty principle, quantum computers, Feynman path integrals, symplectic integration, exponential splitting formulas, Trotter product formula, BQP, Rayleigh Ritz variational method, computable real numbers, operator computability theory, potential singularities, Coulomb's law, analyticity.

Contents

glyph	meaning	§
⊗	Wrong choice; impatient readers may skip	1.1
♣	Defn of nearby term is in appendix A	1
X^*	complex conj. or Hermitian adjoint of X	
$\overline{v}$	overbar denotes expectation value	
C	energy cutoff	6.1
d	#dimensions (degrees of freedom)	3
E_K	kinetic energy	2
E_P	potential energy	6
E	total energy $E = E_K = E_P$	6
g	2^g =#gridlines in each direction	1.2
H	sidelength of containing box	1.2
m	particle mass(es); often the maximum	1.2
n	#timestepping stages	1.2
N	#particles	1.2
N_Q	#qubits in description of initial state	
P	#decimal places of accuracy	2.5
t	timespan system is simulated for	1.2
R	run # (in rerun sequence of simulations)	1.2
V	potential energy function	1.2
V^{reg}	regularized potential energy function	6.1
$\|f\|_p$	L_p norms; if no p then $p = 2$ implied	A

Table 1: Important symbols used in this paper and where they are defined/discussed. Also c, α, m_e, e, h, $\hbar$, G, ϵ_0, μ_e, k_B, λ_C denote standard physical constants following [19].

1 Introduction

Algorithmicity is the most important concept in computer science. Quantum mechanics is the most important branch of physics. So the question of whether the two can be married is important. The purpose of this paper is to show that the answer is "yes": *quantum mechanical systems are algorithmically simulable* (at least, using the simplest versions of quantum mechanics). In contrast, a previous report [83] showed that classical n-body systems under Newton's laws are *not* algorithmically simulable. Thus (surprisingly) quantum mechanics is easier than classical mechanics.

The precise statement of our theorem is necessarily complicated, and even the task of formulating the right conjecture is difficult (e.g. since the notion of "simulation" needs to be precisely defined, and since quantum mechanics is not deterministic). Also, the proof is very difficult (since rigorous quantum mechanics is difficult).

Appendix A explains some of the notation and terminology used in this paper ♣[1].

For discussion of the history of Church's thesis and the motivation for working on it, see [86]. Several topics which are not essential for the simplest version of the proof, but which probably will lead to improvements (e.g. in running time) are delegated to appendices.

1.1 *Outline of the proof*

Here is a quick overview of our results and techniques, and a sketch of the remainder of the paper. The main result is stated in §11. The proof argument involves 7 steps.

1. Quantum mechanics is written in Feynman's path-integral formulation [29] (§7).

2. The potential functions are "regularized" (using "high energy cutoffs"). In other words, we solve a different, and nicer, quantum mechanics problem, than the one we were given. If the form and height of the cutoff are chosen correctly, we show (in §6), partly with the aid of Lieb & Thirring's reformulation of the "uncertainty principle" as an integral "Sobolev inequality," and partly with the aid of flux bounds) that the error introduced by this step may be bounded and is sufficiently small for our purposes. (Since the phrase "may... purposes" at the end of the previous sentence arises often, we abbreviate it by "isn't too large.")

3. The path integral is approximated by finite dimensional integrals (§7,§4). For regularized potentials and if enough dimensions are used, we show (with the aid of operator identities of H.Suzuki; see §5) that the error introduced by this step also isn't too large.

4. We use an optimal tradeoff to select the cutoff in step 2. If the cutoff is made too high, then the error bounds in step 3 become weak, forcing the simulation to consume too much time. On the other hand if the cutoff is made too low, the regularization itself introduces too much error. The best choice equalizes the two kinds of error bounds.

5. The finite dimensional integral is approximated by a Riemann sum. Using standard techniques (§3), the error introduced by this step also isn't too large.

6. The sum above may be computed by a "quantum algorithm[2]" suffering only "polynomial slowdown" (§8). Steps 3 and 5 thus constitute a quantum algorithm for certain kinds of Feynman path integrals. We demonstrate (§9.5) that the error introduced by the use of finite precision arithmetic in this algorithm isn't too large. This demonstrates (our form of) Church's strong thesis if the simulator is a "quantum computer."

7. Finally, by simulating the quantum computer with an ordinary computer (trivially accomplished if one is willing to suffer an exponential slowdown; §10) we demonstrate Church's weak thesis. (Also, we could just evaluate the Riemann sums directly instead of simulating a quantum computer doing it.) This also shows that the *space* usage needs only to be polynomially ♣ large, despite the exponential ♣ slowdown in *time* consumption. That demonstrates the intermediate Church thesis.

Please keep this master plan in mind as you read the various sections, which by themselves might not seem to be moving clearly toward the overall goal. Otherwise reading this paper might become a depressing experience of pawing through 20 pages unrelated-seeming mathematics and annoying argumentation (verging on philosophy and often concerning the *wrong* choices of methods or definitions). The best way to avoid that depression is to realize that getting the right quantum formulation of Church's thesis is a very tricky task all by itself, requiring considering numerous options, most of them wrong. For the reader's convenience we employ the symbol ⊗ to mark choices I consider wrong. The impatient reader will then be able to skip the wrong choices and the argumentation about them and proceed directly to the right choices. Once the formulation (§2) is over, we then consume a large number of pages for the task of gradually building up the tools needed for the proof from different areas, in preparation for the final crescendo in §11.

1.1.1 *Remarks on the proof outline*

A. Our argument remains valid regardless of whether anybody will ever be able to build a "quantum computer."

[1] The symbol ♣ will serve as a reminder that the reader may want to consult appendix A for the definition of a nearby notation or word.

[2] I.e., an algorithm running on a "quantum computer" [81] ♣.

B. The validity of several of the proof steps depends on the potentials being well-enough behaved (at least, after regularization – which means potential singularities were not too nasty before regularization; see rule 5 in §1.2) and the assumption of bounded energy for the initial state (and hence by conservation of energy, also at all times, §6.2) being simulated, and the assumption that the simulator's runtime, for "fairness," is allowed to depend polynomially ♣ upon this energy.

C. The proof sketch above can be simplified greatly (by omitting some steps entirely) if the reader is willing to grant various extra assumptions:

- First, if the interparticle potentials are *bounded*, then there is no need for *any* regularization step. And indeed appendix §B gives an argument suggesting that negative-infinite Coulomb $1/r$, $r \to 0$ singularities really do not exist, or are best treated as though they do not.

- Second, if time and space are discrete not continuous, for example with "grain size" being the Planck length and Planck time (§9.2, §9.3), then various integrals, say, are really sums and we do not have to worry about approximating integrals by sums, since we always had sums.

But I disparage these approaches, especially the second, as cop-outs, for several reasons. First, the Planck length 1.6×10^{-35} meter and Planck time 5.4×10^{-44} second are so small that the computer science view of them as "constant factors" is dubious. Similarly, the anti-Coulomb arguments of §B also only apply at very small length scales. If we are going to view such numbers as mere "constant factors," we might as well take the view that the universe is just a finite state machine (with some very large "constant" number of states) and hence Church's thesis is a triviality. That view totally misses the point of why quantum computers might be interesting and trivializes the whole topic of simulating physics.

My attitude is that our mission is to *stay within* one particular mathematical model of physics and see what are the computational consequences of that. For us, that theory will be the simplest kind of quantum mechanics, i.e. Schrödinger's multiparticle equation, no quantum fields, scalar potentials only, no relativity, no decoherence. If you argue that this theory of physics is not *the* theory of physics (e.g. it is invalid below the Planck scale or at relativistic energies) and therefore quit trying to simulate it, and quit trying to see what it says in the regimes where it is supposedly invalid, then you have abandoned your mission! *It is not within the purview of the simulator to make such judgements* – the simulator is merely in the employ of those who can try to make those judgements later after seeing the simulator's output. Really, if you were going to argue against me, then your only self-consistent stance would be to abandon using Schrödinger's equation entirely and indeed abandon doing all physics until such time as a "final theory" arrives. Why? Because there are small length scales everywhere. Everything is ϵ away from everything else. Quantum gravity happens. How can you *prove* that this does not somehow lead to effects which destroy the validity of Schrödinger's equation even in situations that look innocent? What does "look innocent" mean exactly? How can you *prove* that discretizing spacetime at the Planck scales only introduces error comparable to the errors you get by ignoring the (unknown) laws of quantum gravity that hold at those scales? How can you precisely draw the boundary between when Schrödinger's equation is valid and when it is not, if that boundary depends on some *unknown* physical theory? You cannot! So, you'd have to view everything as suspect. A much more productive point of view is to demand that we simulate Schrödinger's equation *regardless* of its supposed physical validity (*not* copping out) and see what happens.

This is not to express complete disapproval of the Planck scale arguments in §9.2, §9.3 or the anti-singularity arguments in appendix §B. Nevertheless, even if one supports those arguments, one still must admit that it is very comforting that we can simulate quantum mechanics without relying on them. Then, we still have the option open to us of employing such assumptions later – if we later can achieve confidence that their use is justified in certain situations – to get speedups.

D. Actually, although the tradeoff view in step 4 works, it is simpler for the algorithm designer to view the situation differently: Given an accuracy target, one determines the minimum height the cutoff has to be raised to, using theorem 14. One then determines the minimum number n of timesteps needed in the simulation by using EQ 30.

E. I have not examined the question of whether all this can be carried through also in the presence of vector potentials (e.g. magnetic fields; see §11.1) and for relativistic quantum mechanics. Probably it can, but it will be more difficult. The reason I make these speculations is that my methods are spiritually related to the methods previously used by E.Lieb and collaborators to prove the "stability of matter" [63]. Eventually those workers were able to handle arbitrary static magnetic fields of bounded energy and a caricature[3] of relativity, but it was considerably more difficult. Indeed, they showed that matter is only relativistically stable due to numerical *luck*, i.e. the fact that the "fine structure constant" is $\approx 1/137$. If this value had instead been, say, 6, then matter would be unstable, and quantum mechanics would have been extremely different, perhaps not even well posed.

Also, Kato [47] had proved the self-adjointness of extensions of the Schrödinger operators (and hence, essentially, the existence and uniqueness of solutions to Schrödinger's equation) corresponding to $L^2 + L^\infty$ scalar potentials. Kato [49] was later able to extend this result

[3] Quote from Lieb-Yau 1988 [Phys.Rev.Lett. 61 p.1696; reprinted in [63]]: "We are not aware of the existence of any truly relativistic formulation of many-body quantum mechanics and among the many possible caricatures of such a theory we have adopted the simplest possible one..."

to allow a large class of magnetic fields. See [75] vol. 2 (e.g. theorem X.22 and example 4 in X.3) for more such results. This history again suggests that it ought to be possible to extend our results to allow magnetic fields.

F. The question of exactly *how efficiently* we can simulate is now an important question for the future, and some ideas relevant to answering it are discussed in appendices. In the main body of the present paper, however, we will not pay much attention to getting precise performance bounds; merely distinguishing between polynomial ♣ and super-polynomial will be precise enough for us.

G. The algorithms we propose in this paper are very unlikely to be of practical interest for simulating even moderately large quantum systems, *unless* quantum computers are constructed.

1.2 More detailed sketch; and list of "the rules"

Section 2 discusses the subtle issues involved in defining just what it should mean to "simulate" a quantum system, and just what it should mean to say that that system is "physically reasonable."

Among other things, we conclude that

<u>The rules:</u>

1. The simulator should have a random bit generator[4].

2. The goal of the simulation should be to provide Monte Carlo *samples* from probability distributions arising from (the approximate) quantum wave functions of the simulated system. Before sampling, we may demand that the wavefunctions be re-expressed in any (perhaps unusual) basis; this basis transformation should be "quantum computable." (In particular, the eigenbases of the operators corresponding to any polynomial combination of position and momentum operators, e.g. angular momentum about some axis, are allowed.)

3. The accuracy (number of decimal places) of the simulation should be an input parameter P (see §2.5). Simulations accurate enough so that the samples they output are statistically indistinguishable (see rule 7 and §2.3) from the true behavior of the system, are regarded as valid.

4. Assuming a quantum computer performs the simulation, then the initial state of the simulated system has to be specified as a description consisting of a finite number N_Q of qubits (see §2.6, §2.7 for discussion).

5. "Reasonable" physical systems have a finite number N of particles[5], a simply described Hamiltonian, a finite expectation value E for the energy[6], a finite maximum particle mass m, a finite maximum charge, and are contained in a finite size box.

6. The interparticle potentials have to be "nice" ones. Any potential V obeying these conditions will do[7]:

   (a) V and its first 2 derivatives are efficiently computable[8],

   (b) V is bounded above $-C_1 r_{\min}^{-1.19}$ and below $C_2 r_{\min}^{-1.99}$ (for some positive constants C_1, C_2; $r_{\min}$ is the minimal interparticle separation)

   (c) if $r_{\min}$ is bounded below, then $\|V\|_\infty$ and $\|\vec{\nabla} V\|_\infty$ and $\|\nabla^2 V\|_5$ ♣ become bounded,

   (d) The locus of locations for one particle, (when the others are held fixed) in which $|V| > C$, when $C \to \infty$, tend to infinitesimal ball-like shapes (of effectively ♣ bounded isoperimetric ratio),

   Note, all these norm bounds have to be known (i.e. provided as part of the input) or efficiently computable; the same is true of balls approximating the ball-like regions[9].

7. Claims of "polynomial slowdown" should mean polynomial in the input description length (number of bits and of qubits), the number of decimals P of precision, the energy E (more precisely[10] an upper bound on E), the sidelength H of the enclosing box, the number N of particles, the maximum particle mass m, the duration t of time the physical system is to be in operation, and the norm bounds and all the other parameters in rules 5 and 6.

   The runtime should also be polynomial in R, the run number (R is an integer: 1 on the first simulation run, 2 on the second, and so on) in a sequence of rerun simulations designed to give L_2 error proportional to $R^{-1.1}$ on the Rth run. (This 1.1-power error decrease law suffices to make the simulation statistically indistinguishable from the real thing; cf. rule 3 and §2.3.)

Sections 7 and 8 show how to sample from approximate *Feynman path integrals* with a quantum computer. Specifically:

[4] If the simulator is a "quantum computer" with "45° rotation" and "measurement" qubit primitives, then it automatically has access to random classical bits.

[5] Do not confuse N (or N_Q) with n. The latter is employed in §5.3, §8, §9, §11 to denote the number of timestepping stages in an algorithm.

[6] I had once thought the stronger assumption that some exponentially growing function C^E of the energy E, had finite expectation value, was going to be required. For discussion of this highly interesting assumption, see appendix §E.

[7] Alternative assumptions are possible in which, e.g. the derivatives of V need not be computable, provided V is nice in some other ways. It is also possible to handle "finite square well" potentials and "rigid inpenetrable balls." See §6.1 (and §7 for hard balls) for a discussion.

[8] "Polynomial time computable real→real functions" in the sense of Ko [58] ♣.

[9] See §6.1. For the most important case in practice – Coulomb potentials – this mare's nest of restrictions is trivially satisfied.

[10] This bound, and the norm bounds, are assumed to be available as part of the simulator's input.

In time polynomial in n, g, P, and $\log H$, a "quantum computer" can sample from a probability distribution corresponding to a function F of an approximation to a Feynman path integral with n-gonal paths whose vertices lie on a hypercubic grid, with 2^g gridlines in each direction, in a d-dimensional hypercube of side H, and in which the approximate path integral is computed accurate to P decimal places. A wide class of path integrals may be handled (we need not restrict ourselves to ones which arise quantum mechanically), provided the integrand is a phase factor e^{if} with angle f computable to P decimals in polynomial time, and $\|\vec{\nabla} f_{\rm disc}\|_1$ (where $f_{\rm disc}$ is a version of the function*al* f "discretized" to become a function) is bounded. A wide class of functions F may be sampled (we need not restrict ourselves to familiar quantum observables, nor is it necessary for F to depend only on the endpoints of the path; it could depend on the entire path). Certainly the transformation to any eigenbasis corresponding to any operator whose classical version is some computable function of particle positions and momenta (as in rule 2 above), is useable to play the role of F. We argue (§9.5) that the error in the approximation to the true path integral caused by the finiteness of g, and P should be exponentially negligible in practice. Under reasonable assumptions about the integrand (which are forced if our potentials are "regularized") we can make this rigorous.

This path integration quantum algorithm is of independent interest. Previously, two (to my knowledge and/or in my opinion) interesting kinds of algorithms had been invented strongly suggesting that a quantum computer is capable of performing some algorithmic tasks asymptotically far more rapidly than any "classical computer:"

1. Shor [81] showed that a quantum computer could (with high probability of success) find the prime factorization of an N-digit integer in a number of steps polynomial in N.

2. Grover [38] [15] showed how to square-root the expected number of steps taken by naive searching algorithms, by use of a quantum computer.

These results both appeared more relevant after the development of "quantum fault tolerance" techniques [74].

Shor's result was of tremendous theoretical importance since it strongly suggested that "Church's (strong) thesis" (that any reasonable physical system should be simulable by a computer with only polynomial slowdown) is false and therefore that quantum mechanics is intractably hard to do – at least for those of us with classical systems as brains. Looked at another way, it suggested the possible tremendous importance of trying to build a quantum computer.

So ours is the third such algorithm, and it evaluates numerical approximations to "Feynman path integrals." The algorithmic idea is perhaps even simpler than [81] [15]. Since a large amount of human and computer time [52] [53] [61] [26] is currently spent trying to solve Schrödinger's equation or approximations to it and/or to evaluate path integrals, whereas comparatively little time (I would hope) is spent trying to factor integers, this algorithm seems of more inherent usefulness (if a quantum computer ever is constructed) than [81].

For our purposes, of course, the point of the path integration algorithm is that Feynman [28] (and others [80][71], see §7) had reformulated Schrödinger's equation (and related equations) as path integration problems. This allows us to simulate any particular quantum physical system with a quantum computer with "polynomial slowdown."

However, we have to worry about how accurate this simulation is.

In sections 9 and 5, we consider how much error is caused by the finiteness of n, i.e. the discreteness of simulation time. The analysis considers "exponential splitting" formulas related to the famous "Trotter product formula." If the matrices in these formulas are finite dimensional, then such analysis is straightforward. We show that the analysis is also tractable for *operators,* assuming the potentials have been "regularized" and provided we assume the physical "reasonableness" of the wavefunctions being simulated, in particular boundedness of energy.

Provided such definitions of "reasonableness" are accepted, this constitutes a *Quantum version of Church's strong thesis* – any quantum physical system may be simulated by a quantum computer, to accuracy as good as could be obtained by measuring the system itself, with only polynomial slowdown.

Finally, in section 10 we show how a conventional computer can simulate a quantum computer with exponential slowdown but only polynomial amplification of memory space needs (this is almost trivial). This completes the proof of our preferred version of Church's thesis.

Section 11 ties everything together, providing an approximate statement of the simulation algorithm and the main result, and telling which lemmas are used in what ways to get the proof.

Section 12 concludes, and discusses recent improvements to our main results and possible future improvements.

2 Quantum Church's thesis definitions

Thesis 1 Church's weak thesis *is the notion that if any "computer" C (read: "physical system") performs a "computation," then that same computation could have been performed by a Turing machine.*

Thesis 2 Church's strong thesis *adds to the weak thesis the qualification that the time consumed by the Turing machine computation is bounded by a polynomial ♣ of the time and "resources" consumed by C, the description length of C, and the length of its input.*

Thesis 3 Our new "intermediate Church thesis" *adds to the weak thesis the qualification that the memory* *space consumed by the Turing machine computation is bounded by such a polynomial.*

Church's theses were intentionally vaguely phrased. It becomes possible to try to prove or disprove them if we specialize to some particular set of laws of physics and if precise meanings are assigned to all the words in quotation marks. For example, see [83] for proofs and disproofs of Church's thesis in classical mechanics.

If we want to allow the laws of physics to include quantum mechanics, then additional complications arise in the problem definition. We'll now discuss these, one per subsection.

2.1 Nondeterminism

Quantum mechanics is not deterministic. Therefore, certainly any simulator necessarily would need to be equipped with a random bit generator.

2.2 Meaning of Computation

What "computation" is performed by a quantum system? One's first reaction might be: a quantum system computes the wavefunction Ψ which solves Schrödinger's (multiparticle) equation. (Or if we are relativistic, a different equation; this distinction is not significant for our present purposes.) But in fact, that is not the case because it is impossible to measure the wavefunction. The wavefunction constitutes really an infinite number of bits of information, almost all of which are inaccessible during a single measurement (which destroys the wavefunction).

Actually, all we can hope to do is the following: We can choose a set of commuting Hermitian operators $\hat{F}_1$, $\hat{F}_2$, ... corresponding to "compatible physical observables." If we knew the full expansion

$$\Psi = \sum_j c_j \Phi_j \tag{1}$$

of Ψ as a complex linear combination of the common complete orthonormalized eigenfunctions Φ_j of these operators, then we would know Ψ. But upon performing the measurements of those observables, all we get is a *probabilistic sample* from the probability distribution in which j occurs with probability $|c_j|^2$. So, to be fair[11], the Turing machine (or whatever is performing the simulation) should really only be required to produce one such probabilistic sample, and should not be required to output the entire wavefunction (which might require an enormous amount of paper to write down).

[11] "Fairness": The physical system and the simulator are regarded as competitors, and we want neither side to have an "unfair advantage." If the physical system is only required to output one random variable (the result of some measurement) but the simulator were required to produce far more (e.g. the entire wavefunction) that would "unfair." See §2.9.

2.3 Accuracy

Of course by running the Turing machine many times to perform many different simulations of the quantum system, we could extract more information about the wavefunction because we would get many such probabilistic samples – corresponding to the fact that we would need to run the real quantum experiment many times to get this same information.

I'm going to imagine that one in fact does plan to rerun the simulator many times. But the simulator necessarily will not (on any given run) compute the exactly correct wavefunction, and hence will sample from some incorrect probability distribution. If all simulation runs ran the same algorithm with the same input each time, we would eventually be able, by collecting enough samples from the simulator and from the real quantum system, to become *confident* of the fact that they were two different distributions, i.e. that the simulation was wrong.

To avoid this sad fate, we shall *demand* that the simulator actually *tries harder* on the run $R+1$ than on run R, in such a way that a statistician examining the outputs in an infinite sequence of repeated runs, *cannot* build up unboundedly high confidence that the simulation and real experiment were sampling from different distributions.

Now let us analyse this idea more precisely.

The amount of information we get about the $|c_j|^2$, i.e. about the wavefunction, from a single probabilistic sample is roughly $\sum_j |c_j|^2 |\log_2(|c_j|^{-2})|$ bits. (This is just to clarify the point that a measurement will only yield a finite amount of information about the infinite amount of information specifying the wavefunction.) To estimate the true probability p of some event, one must perform n experiments (or simulations) to obtain constant confidence that we have p accurate to $\pm n^{-1/2}\sqrt{(1-p)p}$.

Suppose the true probability (in the real quantum system) of some event is p, whereas the probability of that event in our simulation, is p'. Both p and p' are unknown. How many experiments n and simulation runs n' are required to get some threshhold (say 75%) of confidence that $p \neq p'$? We claim it is impossible to reach a confidence threshhold that $p \neq p'$ if

$$|p-p'|\,(n+n') \ll \sqrt{(1-p)pn + (1-p')p'n'}. \tag{2}$$

This leads to the conclusion that if $n = n'$ and $|p-p'| = o(n^{-1/2}\sqrt{(1-p)p + (1-p')p'})$ (where now p' depends on n), then we cannot gain arbitrarily high confidence that $p \neq p'$, no matter how many experiments n we perform.

Now let us consider two normalized wavevectors Ψ_1 and Ψ_2 (say in some countably-infinite dimensional space) which differ by a small angle θ. It is plain that the best[12] way to distinguish between them is to rotate them both to a configuration in which they are both zero in all coordinates except for two (we assume optimistically – which is OK since we are seeking an upper bound

[12] In the sense of requiring the fewest measurement operations to reach some confidence threshhold.

on the best we can possibly do – that we somehow know such a rotation and can accomplish it[13]), and in that 2-space they differ by the angle θ. We've now reduced the problem to a single "qubit" ♣. Now we can perform a measurement on this qubit, and we can redo this experiment as many times n as we please, assuming we have some way of reproducibly regenerating Ψ_1 and Ψ_2. But we've just analysed the single-bit measurement problem in the previous paragraph. That analysis re-applies here, once you see that $|p - p'| \le \theta$.

But now, if we consider a sequence of simulations which, in run number R, have L_2 error $O(R^{-1.1})$, we conclude that

Claim 4 *If the simulated wavefunction Ψ from the Rth simulation run, in the usual L_2 distance metric[14], is $O(R^{-1.1})$ ♣ away from the true wavefunction, then, no matter what measurements one performs, one will be unable to become arbitrarily confident that the simulation is in error.*

So to be fair, quantum formulations of Church's thesis should not require more simulation accuracy than this.

2.4 Choice of error metrics

Note that using the L_2 metric was essential in the above reasoning. The L_2 metric is the natural one in quantum mechanics because

1. If the L_2 distance (i.e. angle θ) between two wavefunctions is small, then a large number (i.e. $\gtrsim \theta^{-2}$) of measurements are required to gain confidence they are different,
2. It is preserved by basis changes.

The second property is very false for, e.g. the L_1 metric. A small perturbation by a d-vector $(\sqrt{d}, 0, 0, \ldots, 0)\epsilon$ could be transformed by a unitary time-evolution or basis change to $(1, 1, \ldots, 1)\epsilon$, i.e. amplifying this small perturbation (assuming ϵ is small), by an unboundedly large factor ($\sqrt{d}$) in L_1. Such amplification is impossible in L_2. Consequently the first property is also false in other metrics – in the sense that it is impossible to bound n purely in terms of θ without knowledge of d.

2.5 Precision

A different accuracy conundrum (which also arises when considering simulating classical systems) arises because of the fact that times, positions and momenta are real numbers, i.e. an infinite number of bits. One may handle this either

1. ⊗ By assuming a computational model such as a "real RAM" in which real number I/O and arithmetic operations may be performed in one time step;
2. Or by assuming an input oracle which can provide more bits of accuracy upon request, and only requiring the output to be accurate to P decimal places where P is an input parameter, and where any polynomial time or space simulation slowdown guarantees are also allowed to *depend* polynomially upon P and the number of bits transmitted to and received from oracles.

We prefer choice #2[15]. However, because

1. Time evolution in quantum mechanics is *unitary*, causing "chaos" (exponential sensitivity of the final state to infinitesimal changes in initial conditions) to be impossible (very unlike classical mechanics) and
2. because we do not need to compute the final state to high accuracy anyway (as we just discussed[16]; also unlike classical mechanics)

high precision, fortunately, is not an important issue.

Although actually it is not necessary to specify P (since it is deducible from other parameters), as a matter of programming and mental convenience it is desirable to demand that the simulator accept P (which will be the number of decimals to which all arithmetic is supposed to be accurate to, essentially) as another input parameter. Too large P will cause wasteful computation to too many digits of significance (it is pointless to compute an approximate answer to 100 decimal places when we know our approximation is only accurate to 5 places) but this will not matter. We agree to disregard requests for too small P (which could cause misleading answers).

2.6 Input representation

The input has to describe the laws of dynamics, i.e., the Hamiltonian operator, for the simulated quantum system. We shall assume that the *classical* version of this operator corresponds to a known polynomial-time computable function of the particle canonical position and momentum coordinates – polynomial in the number of decimals in all these coordinates and in the number of decimal places to which we want the answer, that is.

The problem of representing the initial quantum state is more serious. If we start a quantum system off from some state, just describing that state may require an enormous number of bits of information, which anyhow we don't know. It seems hardly fair to require the simulator to input an infinite number of bits we don't know before it can even start. There are several ways to define ourselves out of this quandary.

[13] An ingenious "quantum circuit" with $O(N)$ "quantum gates and wires," invented by John Watrous, in fact enables essentially the same quantum state distinguishing power to be gotten (for states made of N "qubits"), as the method I am proposing, also via a single qubit measurement, but *without* need to know, or to use, this optimum rotation. This circuit is described in [17].

[14] I.e. $2\sin\frac{\theta}{2} = \mathrm{dist}_{A,B} = (\int |\Psi_A - \Psi_B|^2 \mathrm{d}x)^{1/2}$. Note: this metric is the same for *any* complete orthonormal basis defined by the eigenstates of *any* quantum observable.

[15] Incidentally, we also assume that the values of physical constants such as $\hbar$ and ϵ_0 ♣ are available from an oracle to arbitrarily high accuracy – in practice a polynomial number of decimal places will suffice.

[16] That is, accuracy $O(R^{-1.1})$ suffices on run $\#R$, which is a far less stringent demand than exponential accuracy.

1. ⊗ We can only allow initial quantum states which are eigenfunctions of some observable – these are concisely describable and also correspond to the maximally precisely specified quantum states which (according to the Copenhagen interpretation) can be prepared. But then, there is the sticky problem of just which observables to allow[17]; more on this below.

2. ⊗ We could assume the input was available in the form of some sort of "oracle." The simulator could ask the oracle questions about the wavefunction and get answers. But then we are confronted with the sticky problem of just what questions to allow. If we were only to allow questions of the form "give me the nth decimal digit of the complex amplitude c_j of Φ_j in the initial state (of form (EQ 1))" then it could take unboundedly long to find the j which has the biggest $|c_j|^2$, and why should we restrict ourselves to one particular allowed set of the Φ_j anyway – maybe the simulator might prefer to use another. But on the other hand, if we allow more powerful questions such as arbitrary Φ_j and asking for the kth biggest c_j, then we run the risk of offloading the bulk of the computation to the oracle. It is difficult to quantify that risk.

3. If our simulator is actually a "*quantum* Turing machine" capable of containing quantum information, we could simply assume that an appropriate internal representation (§2.7) of the initial state information is available in its quantum memory at the start of the computation.

In the present paper, we are going to consider simulating arbitrary quantum systems by quantum computers, and we are therefore going to choose the final one of the above three options. When we consider simulations with a classical computer, we take (a restricted version of) the first option. The strategy we will employ – classical simulation of a quantum computer simulating a quantum system – then provides a natural answer to the question of which operators to allow, or more precisely, which quantum initial states to allow. Namely – the ones we can get our quantum computer to describe in quantum polynomial time after classical initialization with pure-0 or pure-1 bits.

In some sense this restriction is circularly self-justifying: it leads to final states which are in the same class as the permitted initial states. (Hence, one can then feed these into more simulations.) The same kind of circular self-justification applies to the assumption (§2.9) that all states have bounded expectation value for energy (since this value is conserved by Schrödinger time-evolution). This is not a priori necessary since, e.g. a probability density proportional to $1/(x^{1.1} + 1)$ for $x \geq 0$ has no expectation value of x; but it nevertheless seems very plausible that it is true for physically reasonable wavefunctions. The circularity behind these assumptions, or restrictions, does not bother me personally. But it could be argued that really the only possible justification for these properties is the ultimate assumption – that the initial state of the universe obeyed these properties![18]

There is a different and also valid interpretation of all this, which many people will prefer. It is that *any* initial state of bounded expected energy is permissible, whether or not it is easily prepared by a quantum computer. In other words, assuming some all-powerful being were somehow to prepare our quantum computer with some of its qubit registers representing (in our format – see §2.7) *any* initial quantum state, I will claim the quantum computer could then proceed to simulate its time evolution with polynomially bounded slowdown.

Furthermore, a conventional computer could also perform the simulation with exponentially bounded slowdown. Note that in this case, the description of the initial quantum state which it inputs at the beginning of the simulation, would be exponentially many classical bits long.

2.7 Internal representation of wavefunctions

Related to the previous issue of how to represent the initial state, or initial wavefunction, is the larger issue of how to represent *any* wavefunction inside our computer.

We are going to use Fourier modes (more precisely, the solutions of the free particle Schrödinger equation) in a box. We assume, or proclaim, that the system being simulated lives inside a cubical box of sidelength H. The wavefunction is a linear combination of Fourier modes. We cut off the mode sum at some frequency upper limit. We can switch from this mode-sum representation (which is, essentially but not precisely, the momentum basis) to a position basis by using a discrete Fourier transform. In that case, the mode cutoff turns into the fact that we are discretizing continuous space onto a finite grid.

Now to represent a *quantum* state, i.e. a complex-amplitude weighted superposition of Fourier modes, we can use a single *quantum register* [81] ♣ inside our quantum computer (which is performing the simulation). Specifically, representing a superposition of 2^g Fourier modes requires a single g-qubit register. The complex amplitude weights are not represented explicitly anywhere – they just are the 2^g weights for each of the 2^g possible logical states of the quantum register. If the "quantum Fourier transform" [81] of this register is performed, we get, essentially, that the resulting g-qubit register has value 1 if the particle it represents is at grid point 1, value 2 if the particle it represents is at grid point 2, and so on up to gridpoint 2^g. We thus can represent the entire wavefunction of N particles in a 3D box on a $2^g \times 2^g \times 2^g$ grid inside $3N$ quantum registers, each

[17] Nor is there any a priori guarantee that the eigenfunctions of some specified operator even are computable functions [73]; but see appendix §D.

[18] Well – did it? Can any cosmologist tell us?

g qubits wide.

The only issue that remains is the question of how much accuracy we have lost by ignoring the high frequency modes (of wavelength $< 2^{-g}H$). Well, if c_j is the complex amplitude of mode j, then the *kinetic energy* E_K of our wavefunction is

$$E_K = \hbar \sum_j |c_j|^2 \omega_j \tag{3}$$

where ω_j is the frequency of mode j. If we ignore all the modes with $\omega > \omega_{\text{cutoff}}$, then we see that we have altered the wavefunction, in the L_2 distance metric, by

$$\left(\sum_{j,\ \omega_j > \omega_{\text{cutoff}}} |c_j|^2 \right)^{1/2} \tag{4}$$

which is $\leq \sqrt{E_K/(\hbar\omega_{\text{cutoff}})}$. We shall see in §6 that E_K is bounded. Hence our L_2 wavefunction error is arbitrarily exponentially small if we make ω_{cutoff}, i.e. the number of gridlines 2^g, exponentially large.

Hence, the error introduced by the mode cutoff is negligible compared with the errors we are allowing ourselves (§2.3).

2.8 Interparticle potentials

We argue in appendix §B that intercharge potentials in quantum mechanics are best modeled as *noncoulombic* and without attractive singularities. This is important since less-singular potentials lead to faster simulations with better error bounds.

Indeed, our simulation algorithms fail entirely in the presence of attractive interparticle potentials that behave like $-r^{-\gamma}$ with $\gamma > 2$ as $r \to 0$. But it is known that $\gamma > 2$ is impossible in quantum theory since such potentials lead to Hamiltonians without self-adjoint extensions [47][69] – physically, the particles get "sucked in" to each others's singularity.

But fortunately, we do not need to rely on the anti-Coulomb arguments in appendix §B (which is the reason they are in an appendix). That is because our simulation algorithm can handle any $\gamma < 6/5$; see rule 6 of §1.2 for a more precise description of the allowed potentials[19]. (Also, any $\gamma < 2$ seems handleable using the alternative approach sketched in appendix §D[20] – but not in quantum-computer polynomial time.)

Rule 6 places numerous other restrictions on the potential (not all of which, perhaps, are really necessary and hence might be subject to elimination by later authors). I did that for technical reasons arising in the proof. For Coulomb ($1/r$) potentials these restrictions are all easily seen to be satisfied. We can also easily handle some other commonly used potential functions such as H.Yukawa's "screened Coulomb" potential $V_{\text{Yukawa}} = qe^{-kr}/(4\pi\epsilon_0 r)$ and E.A.Uehling's potential

$$V_{\text{Uehling}} = \frac{q}{4\pi\epsilon_0 r}\left[1 + \frac{2\alpha}{3\pi}\int_1^\infty (1+\frac{1}{2u^2})\frac{\sqrt{u^2-1}}{u^2}e^{-2ru/\lambda_C}\,\mathrm{d}u\right] \tag{5}$$

(both expressed as voltages arising from a source charge q). It is also possible (although it goes outside the main stream of our methods and requires some extra tricks) to handle "finite square well" potentials and even "hard balls" (in which the potential is infinite if two particles are separated by < 1 "hard ball diameter") [42].

But there are some other commonly used potentials which we *cannot* handle. These include the Lennard-Jones semiempirical potential (often used to model interatomic forces)

$$V_{\text{LennardJones}} = Ar^{-12} - Br^{-6} \tag{6}$$

where A and B are positive constants, and the potential between a point charge and a point dipole ($\propto -r^{-3}$). The repulsive singularity in (EQ 6) is too severe ($|\vec{\nabla} V| \notin L^1$) for the methods of §3 and §9.5 to handle. The dipole-charge $-r^{-3}$ attractive singularity is too severe for quantum mechanics itself to handle (see footnote 35) which is corroborated by the fact that no point particles with electric dipole moments are known. This same remark (assuming that electrons are point particles with a magnetic dipole moment) makes it impossible for point *magnetic monopoles* to exist in quantum mechanics (which I point out here, since nobody has pointed it out before).

Indeed, quantum mechanics has, so far, only been shown to be well posed for, potentials in, e.g., the Rollnick ♣[21] class. The Lennard-Jones potential is well outside of such classes. Hence, so far as is currently known, it is quite possible that the Lennard-Jones potential is actually inadmissible in quantum mechanics[22]. The same is true for its modification (which includes an angular dependency) the "Stockmayer potential." However, as Lennard & Jones surely would agree, the true asymptotic form as $r \to 0$ of the singularity for the potential between two argon atoms whose nuclei are separated by r, certainly is not $\propto r^{-12}$, but instead behaves coulombically (or even less singularly than that, once the nuclei start to overlap). Most users of the Lennard-Jones potential would argue that the wrong behavior at the singularity "does not matter" because in the energy and r ranges of interest, r^{-12} does describe the truth fairly well. But that argument is totally invalid if the allegedly irrelevant part near $r \approx 0$ actually destroys quantum mechanics or your simulation algorithm.

A dozen empirical potentials of this sort are surveyed in chapter 1.3 of [42].

My personal intuition agrees with the physicists I just disparaged (see §6.1.4) – I suspect even very severe potential singularities ought to be permitted, provided they

[19] If you're wondering how the magic value 6/5 arose, see lemma 13 and (EQ 42-45). Coulomb's law, of course, involves $\gamma = 1$.

[20] And see §12.5.

[21] See theorem X.19 of vol. 2 of [75].

[22] A fact which the authors of ≈ 100 papers on the Lennard-Jones potential, appear to be blissfully unaware of.

are repulsive – but currently no analysis is strong enough to prove such a conclusion.

2.9 Resources

The "resources" in a physical system being simulated are: its[23] energy E, the maximum mass m of any particle, the number N of particles, the[24] maximum charge on any particle, the duration t of the time-evolution being simulated, and the size of the box it is contained in.

If we view Church's *strong* thesis as saying you can't build a physical system that can superpolynomially outperform a classical computer, then it would clearly be "unfair" to the simulator if the physical system it were "competing" against were allowed exponentially enormous masses or exponentially many particles[25]. It also would be unfair if the physical system were unbuildable or inaccessible because it was too huge to fit in your laboratory, i.e. didn't fit in a box of polynomially bounded edge lengths[26].

3 Approximation of integrals by sums

Here is a simple bound on the error ϵ that arises when approximating an integral by the "trapezoidal rule."

Lemma 5

$$\frac{\frac{f(0)}{2} + f(\frac{1}{n}) + f(\frac{2}{n}) + \ldots + f(1-\frac{1}{n}) + \frac{f(1)}{2}}{n} = \int_0^1 f(x)\mathrm{d}x + \epsilon \qquad (7)$$

where

$$|\epsilon| \le \frac{1}{2n}\int_0^1 |f'(x)|\mathrm{d}x = \frac{1}{2n}\|f'\|_1 \le \frac{1}{2n}\|f'\|_2 \qquad (8)$$

This is easy to prove directly (the final inequality in (EQ 8) works only for the interval $[0,1]$ of *unit* length), but also arises as the simplest special case of the Euler-Maclaurin summation formula with error term (discussed in section 9.5 of [36], and other textbooks).

Corollary 6 *If f is real valued,*

$$\left(\frac{e^{if(0)}}{2} + e^{if(1/n)} + e^{if(2/n)} + \ldots + e^{if(1-1/n)} + \frac{e^{if(1)}}{2}\right)/n$$

$$= \int_0^1 e^{if(x)}\mathrm{d}x + \epsilon \qquad (9)$$

where ϵ obeys (EQ 8).

For the corresponding *d-dimensional* integrals over $[0,1]^d$ approximated by sums on cubical grids of $(n+1)^d$ points, and where $f(\vec{x})$ of course now takes a d-vector argument, the error ϵ obeys

$$|\epsilon| \le \frac{1}{2n}\,\|\sum_{j=1}^{d}|\frac{\partial f}{\partial x_j}|\,\|_1 \qquad (10)$$

as may be proved by induction on d.

4 Splitting formulae for exponentials of finite dimensional matrices

As will become apparent in this and the next section, Feynman's definition of his "path integral" (§7) arose from the "Trotter product formula" for operators.

We'll now discuss this and related higher order product formulae. But for now (for simplicity) we'll only discuss these formulae in the context of finite dimensional matrices. That way, this section may be read independently.

It turns out that we will only actually *use* the 2-term and 3-term formulae in the Church's thesis proof, so we'll only discuss them here. But the higher order formulae may be relevant to future improvements of complexity bounds, as well as being of independent interest. They presumably may be used to get better approximate path integrals. Hence we include a survey of them (including new results) in appendix §C.

4.1 Product formulas

Let A and B be two noncommuting general real (or complex) square finite dimensional matrices. Let t be a small real (or complex) number; we are interested in $t \to 0^+$.

4.1.1 2-term product formula

$$e^{(A+B)t} \;=\; e^{At}e^{Bt}[1+O(t^2)]. \qquad (11)$$

This result is immediate from the Campbell-Baker-Haussdorf-Dynkin formula (found by Campbell in 1897 and refined by Baker, Haussdorf, and Dynkin in succession over the next 50 years [98])

$$e^X e^Y = e^{X+Y+[X,Y]/2+([[X,Y],Y-X])/12+[[[X,Y],X],Y]/24+\text{order5}+\ldots} \qquad (12)$$

(We use "commutator" notation: $[X,Y] \equiv XY - YX$.) Expressions yielding the nth order term (which is a homogenous antisymmetric polynomial of degree n in X and Y) in the series are in [98] and [88]. Indeed [88] proves that the series converges if $\|X\|+\|Y\| < (\ln 2)/2$. An even better proof uses the (less well known) Zassenhaus formula

$$e^{X+Y} \;=\; e^X\,e^Y\,e^{[Y,X]/2}\,e^{[[Y,X],X+2Y]/6}\cdots. \qquad (13)$$

[23] More precisely, the *expectation* value of the energy, or some upper bound on it. This quantity is assumed to be provided as part of the input.

[24] Charges in real life appear to be integer multiples of the electron charge e (or $e/3$ if one allows "quarks") and in appendix §B, especially footnote 65, we explain why that integer is always *bounded* by an absolute constant. But our simulation algorithm allows a superset of real-life scenarios, including the possibility of non-integer charges.

[25] Allowing an experimenter to play with 10^{100} grams of mass, or to run an experiment for 10^{100} years, for a cost of only $100, seems unfair!

[26] Incidentally, my "disproof" of Church's thesis in classical Newtonian mechanics [83] had relied upon having an infinitely large, or at least nonrecursively large, containing box. That is one basis for attacking the physical relevance of that result.

Suzuki [88] also gives the general term of this product (the nth term is the exp of a homogenous polynomial of degree $n-1$ in X and Y) and proves convergence if $\|X\|, \|Y\| < \ln(2/\sqrt{e})$.

4.1.2 3-term product formula

$$e^{(A+B)t} \;=\; e^{At/2}e^{Bt}e^{At/2}\,[1+O(t^3)]. \tag{14}$$

This also arises from the C-B-H-D formula (EQ 12) via a rewriting as

$$e^X e^Y e^X \;=\; e^{2X+Y+([Y,[Y,X]]-[X,[X,Y]])/6+\text{order}5+\dots}. \tag{15}$$

5 Error bounds for splitting formulae for operator exponentials

If A and B have commutators which are small in some sense (for example, of bounded norm) then *error bounds* for the 2-term and 3-term product formulae (EQ 11, 14) may be obtained from Suzuki's [89] "identity 1"

$$e^{tA}e^{tB} - e^{(A+B)t} = \tag{16}$$
$$\int_0^t \int_0^v e^{vA}e^{(v-u)B}[A,B]e^{uB}e^{(t-v)(A+B)}\,\mathrm{d}s\mathrm{d}t$$

and "identity 6"

$$e^{(A+B)t} - e^{tA/2}e^{tB}e^{tA/2} = \tag{17}$$
$$\frac{1}{2}\int_0^t \int_0^v e^{vA/2}e^{vB}G(u)e^{vA/2}e^{(t-v)(A+B)}\,\mathrm{d}u\mathrm{d}v$$

where $G(u) =$

$$\int_0^u \left(\frac{1}{2}e^{wA/2}[A,[A,B]]e^{-wA/2} + e^{-wB}[B,[A,B]]e^{wB}\right)\mathrm{d}w. \tag{18}$$

respectively. The obvious[27] bounds that result are, respectively, as follows.

Lemma 7 (Error bounds for exponential splitting) *Let Ψ be normalized (i.e. have unit L_2 norm). Let $\|\cdot\|$ denote L_2 norm for either operators or Hilbert vectors. Let t be real. Let A and B be Hermitian so that e^{itA} and e^{itB} are unitary. Then*

$$\|(e^{itA}e^{itB} - e^{(A+B)it})\Psi\| \le \frac{t^2}{2}\|[A,B]\| \tag{19}$$

and

$$\|(e^{(A+B)it} - e^{itA/2}e^{itB}e^{itA/2})\Psi\| \le \tag{20}$$
$$\frac{t^3}{12}\left(\frac{1}{2}\|[A,[A,B]]\| + \|[B,[A,B]]\|\right).$$

These bounds sometimes are useful, but sometimes are useless.

[27] They arise by using the facts that e^{iA} is unitary if A is Hermitian, and unitary operators have unit L_2 norm, and the sub-multiplicativity and sub-additivity properties of L_2 operator norms.

5.1 Example of uselessness of the error bound for exponential splitting: the hydrogen atom

In particular, consider the case where A is the potential energy operator (i.e. $A\Psi$ causes multiplication of Ψ by a potential function V) and B is the kinetic energy operator, i.e.

$$B = \frac{-\hbar^2}{2m}\nabla^2. \tag{21}$$

More precisely, in the N particle case,

$$B = \sum_{j=1}^{N} \frac{-\hbar^2}{2m_j}\nabla_j^2 \tag{22}$$

where m_j is the mass of the jth particle and ∇_j^2 is taken with respect to its 3 coordinates. This is really the same thing as (EQ 21) if one views the N particles as a single superparticle in a $3N$-dimensional space and is a bit generous in interpreting the division by m (this could be dealt with by appropriate scaling of the coordinates). Hence we'll abuse notation by continuing to use (EQ 21) here and in §5.2 to keep expressions simpler (and anyway $N=1$ for the present purposes: the hydrogen atom with anchored nucleus).

In that case

$$[V,\nabla^2]\Psi = V\nabla^2\Psi - \nabla^2(V\Psi) = -\Psi\nabla^2 V - 2(\vec{\nabla}V)\cdot(\vec{\nabla}\Psi). \tag{23}$$

Now if $r = |\vec{x}|$ denotes the distance to the origin in 3-dimensional space and $V = -k/r$ (potential for a hydrogenic atom; k is a positive constant) this is

$$[\frac{-k}{r},\nabla^2]\Psi = [\frac{k}{r},-\nabla^2]\Psi = -k\Psi\delta^3(\vec{x}) - \frac{2k}{r^3}\vec{x}\cdot(\vec{\nabla}\Psi) \tag{24}$$

where δ^3 is a 3D Dirac delta function ("point mass"). Now if Ψ is the wavefunction of the ground state of the hydrogen atom ([78] IV.16), i.e.

$$\Psi(\vec{x}) = 2a^{-3/2}e^{-r/a} \tag{25}$$

where $a = \hbar^2/(km)$, then $[V,\nabla^2]\Psi$ is not in L^2 and the upper bound (EQ 19) is ∞. While ∞ is a valid upper bound, it is useless! The $\Psi\nabla^2 V$ term, which generates the delta function, is the culprit.

5.2 The error bound for exponential splitting is useful if the potentials are regularized and the wavefunction has bounded energy

The way out of this quandary is: we only allow *regularized* potential functions V, i.e. ones such that $\|\nabla^2 V\|_5$ and $\|\vec{\nabla}V\|_\infty$ are both bounded. In that case, consider the terms in (EQ 23).

- $(\vec{\nabla}V)\cdot(\vec{\nabla}\Psi)$ will automatically have bounded L_2 norm, because $\vec{\nabla}\Psi$ has bounded L_2 norm (proportional to the kinetic energy (EQ 32,34)) and apply Hölder's inequality (EQ 84) with $p=1$, $q=\infty$.

- Similarly $\Psi\nabla^2 V$ will automatically have bounded L_2 norm because Ψ is in $L^{10/3}$, due to the boundedness of the kinetic energy and the Lieb-Thirring inequality (EQ 35), and apply Hölder's inequality (EQ 84) with $p = 5/3$, $q = 5/2$.

We conclude that the error bound (EQ 19) for the 2-term product formula (EQ 11), will always be finite and well behaved (indeed, depending polynomially on the L_2 norm bounds) if the potential function is regularized (Here A and B are the potential and kinetic energy parts of the Hamiltonian and Ψ is a normalized wavefunction of bounded energy):

Theorem 8 (Error bound for 2-term product approximation of Schrödinger time evolution) *Let $\|\cdot\|$ denote the L_2 norm. Let V be a potential function and Ψ be a (normalized) wavefunction. Then the L_2 norm of the error (i.e. $\|\Psi_1 - \Psi_2\|$) obtained by using the 2-term product formula (EQ 11) (with A and B being the potential and kinetic parts of the Hamiltonian operator $A + B$) to approximate the Schrödinger time-evolved Ψ after a time t, is*[28]

$$\leq \ (\|\Psi\|_{10/3} \, \| \frac{\nabla^2 V}{m} \|_5 \ + \ 2 \, \|\vec{\nabla}\Psi\|_2 \, \| \frac{\vec{\nabla} V}{m} \|_\infty) \frac{\hbar^2}{2} \frac{t^2}{2}. \quad (26)$$

It is possible to proceed similarly for the 3-term product formula (EQ 11), but using (EQ 20) instead of (EQ 19). But in this case we need also to make assumptions about the L_∞-boundedness of higher derivatives of the potential function V. Specifically, we need to assume that $|\vec{\nabla} V|^2$ is in L^5, $(\nabla^2)^2 V$ is in L^∞, and

$$\sum_k \frac{\partial^3 V}{\partial x_k^3} \frac{\partial \Psi}{\partial x_k} \quad (27)$$

is in L^∞ ♣, as well as (as we did before) $\nabla^2 V \in L^\infty$. Unlike before, we now also need to assume that $(\nabla^2 V)(\nabla^2 \Psi)$ is in L^2, or at least that $\int (\nabla^2 V)(\nabla^2 \Psi)\Psi^* \mathrm{d}^3\vec{x}$ is bounded. If $\nabla^2 V$ is in L^∞, then this assumption is again justified if Ψ has bounded kinetic energy (which it will, see §6). The other assumptions may be justified by forcibly regularizing V. The error term then behaves proportionally to t^3 times a polynomial function of all the norm bounds.

Presumably the 4-term, 5-term, etc. product formulae (see the appendix §C) also cause bounded L_2 errors, with enough assumptions about Ψ and V and their (sufficiently high) derivatives, but this has not been proven.

5.3 The convergence of the "Trotter product formula" and its ilk

The Schrödinger time evolution operator is $\exp(\frac{-it}{\hbar}\hat{H})$ where $\hat{H}$ is the Hermitian Hamiltonian operator.

We are going to use the Trotter product formula [71][50] for the operator exponential of i times the sum of two Hermitian[29] operators. In our case, the Hamiltonian is the sum of the potential and kinetic energy operators $\hat{H}_P$ and $\hat{H}_K$. This formula states that

Theorem 9 (Trotter product formula)

$$\exp(\frac{-it}{\hbar}\hat{H}) = \operatorname*{slim}_{n\to\infty} (R_n)^n, \quad (28)$$
$$R_n = \exp(\frac{-it}{\hbar n}\hat{H}_P) \exp(\frac{-it}{\hbar n}\hat{H}_K).$$

Here "slim" denotes the "strong limit" of an operator in the L_2 metric, i.e. a sequence A_1, A_2, ... is said to tend to a strong limit if $\forall \epsilon > 0 \ \exists N$ such that if $n, m > N$ then

$$\|(A_n - A_m)\Psi\| < \epsilon \quad \text{for all} \quad \Psi \in L^2(\mathbf{R}^d). \quad (29)$$

In short, the Trotter product formula, specialized to the case of quantum mechanical time-evolution, views time-evolution caused by the total Hamiltonian (which is a sum of a potential and kinetic energy part) as being due to a rapid alternation in time in which the Hamiltonian "switches" between being pure potential and pure kinetic[30]. The theorem is that as the rapidity of the alternation tends to infinity, the right answer is obtained.

This formula may be thought of as the justification of (and/or the interpretation of) Feynman's path-integral formulation of quantum mechanics (see §7).

This formula is useful algorithmically because there are easy exact solutions for time evolution for the pure-kinetic (free particle) and pure-potential (complex phase rotation at position-dependent rates) Hamiltonians. But in order to be useful for our purposes, it is not sufficient that Trotter converges. We need rigorous and computable error bounds saying how quickly it converges.

For that, we use theorem 8. The idea is that Trotter product formula is really just an n-time repeated use of the 2-term product formula (EQ 11). Hence we can just take the nth power of the error bound (converted to a relative error bound by adding 1) in theorem 8; and as we've seen, those error bounds are useful if the potentials are "nice enough," e.g. "regularized."

The result is

Lemma 10 (Trotter product relative error bound)
The error factor in the final wave function (after time t has elapsed) arising from chopping the t-long time interval into n equal subintervals and using the Trotter product formula, is

$$\leq \left[1 + (\frac{t}{n})^2 \widetilde{B}\right]^n \quad (30)$$

where $\widetilde{B}$ is a factor arising from norm bounds as in theorem 8, i.e. the cofactor of t^2 in (EQ 26).

[28] See the remark on abuse of notation near the beginning of §5.1.

[29] Also valid more generally for operators whose spectrum lies in the closed lower complex z-halfplane $\mathrm{Im} z \leq 0$.

[30] Actually, the total time spent in each half is t, not $t/2$, so this mental alternation picture is not precisely correct; there is also a factor-2 time dilation.

Note, if t, and an upper bound on the kinetic energy, and the cutoff height for the potential regularization are fixed, then as $n \to \infty$ our error goes to zero (i.e. the error *factor* goes to 1), indeed, is asymptotic to $t^2 B/n$.

Note that if B is only polynomially large, then it suffices to take n (proportional to the amount of computational work) only polynomially large to assure an error proportional to $R^{-1.1}$ (or smaller) on the Rth simulation run. This is, in fact, exactly what we are going to do.

5.3.1 An Alternative

Instead of the original Trotter formula based on the 2-term product formula (EQ 11), we could use something else based on repeated use of the 3-term product formula (EQ 14). This could yield better error bounds, but extra assumptions about the niceness of the potential would be needed.

6 The Lieb-Thirring form of the uncertainty principle and its uses

For a single particle of mass m and wavefunction Ψ in 3-dimensional space, Lieb and Thirring's formulation [65][63] of the uncertainty principle[31] states that

$$\left(\int \rho^3 \mathrm{d}^3\vec{x}\right)^{1/3} \leq \left(\frac{4}{\pi^2}\right)^{2/3} \frac{2m}{3\hbar^2} E_K. \tag{31}$$

where $\rho = |\Psi|^2$ is the probability density corresponding to the wavefunction $\Psi(\vec{x})$ and E_K is the[32] kinetic energy

$$E_K = \frac{\hbar^2}{2m} \int |\vec{\nabla}\Psi(\vec{x})|^2 \, \mathrm{d}^3\vec{x}. \tag{32}$$

There are also Lieb-Thirring inequalities valid if we have $N \geq 1$ particles. For example, let ρ be the probability density function for finding a particle at $\vec{x}$:

$$\rho(\vec{x}) = \sum_{k=1}^{N} \int \left| \Psi \underbrace{(\vec{x}_1, \ldots, \vec{x}_N)}_{\vec{x}_k \text{ replaced by } \vec{x}} \right|^2 \underbrace{\mathrm{d}^3\vec{x}_1 \ldots \mathrm{d}^3\vec{x}_N}_{\text{every } \vec{x}_j \text{ except } j=k} \tag{33}$$

and $\int \rho(\vec{x})\mathrm{d}^3\vec{x} = N$. The kinetic energy is

$$E_K = \int \cdots \int \sum_{j=1}^{N} \frac{-\hbar^2}{2m_j} \nabla_j^2 \Psi \ \mathrm{d}^3\vec{x}_1 \cdots \mathrm{d}^3\vec{x}_N. \tag{34}$$

Then

Theorem 11 (Lieb-Thirring multiparticle kinetic energy bound) *If all N of the particles are* fermions *each with at most*[33] *s spin states and each of mass $\leq m$, then the kinetic energy E_K obeys*

$$E_K \geq \frac{3}{2}\left(\frac{3\pi}{2s}\right)^{2/3} \frac{\hbar^2}{2m} \int \rho(\vec{x})^{5/3} \mathrm{d}^3\vec{x}. \tag{35}$$

If some of the N particles may be bosons, then this same inequality remains valid if bosons are regarded as fermions with $s = N$ spin states[34].

Lieb and Thirring's inequalities, together with our assumptions about the form of the potential function, imply that the wavefunction cannot be "too concentrated" anywhere, i.e. only a small amount of its probability can lie in a small region. Too much concentration would force a huge kinetic energy, which could not possibly be compensated by the potential energy (which depends less severely on the probability density because $1 < 5/3$), and hence would force a huge total energy. But that is forbidden if we assume the total energy is bounded.

More precisely: From (EQ 31) and the concave-∪ nature of the function x^3 for $x > 0$, it follows that the amount M of probability ($0 \leq M \leq 1$) lying in a ball of radius r obeys

$$M \leq \frac{2^{11/3} m}{3^{5/3}\pi^{2/3}\hbar^2} E_K \, r^2 \tag{36}$$

if we have only one free particle. If we have N particles, then (EQ 35) as $r \to 0$ and the concave-∪ nature of the function $x^{5/3}$ for $x > 0$ shows that

Lemma 12 *If M is the (integrated total) amount of particle probability density ($0 < M < N$) that a particle lies inside a ball of radius r, then*

$$\frac{3^{1/3} 9\hbar^2}{16 s^{2/3} m} \frac{M^{5/3}}{r^2} \leq E_K \tag{37}$$

or equivalently

$$M \leq \left(\frac{16^3 s^2 m^3}{3^7 \hbar^6}\right)^{1/5} E_K^{3/5} r^{6/5} \tag{38}$$

where E_K is the kinetic energy, N is the total number of particles, m is the maximum particle mass, and s is a constant for Fermions (e.g. 2 for electrons) but $s = N$ for bosons.

Suppose the interparticle potentials can have attractive singularities at worst like $-r^{-\gamma}$ for $r \to 0$, for some γ, $0 < \gamma < 2$. (In the usual Coulombic case, $\gamma = 1$.) Then Lieb's bounds also imply that the potential energy E_P cannot be very large if the total energy E is bounded. Because: that would imply huge probability density concentration, which would force E_K to grow at least as fast as $E_P^{2/\gamma}$, i.e. faster than E_P, leading to a contradiction

[31] As Lieb and Thirring pointed out, this is really just a special case of a "Sobolev inequality" [93].

[32] More precisely, E_K is the expectation value of the kinetic energy. We often abuse notation throughout this paper by saying "energy" when we more precisely mean "expectation value of energy."

[33] Actually, really (EQ 33) was stated only for $s = 1$. Sums over spin states, and additional Ψ dependence on those spin states, have to be put in to EQ 33 if we want $s > 1$ [65][63].

[34] I.e. so that the Pauli exclusion principle has no effect.

with the demand that the total energy $E = E_K + E_P$ be bounded.

Lieb's bound also implies that the kinetic energy E_K cannot be very large if the total energy E is bounded, because that would force very large potential energy due to $E = E_K + E_P$, which we just said was impossible.

More precisely: First, consider 1 particle in a spherically symmetric potential $-Cr^{-\gamma}$. Assume the probability density function ρ is spherically symmetric (we will see later that this assumption may be dispensed with). Then the potential energy $E_P = \int \rho V \mathrm{d}^3\vec{x}$ obeys

$$-E_P = C \int_0^\infty r^{-\gamma} \rho(r) 4\pi r^2 \mathrm{d}r \tag{39}$$

and the lower bound (EQ 31) on the kinetic energy is

$$E_K^3 \geq (\frac{\pi^2}{4})^2 (\frac{3\hbar^2}{2m})^3 \int_0^\infty \rho(r)^3 4\pi r^2 \mathrm{d}r \tag{40}$$

Now the requirement that the total energy $E = E_K + E_P$ be bounded prevents either E_K or $|E_P|$ from being too large. For example if we maximize $-E_P$ by choice of ρ subject to

1. an upper bound on {our lower bound on E_K}
2. the normalization constraint $4\pi \int_0^\infty \rho r^2 \mathrm{d}r = 1$,
3. $\rho(r) \geq 0$,

then we find by the method of Lagrange multipliers that the severest $\rho(r)$ obeys

$$\rho = A\sqrt{(r^{-\gamma} - B)_+} \quad \text{where} \quad x_+ \equiv \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{41}$$

for some positive constants A and B chosen to force the constraints to be satisfied. Then the integral (EQ 39) defining E_P is *bounded* if $\gamma < 2$. Hence[35] for one particle, E_P is bounded by some function depending polynomially ♣ at worst on E_K, $\hbar$, the particle masses m, and the number N of particles. (Actually, the dependence on E_K must be linear for large E_K with E fixed, of course.) The assumption of spherical symmetry of ρ may be dispensed with by using the concave-∪ nature of x^3 for $x > 0$ to argue that the *spherical average* of rotations of any alleged non-symmetric maximizing ρ, would in fact be at least as good[36]. Also, more generally, for any potential V bounded above $-Cr^{-1.99}$ for some positive constant C, one can easily see that this argument still goes through.

We'll now redo this argument but with N particles instead of 1. Assume they are each in spherically symmetric potentials $-Cr^{-\gamma}$. Again assume the probability density function ρ is (multiply) spherically symmetric. Then the potential energy (assuming maximally conservatively that this same scenario is repeated N times, once for each of N particles, and each of the particles sees an N-times larger potential well) is

$$-E_P \leq N^2 C \int_0^\infty r^{-\gamma} \rho(r) 4\pi r^2 \mathrm{d}r \tag{42}$$

and the lower bound (EQ 35) on the total kinetic energy is

$$E_K \geq N \frac{3}{2} (\frac{3\pi}{2s})^{2/3} \frac{\hbar^2}{2m} \int_0^\infty \rho(r)^{5/3} 4\pi r^2 \mathrm{d}r \tag{43}$$

Now the requirement that the total energy $E = E_K + E_P$ be bounded prevents either E_K or $|E_P|$ from being too large. For example if we maximize our upper bound on $|E_P|$ by choice of ρ subject to the same 3 constraints as before, then we find by the method of Lagrange multipliers that the severest $\rho(r)$ obeys

$$\rho = A(r^{-\gamma} - B)_+^{3/2} \quad \text{where} \quad x_+ \equiv \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

for some positive constants A and B chosen to force the constraints to be satisfied. Then the integral (EQ 42) defining E_P is *bounded* if $\gamma < 6/5$. Hence

$$E_P = O(E) \quad \text{if} \quad \gamma < \frac{6}{5}, \tag{45}$$

where the constants in the "O" depend polynomially at worst on $\hbar$, the particle masses m, and the number N of particles. Again the assumption of spherical symmetry of ρ may be dispensed with by using the concave-∪ nature of $x^{5/3}$ for $x > 0$ to argue that the spherical average of rotations of any alleged non-symmetric maximizing ρ, would be at least as good. Also, more generally, for any potential V bounded above $-Cr_{\min}^{-1.19}$ for some positive constant C, one can easily see that this argument still goes through[37]. So:

[35] The Schrödinger Hamiltonian corresponding to the $-k/r^2$ potential (where r is the distance to an attracting singularity) in $\mathbf{R}^3$ is *not* self-adjoint if k is any sufficiently large constant [71] [69] [[60], §35 "fall of a particle to the centre"]. Thus, physical systems involving such potentials are not "reasonable," according to our definition of the term. Quantum mechanics is essentially unuseable if such Hamiltonians are admitted [47]. It is interesting that this fundamental breakdown of quantum mechanics happens at exactly the same place as the breakdown of our energy-bounding lemma in the single particle case.

[36] An alternative proof would involve "Steiner symmetrization" [21] of ρ about a hyperplane. This decreases E_P for any potential monotone in r, but leaves the lower bound (EQ 31) unaffected.

[37] Note that as far as *this* section is concerned, only *negative-infinite* (attractive) potential singularities need to be excluded. Repulsive singularities cause no problems. This is because kinetic energy is positive. Hence the only thing we need to worry about here is large negative potential energies combined with large positive kinetic energies. But *later*, in §9.5, we shall also need to forbid repulsive potential singularities $\propto r^{-\gamma}$ with $\gamma \geq 2$, in order to get good bounds when approximating integrals by sums using the bounds in §3 applied to integrands which are complex exponentials of "classical actions" (EQ 62). (Also, the condition $\gamma < 2$ suffices to force the interparticle potentials to be Rollnick ♣, which by theorem X.19 of vol. 2 of [75] forces the Schrödinger operator to have a self-adjoint extension so that quantum mechanics is well posed.) Note that we permit more severe repulsive singularities (our requirement is $\gamma < 2$) than attractive ones ($\gamma < 6/5$). Nevertheless we must still exclude the Lennard-Jones potential (EQ 6) because of its repulsive r^{-12} singularity. Perhaps there is some way to further strengthen our later analysis so that more severe repulsive singularities are permissible. However in most scenarios such a strengthening wouldn't make much difference because repulsive singularities are attractive for opposite charges and hence would need to be regularized anyway.

Lemma 13 (Energies are bounded) *Both the kinetic energy E_K and and the potential energy E_P have absolute values bounded by a polynomial function of the number N of particles, the maximum particle mass m, and the total energy E, if the interparticle potentials are bounded above $-Cr_{\min}^{-\gamma}$ for some $\gamma < 6/5$.*

Hence for the purposes of looking at Church's thesis (even the strong variant), from now on we can just regard E_K and E_P as bounded in all our analysis.

6.1 Regularizing potentials

These bounds are going to be key for allowing us to regularize the potential function V. The idea is to cut V off at some energy cutoff. The cut should not be a flat plateau, since that would cause "corners;" we actually want to "round off the corners" in order to keep $\|V\|_\infty$, $\|\vec{\nabla} V\|_\infty$ and $\|\nabla^2 V\|_5$ all bounded.

Unfortunately, the quantum system with regularized potentials and the original quantum system will evolve differently. We need to prove the difference will not be very large in the L_2 norm.

The proof of that is

1. By lemma 12, there cannot be very much probability M in the (necessarily small volume) "bad" region in which the regularized and original potentials differ.

2. Regard probability density as a conserved[38] "fluid."

3. Even if all the fluid in the bad region for the regularized potential becomes totally orthogonalized (or, even worse, antiparallel[39]) to the true wavefunction and stays that way forever (this is a worst case assumption), still the most the two wavefunctions will be able to differ in the L_2 norm, will be bounded by the total fraction F of fluid that ever enters the bad region during the entire duration of the experiment/simulation.

4. Starting with mass 1 of fluid, then in order for a fraction F of it to enter a bad ball of radius r (and which always contains at most a small amount $M < F$ of fluid mass inside it; and the same for the concentric ball of radius $2r$) during a timespan t, the fluid must have, at some moment, large total kinetic energy E_K. Specifically, $E_K \geq mv^2/2$, where
$$v \;\geq\; \frac{F-M}{4\pi r^2 t}\,\frac{4\pi r^3}{3M} \;=\; \frac{(F-M)r}{3Mt}. \qquad (46)$$
This inequality was derived by realizing that it would be an equality if the fluid flowed into the ball, of surface area $4\pi r^2$, at a rate uniform everywhere on the ball's surface and not varying with time. The flow velocity would then be the flux $(F-M)/(4\pi r^2 t)$ into the surface divided by the fluid density $\rho = 3M/(4\pi r^3)$, which we also take to be uniform. (Actually, the fluid must also manage to get *out* of the ball, a fact we may and shall ignore since its consideration would only make our inequality more true. We instead for simplicity imagine that once the fluid enters the ball it vanishes into some other dimension.) The point is that if a fixed amount $F - M$ of probability-fluid of uniform density ρ manages to move into a ball of small radius r in a small time t, it must have moved fast. Now to allow for the fact that the flux and density might actually *not* be uniform:
   (a) We used the concave-$\cup$ character of the function v^2; the mean-square speed of a fluid particle always is at least the square of the mean speed, so that distributing the fluid flux nonuniformly in space or time would only make our inequality more true.
   (b) Due to the concave-$\cup$ character of $1/\rho$ (or $1/\rho^2$), the same remark applies for distributing the probability density nonuniformly in space or time.

   Now by using lemma 12 to bound M
$$v^8 \;\geq\; \frac{3^7\hbar^6}{16^3 s^2}\,(\frac{F-M}{3mt})^5\, r^{-1}. \qquad (47)$$

5. But the kinetic energy E_K is bounded by assumption (and see lemma 13).

6. Hence, regularizing the potential will introduce boundable L_2 errors into the computation; in (EQ 47), if E_K is bounded, F is also.

We conclude

Theorem 14 (Relation between regularization energy cutoff and wavefunction error) *If the potential energy function is altered when one or more particles lie within a ball of radius r (*any *alteration is permissible provided the Schrödinger operator remains essentially self adjoint and the difference between the altered and unaltered potentials has gradient in L^1 and singular behavior bounded above $-Cr^{-\gamma}$ for some $\gamma < 6/5$ and $C > 0$), then the total L_2 perturbation in the wavefunction after time-evolution for a time t will be*
$$\leq\; (\frac{16^4 s^2}{3^7\hbar^6}(\frac{E_K}{m})^4)^{1/5} 3tr^{1/5} + (\frac{16^3 s^2}{3^7\hbar^6})^{1/5}(\frac{2E_K}{m})^{3/10} r^{6/5}. \qquad (48)$$

[38] Since Schr̆odinger time evolution conserves it [78] both locally and globally.

[39] This argument about the "fluid" might seem to be merely heuristic. But it may be justified by using the linearity of Schrödinger time evolution, i.e. Green's functions (also called "propagators"). The wavefunction t later is a superposition of Green's functions weighted according to the wavefunction now. If all the Green's functions for points in the "bad" region are arbitrarily scrambled up by some unitary transformation, but the others are unaffected, then if the bad region has small amount of probability we can conclude (in view of L_2 norm conservation for Green's function – or any other – time evolution) fully rigorously that the wavefunction time t later really does get altered by a small amount in the L_2 metric.

6.1.1 Regularizing Coulomb potentials

For Coulomb potentials, the bad regions will be balls of radius r centered at the point charges. We regularize the potentials by replacing each point charge by a uniform distribution of charge within that ball. In that case, consider the potential V between two charges q_1 and q_2 separated by distance s:

$$V_{12}^{\text{reg}} = \frac{q_1 q_2}{4\pi\epsilon_0} \cdot \begin{cases} s^{-1} & \text{if } s \geq r \\ 2r^{-1} - s^2 r^{-3} & \text{if } s < r \end{cases} . \tag{49}$$

Actually we have given the formula under the assumption that q_1 is a radius-r ball but q_2 is still a point. However, this is still a perfectly good method of regularization (provided we use $V^{\text{reg}} = \frac{1}{2}\sum_{j\neq k} V_{jk}^{\text{reg}}$ to average the j-ball & k-point and j-point & k-ball potentials) and it is simpler than the true ball-ball potential would have been. Then

$$\|V_{12}^{\text{reg}}\|_\infty \leq \frac{q_1 q_2}{4\pi\epsilon_0} r^{-1}, \tag{50}$$

$$\|\vec{\nabla} V_{12}^{\text{reg}}\|_\infty \leq \frac{q_1 q_2}{4\pi\epsilon_0} r^{-2} \tag{51}$$

and for any p with $0 < p \leq \infty$,

$$\|\nabla^2 V_{12}^{\text{reg}}\|_p \leq \frac{3 q_1 q_2}{4\pi\epsilon_0} (\frac{4\pi}{3})^{1/p} r^{3/p-3}. \tag{52}$$

(Actually, these bounds are equalities if only the first charge is regularized and the second remains a point.)

Of course if we have N charges then these bounds could increase by a factor at most N.

6.1.2 General 2-body distance-dependent potentials

If the potential is a sum of 2-body potentials, each of which depends only on interparticle pair separations s_{jk}, then regularization is easily accomplished in exactly the same manner as before; we cut off the interparticle potentials V_{jk} below some critical value r of s, and replace them in the region $s < r$ with an appropriate "Harmonic oscillator" potential (that is, a quadratic function of s) chosen to match $V_{jk}(s)$ and $V'_{jk}(s)$ at $s = r$.

6.1.3 Finite square wells

"Square well" potentials, such as

$$V = \begin{cases} A & \text{if } a < r < b \\ B & \text{otherwise} \end{cases} \tag{53}$$

(a, b, A, B ($A < B$, $0 < a < b$ are constants) can be handled by using a regularization method which "rounds off the corners." In this case, the "bad" regions are *not* balls, but rather spherical shells of small thickness and radii $\approx a$ and $\approx b$. However, it is possible to modify the statement and proof of theorem 14 to handle this, with the result that the reciprocal of the shell thickness needs only to be polynomially large to make the error caused by regularization suffciently small.

6.1.4 Infinitely high walls

For particles which are "rigid impenetrable balls" (such as in the "Sutherland empirical potential" [42]) the potential is $+\infty$ below some radius. Surprisingly, it even seem possible to handle these, but in a different manner. We do not do any regularization, but instead forbid the region with $V = +\infty$ via a boundary condition. See §7.

6.1.5 Regularizing more general potentials

It is trickier to regularize more general potentials which may not just be sums of 2-body potentials each of which have spherical symmetry.

The method we employ must simultaneously satisfy the constraints that

1. $\|V\|_\infty$ and $\|\vec{\nabla} V\|_\infty$ and $\|\nabla^2 V\|_5$ are bounded after regularization.
2. The regularized and original versions of V differ only in at most a polynomial number of regions of small diameter.
3. V is computable efficiently.

One method which often works is as follows. Consider the 2-time differentiable function

$$\text{squish}(x) = \begin{cases} -2 + \exp(1+x) & \text{if } x \leq -1 \\ x & \text{if } |x| \leq 1 \\ 2 - \exp(1-x) & \text{if } x \geq 1. \end{cases} \tag{54}$$

Replace V by $V^{\text{reg}} = C\text{squish}(V/C)$ where C is an appropriately chosen energy cutoff. The resulting potential will agree with the original one wherever its absolute value was $\leq C$, but will obey

$$\|V^{\text{reg}}\|_\infty \leq 2C. \tag{55}$$

If V happened to have the property that its gradient and/or Laplacian only became unbounded where V itself did, and then only as a power law (or sufficiently slow-growing exponential) in V, then this regularization also will suffice to keep $\|\vec{\nabla} V^{\text{reg}}\|_\infty$ and $\|\nabla^2 V^{\text{reg}}\|_\infty$ bounded.

However this particular "squish" method fails for potentials such as $V(x) = 5 + \sqrt{x}$ and $V(x) = \ln(x)^2$ as $x \to 0^+$. In those cases, though, alternative squishing functions do work.

6.1.6 General purpose regularization method

The following method works in all cases in which $\vec{\nabla} V$ and $\nabla^2 V$ are differentiable and efficiently computable[40]. If $|V| > C$ then[41] replace V by

$$V^{\text{reg}} = (\operatorname{sign} V) \times \left[2C - C\exp\left(1 - \frac{|V|}{C} - [|V| - C]^4[|\vec{\nabla} V|^4 + |\nabla^2 V|^4]\right)\right]. \tag{56}$$

[40] E.g. polynomial time computable in the sense of Ko [58] ♣.

[41] Actually, it may not be efficiently decidable whether $|V| > C$ (or, for that matter, whether $a > b$ for any two computable reals a, b) [58]. However, these issues do not matter in this situation; the overall function V^{reg} is easily seen to be efficiently computable if V is.

The idea is that functions like $\exp(-1/x)$ are extremely well behaved as $x \to 0^+$.

It is essential in all of these regularization techniques (in order for step 4 of the proof of theorem 14 to work) that the potential be "bad" in a *ball-like* region, i.e. with surface area proportional to volume$^{2/3}$ in the limit of small volume, and this notion has to be an effective ♣ one. (Actually this requirement may be relaxed, e.g. see §6.1.3, but we'll restrain the urge for maximal generality.)

6.2 Energy conservation

If we use the Trotter product formula (§5.3) to simulate Schrödinger time evolution in a regularized potential, then the *true* wavefunction's evolution in that potential both conserves L_2 norm and conserves energy. The *simulated* wavefunction's evolution also conserves L_2 norm (at least, up to roundoff error; because the Trotter product formula is unitary) but it does *not* necessarily conserve energy. It thus remains conceivable at this point that the simulated wavefunction could acquire some huge energy, at which point all our arguments (e.g. lemmas 12,13) based on boundedness of energy would lose their force.

Fortunately, this possibility can be ruled out. The idea is that in the "alternation in time" view of the Trotter product formula (§5.3), we alternate between two phases, one of which multiplies the wavefunction by a position-dependent phase factor and hence exactly conserves potential energy E_P, and the other of which is free particle propagation and hence exactly conserves kinetic energy E_K. Now the E_P-conserving phase will, in a potential V with bounded $\|\vec{\nabla} V\|_2$ (call this norm bound B), in a span of time Δt, cause kinetic energy to increase from E_K to at most $(\Delta t B/\hbar + \sqrt{E_K})^2$. This[42] is easily seen to follow from (EQ 32) and (EQ 34).

Thus, the kinetic energy cannot become superpolynomially large. Hence, the lower bounds (EQ 31, 35) on E_K cannot get large either, which in turn prevents E_P (for either the regularized or original potentials) from becoming large by the same argument as we used before in the proof of lemma 13. We conclude that E_K, E_P and the total energy, while not exactly conserved, nevertheless are incapable of becoming huge. That suffices. More precisely, any bound depending at worst polynomially on B and E_K and t, suffices for our purposes[43] because we can just agree to use our bound on E_K^{maximal} rather than E_K^{initial} everywhere.

6.2.1 An aside

I must comment that my argument here looks appallingly weak. Really, at least in most practical scenarios, I would expect energy to be approximately conserved by the simulation. Indeed I would hope[44] that nonconservation could be shown to be arbitrarily small as $n \to \infty$ (where n is the number of timesteps performed by the simulation) like some negative power of n. However, perhaps due to some mental weakness, I have not proven that.

More generally, it is an interesting point that our simulations could be perfectly good, in the sense of §2.3 of being statistically indistinguishable from the real thing, despite featuring expectation values of certain quantities which differ hugely from the real thing. This is because *it is impossible to measure an expectation value* (for example E)! Although it is of course possible to estimate expectation values indirectly from the results of other repeated measurements, our usual arguments then suffice to show indistinguishability.

7 Feynman's path integrals: definitions and reformulation of quantum mechanics

Consider a continuous "path" $\vec{x}(u)$ where u lies in some real time interval $[0, t]$ and $\vec{x}$ is a d-dimensional moving point. It is possible to consider real valued functionals f of paths. The functional (EQ 57) arising in Feynman's path integral treatment of quantum mechanics is the "classical action" functional (measured in units of $\hbar$)

$$f[\vec{x}(u)] = \frac{1}{\hbar} \int_0^t \left[\frac{m}{2} |\dot{\vec{x}}(w)|^2 - V(\vec{x}(w)) \right] \mathrm{d}w \qquad (57)$$

where V is the potential energy function and more generally the term $\frac{m}{2}|\dot{\vec{x}}|^2$ should be replaced by a kinetic energy, and a different action formula is needed in the presence of vector potentials, e.g. for charged particles in magnetic fields. (The right hand side is the "classical action" measured in units of $\hbar$.) However, we will allow a much more general class of functionals than this. Anything computable for piecewise constant velocity motions along polygonal paths to P decimals of accuracy in polynomial (in P and the number of sides of the polygon) time, will do (and even this is vastly generalizable, see the end of this section).

Then Feynman [28] defined a notion

$$\int_0^t e^{if[\vec{x}(u)]} \mathcal{D}\vec{x}(u) \qquad (58)$$

of a "path integral." (It is also possible to consider integrands which are not necessarily complex phase factors. However, in this paper, we will restrict our attention to phase factor integrands, since they lead to propagators which are unitary transformations.) What might be

[42] Note if $\Delta t = t/n$ then $[(\Delta t B/\hbar + \sqrt{E_K})^2/E_K]^n$ grows only polynomially with $n \to \infty$.

[43] These purposes are: (1) bounding error due to regularization of the potential, and (2) bounding error in approximating integrals by sums.

[44] I can show this under certain unproved hypotheses that seem plausible in most practical scenarios. However, [97] showed that it is impossible for any simulation of a quantum "non-integrable" system by a quantum computer to conserve energy exactly, while at the same time featuring exactly unitary evolution. This is a quantum analogue of the "Ge-Marsden theorem" [33] that no exactly "symplectic" integrator can exactly conserve energy (for simulating classical non-integrable Hamiltonian systems of a certain type).

called the (n,g,P,H)-approximation to this path integral is the sum

$$\sum f[\vec{x}(u)] \quad (59)$$

where we sum over the f-values corresponding to polygonal paths with n line segments, in which the vertices of the line segments are nodes of a $2^g \times 2^g \times \cdots \times 2^g$ hypercubic d-dimensional grid lying inside a hypercube of side H, and in which the velocity $\dot{\vec{x}}$ is a piecewise constant function of u, with n pieces, i.e. the curve $\vec{x}(u)$ is an n-gonal path, chosen so that the $n+1$ vertices are arrived at at n moments equally spaced in the time interval $[0,t]$ (with the first and last vertices being at 0 and t respectively). Finally, to explain P, we assume that the computation of this sum, is only done to P places of decimals. This could be accomplished by evaluating each summand e^{if} to $P^+ = P + O(gd)$ decimal places, i.e. by carrying $O(gd)$ "guard digits." Feynman essentially defined his path integral (EQ 58) as

$$\lim_{n\to\infty} \lim_{H\to\infty} \lim_{g\to\infty} \lim_{P\to\infty} \left[H 2^{-g} Q(n) \sum f[\vec{x}(t)] \right] \quad (60)$$

where $Q(n)$ is a specific "normalizing factor," which in Feynman's treatment of quantum mechanics was $[mn/(iht)]^{n/2}$. If we take the inner three limits this is

$$\lim_{n\to\infty} Q(n) \int\int \ldots \int \exp\left[i f_{\rm disc}(\vec{x}_0, \vec{x}_1, \vec{x}_2, ..., \vec{x}_n) \right] \mathrm{d}^d\vec{x}_1 \mathrm{d}^d\vec{x}_2 \ldots \mathrm{d}^d\vec{x}_n \quad (61)$$

where with Feynman's f from (EQ 57),

$$f_{\rm disc}(\vec{x}_0, \vec{x}_1, \vec{x}_2, ..., \vec{x}_n) = \sum_{j=1}^{n} \frac{mn}{2t} |\vec{x}_j - \vec{x}_{j-1}|^2 - \frac{t}{n} V(\vec{x}_j), \quad (62)$$

and more generally $f_{\rm disc}$ stands for the functional for the n-sided polygonal path with vertices $\vec{x}_0, \vec{x}_1, \vec{x}_2, ..., \vec{x}_n$.

If only paths starting from a specific point are allowed, the path integral is a function of that starting point. If starting and ending points are both prescribed, the path integral becomes a "propagator;" and (at least, after assuming enough convergence and uniqueness) Feynman demonstrated the identity of this propagator with the Green's function (i.e. the kernel of the time evolution operator) for Schrödinger's equation of the quantum system with the given classical action function. By starting at a complex superposition of starting points (i.e. a wavefunction) the path integral gives (as a function of the path's endpoint) the time-evolved wave function.

Although Feynman's argument [28] had indicated that his path integral would obey the correct Schrödinger time evolution equation, there is still the matter of boundary conditions. These are easily dealt with. If we are dealing with a quantum system confined to some spatial region, then simply define the potential energy to be infinite outside that region. (Equivalently: all paths passing through the forbidden region have their phase angles completely randomized.) Of course multiparticle systems are viewed as a single "superparticle" in a higher dimensional space. Fermionic symmetry may be dealt with by regarding it as a boundary condition too (that is, if there are two identical fermions in the same location, the "potential" is infinite; indeed we may require identical fermions to be in sorted order according to any ordering notion we desire, and define the potential to be infinite otherwise, thus implementing the Pauli exclusion principle as a boundary condition).

"Hard balls" may also be regarded as such a boundary condition, i.e. for paths involving interpenetrating balls at some point, the phase angle is totally randomized. (The later discussion in §9.5 makes it seem that by using exponentially fine grids and exponential numerical accuracy, such randomization will, with very high probability, yield the desired effect...) But this is fundamentally different in the sense that Fermion boundary conditions and the condition that all particles must lie in a convex box, both represent a *convex set* of allowed configurations, i.e., it is easy when considering paths made of line segments, to never consider a forbidden path. For hard balls, we actually do consider forbidden paths, but hope to make their contributions approximately cancel out (to exponentially good precision) with intentional phase-angle randomization.

Thus, time evolution in quantum mechanics is just a path integration problem, and the problem of simulating quantum mechanics in the sense of §2 is just a problem of sampling values from the distribution of a function F of a path integral – precisely the problem we will consider in §8[45].

7.1 *Feynman's path integral is an instance of the Trotter product formula*

Clearly the combination of (EQ 61) and (EQ 62) may be interpreted as involving an n-term product of matrices with continuous indices, i.e. operators.

More generally that is true if $f_{\rm disc}$ is a sum with n terms as in (EQ 62), where the jth term is a fixed function of x_j and x_{j-1}.

Also, in our particular case (quantum mechanics, EQ 62), we may view the Schrödinger Hamiltonian operator as a sum of a potential energy term and the free-particle Hamiltonian. In that case, by using the 2-term product (EQ 11) to approximate each term of the product we can view path integral (EQ 61) as just an instance of the *Trotter* product formula (EQ 28). (And this observation immediately proves its convergence.)

The potential energy part of the Hamiltonian is just a multiplication by a position dependent function in the position basis. The free-particle part is just a multiplication by a momentum dependent function in the momen-

[45] It would also be possible to consider, rather than a sum over polygonal (piecewise differentiable) paths, instead a sum over paths which were piecewise 2-time or 3-time differentiable, e.g. cubic splines. In this case, the definition of the functional f could also involve the second time derivative $\ddot{\vec{x}}(t)$. And so on: still more derivatives could be allowed. Without dwelling on precisely how this should be done, we claim that everything in this paper is still applicable to any such redefinitions of path integration.

tum (more precisely, Fourier mode) basis. Thus (with the aid of the "quantum Fourier transform" of Shor [81] to convert between these two bases) it is easy, algorithmically, to *apply* the operators in these products. Hence we can algorithmically use the Trotter product formula to time-evolve a wavefunction.

Alternatively we could stay in the position basis by viewing (EQ 62) as a position dependent function. (Either way works.) This idea will be the basis of §8.

In either case: because, e.g. we are only multiplying our wavefunction by V *at grid points* rather than everywhere, we are only getting an approximation to applying the V operator, essentially because we are replacing an integral by a Riemann sum. Errors due to such approximations may be bounded immediately using the techniques in §3 and §2.7 and using norm bounds about V (after regularization) and Ψ and their derivatives.

8 Quantum algorithm for approximate path integration

We will show how to perform a "timestep." The entire path-integration algorithm consists of n timesteps, each one of which consumes $O(gd)$ new qubits. The total space consumption of the n-timestep algorithm is $O(ngd+P+g+\log H)$ qubits. These qubits are assumed initially to be available in the all-0 state, i.e. the state $|000...000\rangle$ with amplitude 1 and all other states with amplitude 0.[46]

Step 1: "Randomize" $N = gd$ of the qubits by performing a "45° phase rotation" on each of them, i.e. applying

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} / \sqrt{2} \tag{63}$$

in the 2-state basis $\{|0\rangle, |1\rangle\}$. This causes the new state to be: whatever it previously was on all the other qubits, tensored with the state

$$\frac{1}{2^{N/2}} \sum_{2^N \text{ bitstrings } s} |s\rangle. \tag{64}$$

Step 2: Regard s as the $N = gd$ bits of the d coordinates, in base 2, of a point in a d-dimensional hypercubic grid with 2^g gridlines in each direction[47]. Compute $f(s)$ to $P^+ = P + gd$ decimal places and store it in $O(P^+)$ more qubits. By assumption, this may be accomplished in a number of steps polynomial in N and P^+. (Note: we are using "quantum parallelism" here since all 2^N of the $f(s)$ values are automatically computed "simultaneously.") Incidentally, although we've written it that way for concreteness, there is no requirement that f be a function of s alone. It could also depend on the s values from previous timesteps, the fixed value H, the number n of timesteps we intend to perform, and any number of precomputed constant values. Really, of course, all this information as well as an $|000\rangle$ for the blank as-yet unused quantum registers, is also in the state description below, although we've adopted the convention of not writing that down. We now have[48] the state

$$\frac{1}{2^{N/2}} \sum_{2^N \text{ bitstrings } s} |s, f(s)\rangle. \tag{65}$$

Step 3: Rotate the complex amplitude of the state $|s, f(s)\rangle$ by multiplication by $e^{if(s)}$. This may be accomplished with the aid of $O(P^+)$ different 1-qubit quantum gates which perform multiplication by $\exp(i2^{-j})$ for $j = 1, 2, \ldots O(P^+)$; the jth such gate is invoked if the 2^{-j} bit in the binary representation of $f(s)$ is a '1.' We now have the state

$$\frac{1}{2^{N/2}} \sum_{2^N \text{ bitstrings } s} e^{if(s)} |s, f(s)\rangle. \tag{66}$$

Step 4: "Uncompute" the $f(s)$ bits (causing them to return to all-0's) to get the state

$$\frac{1}{2^{N/2}} \sum_{2^N \text{ bitstrings } s} e^{if(s)} |s\rangle. \tag{67}$$

Readers who prefer not to be ensnared in unnecessary detail can view all 4 of these steps as just 1 step with the net effect of transforming (EQ 64) to (EQ 67).

Recap: We have now performed a timestep. After n such timesteps we have the state

$$\frac{1}{2^{nN/2}} \sum_{\substack{2^{nN} \text{ bitstrings} \\ s_1, s_2, \ldots, s_n \\ \text{(each } n \text{ bits)}}} e^{if(s_1)} e^{if(s_2)} \cdots e^{if(s_n)} |s_1 s_2 \cdots s_n\rangle. \tag{68}$$

We now compute some function F of the bits s_n describing the final state (and again, F could also depend upon all the non-final s_j) to P decimal places, and measure the $O(P)$ qubits describing the value of F. We are done.

To see that any function F expressible in terms of particle position and momentum data can be handled, we observe that we can simply store bits of the position coordinates of the final state $|s_n\rangle$ in some quantum register, and we can then switch bases to a momentum basis efficiently because – it turns out – the "quantum Fourier transform" of Shor [81] is precisely the operation needed to convert between the momentum and position bases. Then we can store the momentum describing bits for the final state in some quantum register. Then, we compute whatever function we want of these two (position

[46] If the qubits were implemented inside a real quantum computer using electron spins, then this initial "N zeroes" state could be prepared with high probability of achieving any fixed accuracy by, e.g. slowly cooling the electrons in a high magnetic field $\vec{B}$ to a temperature T at which $k_B T \log N \ll \mu_e |\vec{B}|$ where k_B is Boltzmann's constant and μ_e is the magnetic moment of an electron.

[47] A fixed scaling factor occupying $O(g + \log H)$ classical bits describes the sidelength of the hypercube enclosing this grid.

[48] This has been analogous to the way Shor [81] got from his equation 5.1 to 5.2 on his page 498-9. We say this to make it clear our steps are not forbidden by the "no-cloning theorem." ♣

and momentum) quantum registers, and store the result in some other quantum register. Finally we classically measure this to get our sample. For example, the usual energy operator $V(\vec{x}) + \sum_j |\vec{p}_j|^2/(2m_j)$ is expressible in terms of the momenta $\vec{p}_j$ of the jth particle, plus the position coordinates $\vec{x}$, and hence we can get a energy sample from a simulated physical system. Similarly for angular momenta and so forth.

Result: we now have extracted a sample from the probability distribution of F corresponding to the time-evolved final state s_n. We assume at the beginning that the desired complex amplitudes corresponding to the possible initial states $|s_0\rangle$, were available in our quantum memory. The algorithm is done.

9 How good are approximate path integrals?

There are three kinds of errors we have to worry about when approximating path integrals. The first kind of error arises from our discretization of space (2^g gridlines in each direction). The second kind arises from discretizing time (n timesteps). The third kind of error arises from the use of finite precision arithmetic instead of exact real arithmetic (getting a result accurate to P decimal places at best).

Actually there is a fourth kind of error too, if we use a finite box of sidelength H instead of $H \to \infty$ (an infinite universe). However, in this paper we have defined this error to be zero by *proclaiming* that the physical system being simulated lives inside a finite size box. This proclamation probably could have been avoided, but that task will be left to future authors.

9.1 *Discreteness of g, P and H*

Because we may take n, g, P and H to ∞ along a wide variety of 4D curves, we have a great deal of algorithmic flexibility in approximating path integrals. Because we may use an exponentially fine grid (e.g. a grid size of 10^{-100} meters) in an exponentially large hypercube (say 10^{100} meters on a side) and exponential numerical accuracy (say carrying 1000 decimal places for accuracy $\pm 10^{-1000}$) with only *polynomial* computational overhead, it seems very plausible that the error in the approximation of a Feynman-Schrödinger path integral caused by the finiteness of g, P, and H, will be negligible for all practical purposes.

There are two ways to justify this.

The first is to use Planck scale arguments $\otimes$ (§9.2, §9.3). Although this is an "easy way out," I do not like it for reasons described in §9.4. The second, which I prefer, is to use the estimates in §3 to bound the error introduced by space discretization; and Trotter product error bounds for regularized potentials (§5.3) to take care of time discretization. We'll describe that in §9.5.

9.2 *Finiteness of n; Planck length argument* $\otimes$

If we restrict our attention to the path integrals arising when simulating real physics, it is generally agreed that the spacetime metric must be "fluctuating" and "wavy" at length scales on the order of the *Planck length* $\ell_p = \sqrt{\hbar G/c^3} \approx 1.6 \times 10^{-35}$meters. Nobody currently knows effective ways to treat that phenomenon, and hence it is generally ignored. In the event that ignoring this phenomenon was justified, presumably no appreciable error results from using a grid size of $\leq \ell_p$ (perhaps even better, exponentially less than ℓ_p). In the event that ignoring this is not justified (S. Hawking has conjectured, for example, that quantum gravitational effects are one unavoidable cause of "decoherence" – a highly noticeable effect), presumably by using a grid of size scale comparable to ℓ_p we can approximate the truth at least as well as the usual course of ignoring ℓ_p entirely.

These arguments suggest that the *only* error of any interest (in practice) will arise from the finiteness of n, i.e. the fact that we have approximated an infinite dimensional path integral by an n-dimensional one, or equivalently that we have used n finite timesteps rather than continuous time. The other approximation, of the finite dimensional integrals by Riemann sums on exponentially fine grids, will introduce negligible errors.

9.3 *Finiteness of n; Planck time argument* $\otimes$

Those who like the Planck length argument should be equally happy about the idea that using timesteps of order smaller than the Planck time $t_p = \ell_p/c \approx 5.3 \times 10^{-44}$sec is presumably pointless. Note, for theoretical purposes, this only introduces a "constant factor slowdown" (admittedly, a huge one) and hence the Quantum Strong Church's Thesis seems justified immediately...

9.4 *Anti-Planck counterarguments*

I am not in favor of the Planck scale arguments of §9.2 and §9.3.

First, those arguments carry no weight if the path integrals under consideration have nothing to do with physics. If we want an approximate path integration algorithm of more general applicability than just to physics problems, therefore, we can't take that way out.

Second, I believe it is the duty of the simulator to simulate Schrödinger's equation with rigorous error bounds. It is *not* to simulate some plausible-sounding heuristic notion of what physics hopefully is like, completely in the absence of rigorous error bounds.

Third, it is a bit embarrassing to say that simulating t seconds of physics "only" requires $2 \times 10^{43} t$ timesteps, i.e. "$O(t)$." It seems preferable to me to use only $2t$ timesteps, where "2," more precisely, represents some reasonable factor which increases polynomially if you want smaller error bounds. We realize full well that in any feasible computation, one is never going to reach the regime where this factor gets to be 2×10^{43}.

9.5 Error in path integration caused by finiteness of n, g, and P

Errors in (EQ 61) caused by the finiteness of g (the $\log_2$ of the number of grid points in each coordinate direction) are exponentially negligible if g is made polynomially large while $\|\vec{\nabla} V\|_1$ is bounded[49], by the results of §3. Hence, this error is readily made effectively ♣ negligible compared with the errors we are allowing ourselves (in claim 4 of §2.3) anyway.

The same is true for the errors caused by using P-decimal arithmetic; by making P polynomially large we get exponential precision. Since essentially the only arithmetic operation under consideration here is *summation* of 2^{gd} summands in fixed point arithmetic, it suffices to use $P^+ = P + O(gd)$ decimals internally, i.e. $O(gd)$ "guard digits."

Finally, we've already discussed n in §5, e.g. lemma 10, with the conclusion that for regularized potentials, the L_2 error would be a reciprocal polynomial ♣ function of n.

In a more general setting, in which we allow path integrals not necessarily arising from physics, we can also get error bounds just like in lemma 10 provided we for some reason know that the unitary operators corresponding to individual integrals in (EQ 61) induce L_2 errors in the wavevector they operate on (versus the value of the true path integral over their t/n timespan) of order n^{-2} as a function of n as $n \to \infty$. Unfortunately, we don't know precisely what class of path integrals are permitted by this criterion, but clearly it is a large class and we know from the Trotter product re-interpretation that it includes quantum mechanics.

Theorem 15 (The path integration algorithm works)
In cases (in particular, quantum mechanics!) in which the finite-dimensional integral may be viewed as a Trotter product, our operator error bounds of lemma 10 will apply to produce bounds showing that it approximates the Feynman path integral with error bound proportional to $1/n$, and the further approximation by a Riemann sum (on a grid of exponential fineness in the hypercubic domain, and carrying a polynomial number of decimal places in the arithmetic) introduces exponentially ♣ small error.

10 The final layer: simulation of a quantum computer by a classical one

It is well known that

Lemma 16 BQP⊆PSPACE. *I.e., less precisely, polynomial time algorithms on a quantum computer may be simulated by a conventional computer in polynomial memory space.*

Proof. In PSPACE, one simply follows all possible computational paths to either a '1' answer or a '0' answer and adds up the complex probability amplitudes for each one to get a net amplitude for getting a '1' or '0' as the final outcome of the computation.

If we are considering computations allowed to output arbitrary bitstrings instead of just yes-no answers, then one could also have an outer loop over all possible outputs, computing – but not storing – the probability amplitude for each one. One needs to produce a sample from the probability distribution, but that is easy. (To produce output n with probability p_n, we simply scan thru the possible outputs, halting at m with probability $p_m / \sum_{n\ge m} p_n$. We may use the fact that $\sum_{n\ge m} p_n = 1 - \sum_{n<m} p_n$ to compute this probability "on the fly.") One also needs to make sure that computing the probability amplitudes in finite precision arithmetic does not cause large error – but that is also easy. Indeed, simulating the amplitudes of individual "quantum gates" accurately to $\pm t^{-O(1)}$ (i.e. carrying $O(\log t)$ bits of precision) suffices [10]; this is not at all onerous considering we can afford to carry $t^{O(1)}$ bits while still staying in PSPACE. □

Indeed, stronger results are known[50].

Theorem 17 (Complexity class inclusions)

$$\mathrm{BPP} \subseteq \mathrm{BQP} \subseteq \mathrm{PP} \subseteq \mathrm{P}^{\#\mathrm{P}} \subseteq \mathrm{PSPACE}, \tag{69}$$

$$\mathrm{PP}^{\mathrm{BQP}} = \mathrm{PP}, \tag{70}$$

$$\mathrm{NP} \cup \mathrm{coNP} \subseteq \mathrm{PP}. \tag{71}$$

Proof. The first claim was shown by Adleman et al. [2], and the second by Fortnow and Rogers [32]. The third is well known [44]. □

For guidance about the acronyms of complexity class theorists, see [44]. The particular classes mentioned here are (roughly): "BQP" is the set of yes-no problems soluble by a quantum computer in polynomial time with error probability $< 1/3$ for the answers it produces. (Of course "1/3" may be made exponentially small by a polynomial number of repeated runs). "PP" is the set of yes-no problems soluble by a conventional computer (equipped with a random number generator) in polynomial time with errors allowed, but for which the probability of getting the right answer always exceeds 1/2. "P" is polynomial (deterministic, worst case) time. "PSPACE" is the class of problems soluble with a polynomial number of bits of memory. "#P" is the class of problems soluble in polynomial time on a "counting Turing machine." "$P^{\#\mathrm{P}}$" means "P with access to a unit-time oracle for solving #P-complete problems on demand." "NP" is the class of problems soluble in polynomial time on a "nondeterministic" Turing machine which always "guesses right;" another characterization is it is the class of yes/no problems whose yes answers come with proofs which may be checked in polynomial time.

[49] This will happen, e.g., if all potential singularities $\propto r^{-\gamma}$ have γ bounded below 2.

[50] Actually theorem 17 leads to a stronger kind of Church thesis than we've claimed in the introduction to this paper. Not only do we get the "intermediate" Church thesis, involving a PSPACE simulation, in fact we get PP and $\mathrm{P}^{\#\mathrm{P}}$ simulations.

10.1 Alternative view #1

It is also possible to avoid simulating a quantum computer by a classical one by never using a quantum computer in the first place. We simply evaluate all the Riemann sums over exponentially large sets directly in PSPACE. Similarly such operations as discrete Fourier transforms on exponentially large vectors and sampling from approximate path integrals are readily seen to be in PSPACE directly.

10.2 Alternative view #2

The initial state of a N qubits inside a quantum computer may be thought of as being described by 2^N complex numbers. Thus, one could argue that the input for any classical simulator, would have to be at least 2^N bits long. In that case, a quantum computer algorithm involving G quantum gate operations, where $G \leq$polynomial(N), could be simulated in time of order roughly 2^{G+N} by a classical simulator, which is *quasipolynomial* ♣ in the classical input length.

The *moral* of this viewpoint seems to be that the *only* reason that quantum computers are (apparently) more powerful than conventional computers, is that their input is available in a "sparse format."

11 Putting it all together – Main Result

The simulation algorithm is roughly as follows.

1. Inputs include (in addition to the initial state, described by N_Q qubits, for which see §2.6, §2.7) a bound on the kinetic energy E_K, the side length H of the enclosing box, the number N of particles, a description of the potential function V, the maximum particle mass m, the duration t of time being simulated, and the (positive integer) run number R.

2. Regularize the potential V (using a cutoff radius selected according to theorem 14 to be small enough to get the desired L_2 error bound) as discussed in §6.1. Compute bounds on $\|V^{\rm reg}\|_\infty$, $\|\vec{\nabla} V^{\rm reg}\|_\infty$, $\|\nabla^2 V^{\rm reg}\|_5$.

3. Use lemma 10 to find the number of timesteps n we need, based on R (the run number, see rule 7 of §1.2 and §2.3), our bound on the kinetic energy E_K, m, and the norm bounds, in order to assure L_2 error bound $< R^{-1.1}$.

4. Similarly find suitable P (number of decimals we will need to keep our arithmetic accurate to) and g (for a 2^g gridline grid in each direction) values.

5. Do the Trotter product formula n-step loop (as in (EQ 28) and §7.1) method for simulating the Feynman path integral that accomplishes the time-evolution. Each step of this loop involves a free-particle time evolution for time t/n (i.e. multiplying each Fourier mode of energy E_m by a phase factor $e^{-itE_m/\hbar}$) and a pure-potential time evolution for time t/n (i.e., multiplying the wavefunction by a position dependent phase factor $e^{-itV^{\rm reg}/\hbar}$). It seems mentally simpler to use this method from §7.1, rather than the general purpose path integration algorithm described in §8, to time-evolve the wavefunction.

6. Sample from the final wavefunction as discussed at the ends of §8 and §1.2.

<u>The main claim of this paper:</u>

A. The final wavefunction will have L_2 error (versus the true wavefunction) that is $\leq R^{-1.1}$ on the Rth simulation run. This satisfies the accuracy goal of rule 7 of §1.2 (see also §2.3) and assures that the probability distributions output by the simulator and the genuine quantum system will be statistically indistinguishable over any unboundedly long sequence of rerun simulations and experiments.

B. The entire simulation algorithm will run on a quantum computer in time bounded by

$$\text{polynomial}(H, N, m, E, \|V^{\rm reg}\|_\infty, \|\vec{\nabla} V^{\rm reg}\|_\infty, \|\nabla^2 V^{\rm reg}\|_5, t, N_Q, R, P). \quad (72)$$

The number of calls to "oracles" for providing more decimals of fundamental physical constants, will also be polynomial. If the simulation is running on a conventional computer (i.e. Turing machine equipped with a random bit generator) it will occupy a number of bits of memory bounded by (EQ 72) and hence will run in a number of steps bounded by 2 raised to this power. As pointed out in §10.2, the runtime will also be bounded by[51]

$$\text{quasipolynomial}(H, N, m, E, \|V^{\rm reg}\|_\infty, \|\vec{\nabla} V^{\rm reg}\|_\infty, \|\nabla^2 V^{\rm reg}\|_5, t, 2^{N_Q}, R, P). \quad (73)$$

In these bounds, N_Q, P, and the norm bounds are *special* because they are really bounded by functions of the *other* arguments of the polynomial (EQ 72). Specifically, $N_Q = O(N \log \frac{H\sqrt{mE}}{\hbar})$ and P is polynomially bounded.

C. The potentials have to be "nice enough" – in the sense discussed precisely in §6.1 and §9.5. Let us merely say here that all the potentials physicists usually write down which involve attractive singularities bounded above $-C_1 r^{-1.19}$ and repulsive singularities bounded below $C_2 r^{+1.99}$, where C_1 and C_2 are known positive constants, will do. Specifically it will suffice if V, $\vec{\nabla} V$, and $\nabla^2 V$ are polynomial time computable real functions in the sense of Ko [58] everywhere except possibly when the minimal separation $r_{\min}$ between two particles is small, and then the dependence on $r_{\min}$ must obey the preceding two bounds and the dependence of the runtime to compute V accurate to $\pm 2^{-x}$ must be bounded by a polynomial of x and $1/r_{\min}$. It is also possible to

[51] Warning: in my opinion (EQ 73) should not be regarded as a quasipolynomial Church thesis (i.e. nearly as strong as the strong, i.e. polynomial, Church thesis) because although it achieves runtime quasipolynomial in the length of its input, I certainly do not regard it as having achieved only quasipolynomial *slowdown.*

drop any requirement that $\vec{\nabla} V$ and $\nabla^2 V$ be computable, if V is nice enough in other ways (see §6.1). If these requirements are satisfied, then the norm bounds $\|V^{\rm reg}\|_\infty$, $\|\vec{\nabla} V^{\rm reg}\|_\infty$, $\|\nabla^2 V^{\rm reg}\|_5$ will be bounded by polynomials of all the non-special arguments of (EQ 72).

D. All the rules in §1.2 and desiderata in §2 are satisfied.

11.1 What about magnetic fields?

Unfortunately (and I have not seen this easy fact pointed out before) *magnetism destroys nonrelativistic quantum mechanics*. Specifically, the classical energy of a point magnetic dipole $\vec{\mu}$ in a magnetic field $\vec{B}$ is $-\vec{\mu} \cdot \vec{B}$, and the magnetic field of (another) point magnetic dipole μ (located at the origin of x, y, z-space and pointing in the z direction) is

$$\vec{B} = \frac{\mu\mu_0}{2\pi} \frac{(-3xz, -3yz, x^2 + y^2 - 2z^2)}{(x^2 + y^2 + z^2)^{2.5}}. \quad (74)$$

In other words, the potential energy function, for a magnetic dipole in the presence of another, stationary, magnetic dipole[52] goes infinite, proportionally to $\pm r^{-3}$ where r is the distance between raised them (and the sign of the infinity depends on the direction of approach and the dipole's orientation; both signs happen).

Thus I claim that a point electron and point nucleus (both of which are magnetic dipoles), obeying the Schrödinger equation or Pauli equation, *would* collapse (even for uncharged electrons and nuclei), because collapsing is what two particles obeying a power-law potential going to $-\infty$ more rapidly than $-r^{-2}$ do [60][69].

But what about the theorem by Lieb, Loss, and Solovej [66] "proving the stability of nonrelativistic matter" in the presence of magnetism? Answer: it is highly misleading in this respect. That theorem depends on including magnetic field energy $\int \vec{B}^2 d\vec{x}$ in the Hamiltonian and then proves a finite lower bound on the ground state energy. However, the self-energy of the $\vec{B}$-field in EQ 74 is $+\infty$, so the Lieb-Loss-Solovej theorem *in no way prevents the collapse of matter* to zero radius, since at arbitrarily close approach the "energy" would *still* be $+\infty$, – certainly obeying the (finite) LL&S lower bound.

So the reader may well now ask: why, in reality, don't hydrogen atoms collapse? Probably for the following reason. It is true it would be arbitrarily energetically favorable to bring the electron very close to the point nucleus (assuming it is brought toward the North pole of the nucleus while the South pole of the electron points South); indeed (to do some dimensional analysis) at electron-proton separation of order

$$r \approx \frac{\mu_e \mu_p \mu_0 m_e}{\hbar^2} \approx 1.3 \times 10^{-17}\,\text{meter}, \quad (75)$$

or below, the magnetic attractive energy would be enough to overcompensate for the Fermi energy required to compress the electron into that small a volume – and so our H atom will collapse. *But*, this separation is about 100 times smaller than the radius of the proton (in reality not a point), about 10^5 times smaller than the Compton wavelength of the electron, and about 10^7 times smaller than the Bohr radius of the H atom.

Therefore relativistic effects, and/or the fact that the nucleus is not a point, will have significant effects, invalidating the application of the Pauli and Schrödinger equations in this region and replacing it presumably with some other equation – which presumably does not collapse.

The usual Schrödinger electron wavefunction for the H ground state has very small overlap with the tiny volume in which this magnetic effect matters much. Hence, the Pauli equation would have predicted a slow rate of "tunneling" or "leakage" of the electron down into that region, at which point collapse would occur (so H atoms would be unstable, but with a fairly long half-life). But the true equation presumably predicts that this tunneling/leaking will not be an energy win, hence the H atom in fact will have infinite half-life and will be stable, and hence the usual Schrödinger approximation (that there are no magnetic effects) will be quite accurate.

So I claim the true explanation for the stability of matter with magnetism is *not* the Lieb-Loss-Solovej proof about point nuclei and nonrelativistic dynamics (that theorem misleads us) but instead must be due to finite size nuclei and/or relativistic effects.

Magnetic dipoles cannot be included in nonrelativistic quantum mechanics per se without destroying it.

But if a uniformly bounded *externally imposed* magnetic field corresponding to a uniformly bounded vector potential, is added to our picture, but the particles remain point charges regarded as having *zero* magnetic moments and regarded as not generating – but only responding to – magnetism, then let me claim (without full proof) that the methods of this paper go through with only minor alterations to show QM remains algorithmic.

12 Conclusion and open problems

12.1 Why is quantum mechanics easier than classical mechanics?

At first it seems incredible that quantum mechanics could possibly be *easier* to simulate than classical mechanics, but comparing the present paper with the earlier [83] shows that there is a sense in which that is the case. Specifically, [83] had shown that unsimulable initial configurations of N point masses existed – i.e. the masses, following Newton's laws of gravity and motion, and starting from a certain compact ball in the space of possible initial configurations, would describe one of *uncountably infinitely many* possible topologically distinct trajectories in 1 minute. But (as we now see) the equivalent quantum mechanical N-body problem is always simulable.

There are several reasons for this surprising conclusion. First, quantum mechanics is "fuzzy," i.e. after running a

[52] Electrons are often regarded as point dipoles.

quantum experiment, you only get a probabilistic sample out. This means the simulator has the easier task of only simulating a sample (§2.3). We can ignore small amplitude effects and still get an acceptably good simulation. Classically, in contrast, huge precision can be *required*. Another aspect of the same thing is that classical mechanics is generically "chaotic" (or at least, is generally conjectured to be), i.e. exponentially sensitive to infinitesimal changes in initial conditions. Quantum mechanics in contrast is incapable of chaos [39] since time-evolution is unitary and hence preserves L_2 distances.

Second, uncertainty principles, Fermi gas pressure, bounded wave function norms, and bounded kinetic energy if total energy bounded (all aspects of same thing – see §6) prevent the probability amplitudes from being large for very small interparticle separations, i.e prevent "huge speeds" and "huge energies" from being probable. Those were precisely the features of Newtonian mechanics used to get unsimulatable dynamics in [83].

Both the classical and quantum N-body problems are simulable for *regularized* potentials [83]. However, only in the quantum problem is it true that regularizing the potential (to remove singularities) has provably small effect on all answers.

A short discussion from a different point of view about how and why quantum mechanics is easier than classical is in [75] vol II, section X.14. They draw attention to the question (which has also been worked on by other people; and was also highlit in [83])

Conjecture 18 (Measure zero set evolves to singularity?) *The set of initial conditions leading to a singularity of the motion for the classical Newtonian N-body problem, is of measure zero in any ball in momentum-position phase space.*

Open problem #1: What happens to the methods of the present paper in the "classical limit" $\hbar \to 0$? In particular: will that will provide a proof of conjecture 18?[53]

12.2 How the proof was developed

Many proofs are presented as fait accompli with all trace of the processes used to find that proof, carefully hidden. But in the present case, I think it's desirable to explain some of my thinking,

1. to allow the reader to avoid some tempting "wrong turns,"
2. so I can present some interesting problems arising in the investigation of methods which ultimately were unused in the proof,
3. to elucidate some of the weaknesses in present day mathematics relevant to future investigations of Church's thesis.
4. Because the proof is so complicated, this may aid understanding it.

12.3 Initial thinking

Since quantum mechanics is governed by partial differential equations (PDEs) the approach initially seems obvious – use numerical methods for approximately solving PDEs! However, in general this approach does not work because essentially *nothing* is currently known about rigorous error bounds on numerical solutions of PDEs. For example, consider the Einstein-Maxwell PDEs governing General Relativity, or the Navier-Stokes equations of fluid mechanics. In neither case is existence and uniqueness of a solution known. The best that has been proven is that, given certain reasonable constraints on the initial state, a solution will exist for some nonzero amount of time. If we cannot even prove existence, then we cannot hope to get rigorous algorithmic results.

12.4 The proofs

Fortunately, the PDEs of quantum mechanics are nicer than the Navier-Stokes and General Relativity PDEs in at least three ways:

1. Quantum mechanics is linear.
2. Quantum mechanics features uncertainty principles.
3. A solution may be explicitly written down, if one is willing to use Feynman "path integrals."

Idea 1 allows the use of the heavy machinery of linear operator analysis. Using this machinery, Kato [47] was able to show existence and uniqueness for Schrödinger's equation with L^2+L^∞ interparticle potential functions in $\mathbf{R}^3$. Also[54], the *unitary* nature of time evolution, and the fact

[53] Here is intuition about my idea of trying to use the classical limit of QM to prove the open conjecture that a 0-measure set evolves to signularity for the classical N-body problem. Feynman (path integrals) has a mental picture that in the classical limit $\hbar \to 0$, all paths cancel out except for those "near" the classical paths. I do not know how to make that rigorous. But, suppose it could be managed somehow, even in some very weak limiting/asymptotic bounding sense. Now if a nonzero fraction of classical initial states became singular (acquired infinite kinetic energy E_K) then maybe one could show the quantum expectation value of E_K approached infinity in some limit involving $\hbar \to 0$ and time increasing (I am being vague about precisely what limit that is). But via Lieb's uncertainty principle we can (and have) shown E_K cannot ever get very large in QM (for any fixed $\hbar > 0$). This hopefully will yield a contradiction.

[54] The reader may wonder how it can be that Kato's condition ($L^2 + L^\infty$) can nowhere say anything about derivatives. After all, the Schrödinger equation explicitly involves derivatives, so surely the correct setting in which to study it, is, e.g., a Sobolev space – in which explicit norm constraints are imposed on derivatives? Without such constraints, the Schrödinger equation is surely inapplicable? However, that view is wrong.

Suppose the Schrödinger operator has a complete orthonormal set of eigenfunctions on some space "S." Then QM is defined for *any* wavefunction in S whose expansion as a linear combination of the eigenfunctions, converges absolutely in norm(S). Because, there is a natural, unique way to time-evolve it (namely to multiply each eigenfunction by $e^{itE/\hbar}$ where E is its energy), and that way will take over the role of "the Schrödinger equation." In this view there is nothing special about the particular formulation of the that (for the purposes of Church's thesis) quite large errors in the simulated wavefunction are permissible (§2.3), and Lieb's reformulation [63] of the uncertainty principle as a powerful tool (§6), are all extremely helpful in getting error bounds. The most direct outgrowth of these facts is the Rayleigh-Ritz approach sketched in §D.

Idea 3 initially seems insane because path integrals seem to be extremely poor choices for any algorithmic purpose. It turns out, however, that they are naturally implemented on a "quantum computer" (§8). This insight caught my imagination because it leads to a quantum computer version of Church's strong thesis.

The key question then became: how to get rigorous error bounds for approximate path integration algorithms? The rigorous convergence proof for Feynman path integrals, by viewing them as an instance of the "Trotter product formula" [71] seemed the obvious place to start looking for the answer, but it alone does not suffice to obtain any sort of error bound. Eventually I realized that combining Suzuki's operator identities (EQ 16,17) Lieb's uncertainty principle inequalities (§6), probability flux bounds (§6.1), and a strategy involving optimally-sliding cutoffs on potential singularities (§6.1), would in total generate enough power to get the job done.

At that point (1998) I wrote this paper and it duly appeared as NECI technical report dated 26 Feb. 1999 and available via my web site. It then was submitted to J.Math'l Physics, which ignored it for several years. During those years, I received various comments from readers, and also I thought of several ways to extend my results. In 2002 I revised the paper to at least mention these in §12.5 (although without making any attempt to use them at their maximum strength). The main improvement is to use more powerful formulations of the uncertainty principle than the one (E.Lieb's) I was using. These should allow producing a stronger version of the algorithmicity theorem, in which potential singularities as strong as $-r^{-1.99}$ are permitted (rather than merely $-r^{-1.19}$), and/or in which my assumption that the initial expectation-value of energy, $\overline{E}$, is bounded, may be weakened to merely assuming that $\overline{E^p}$ is bounded, for any fixed $p > 0$. Appendix E argues that the ultimate goal probably should be to try to weaken this still further, to merely assuming $\ln \overline{\exp(cE)}$ is bounded, for some constant $c > 0$.

Schrödinger equation as a PDE in position-time space (indeed perhaps that formulation should even be deprecated) and it is almost irrelevant whether the wavefunction is differentiable. Who cares if there is a "Schrödinger equation" that "looks like it usually looks?" The only thing that matters is, something plays its role as a time-evolver. In that case there is something we can call "quantum mechanics," and it features unique time evolution, which exists. It was precisely in this sense that T.Kato proved the existence and uniqueness of solutions in quantum mechanics – he showed that with interparticle potentials in $L^2 + L^\infty$, the Schrödinger operator had a complete set of orthonormal eigenfunctions. I view Kato as a great hero in this subject, though sadly a hero whom most physicists have never heard of.

12.5 *Our assumptions about energy and potential singularities can be weakened by using more powerful uncertainty principles*

The present paper contents itself with proving a (quantum) polynomial time bound for QM-simulation for the most basic variety of quantum mechanics. I wrote it in 1998. We now discuss several ways to go further, invented during 1998-2002.

There are two slight, but annoying, weaknesses about our main result (§11).

First, to get algorithmicity, we had to assume that the initial conditions were such that the expectation value of energy was[55] finite (and indeed the QTM runtime bounds are expressed in terms of $\overline{E}$). This was subject to the criticism that probability densities exist which do not have expectation values – e.g. the density proportional to $1/(1 + |x|^{1.1})$.

Second, we had to assume that the interparticle potentials had attractive singularities behaving, at worst, like $-r^{1.19}$ (more precisely, any exponent below $6/5 = 1.2$ will do). Nut the methods of proof permitted *repulsive* singularities as bad as $r^{1.99}$ (any exponent below 2 will do). Furthermore, the existence and uniqueness proof by Kato [47] had allowed *both* attractive and repulsive singularities behaving like $\pm r^{1.99}$ (since these are in $L^1 + L^\infty$), while there is nonexistence in the presence of $-r^{-2}$ attractive potentials. So it seemed plausible[56] that the limit 1.2 was merely an artifact due to a weakness in the proof technique; really it ought to be increasable to 2, which would be tight.

Revisiting this (in 2002), it now appears that both of these annoyances can be obliterated. The key is to employ more powerful forms of the "uncertainty principle." The original 1998 version had realized that the usual (product of momentum and position variances) form of the uncertainty principle was not powerful enough, hence instead employed Lieb's form (§6). But there are still more powerful forms! We now discuss them and their consequences.

[55] If it was finite at the beginning of the universe, then it still is. However, was it? Most physicists are perfectly happy with the $\overline{E}$-finiteness demand (and indeed see Appendix E for the suggestion that it might be legitimate, or interesting, to *strengthen* our $\overline{E}$-boundedness assumption to require that, further, $\ln \overline{\exp(cE)}$ be bounded, for some constant $c > 0$), but mathematicians might be interested in the true mathematical limitations on the algorithmicity of QM. It may be of interest that the ground state of the hydrogen atom (EQ 25) is essentially of form $\Psi(\vec{x}) = \exp(-r)$, whose 3D Fourier transform is $\tilde{\Psi}(\vec{w}) = 8\pi(w^2 + 1)^{-2}$. (This may be shown by using the Hankel transform theorem, theorem 1.12 p.19 of [16].) Note this behaves (for large w) like w^{-4}. Therefore if we take the square of the fourier transform multiplied by w^{2p} to get the expectation of (kinetic energy)^p, then we get an integrand that acts near infinity like w^{2p-8}, so that this integral diverges when $p \geq 2.5$. Also, the expectation of C^E is also divergent, for any $C > 1$.

[56] Again, most physicists do not care about this issue, since all the particle-potentials so far proposed by physicists have singularities behaving no worse than $\pm r^{1.01}$. But again, QM seems sufficiently important that it is worth trying to pinpoint its ultimate *mathematical* limitations.

First: The following theorem was discovered in 1983-1984 by 5 different authors [9] [41] [46] [68]. (I have attempted to compress their results into one unified statement.)

Theorem 19 (Ultimate-strength uncertainty principle?) *Let $1 \le p \le q < \infty$ and let $(1/p) + (1/p') = 1$ so $1 < p' \le \infty$. Let $a(r)$ and $b(r)$ be fixed nonnegative-valued functions of $r \ge 0$ with $a(r)$ non-decreasing and $b(r)$ non-increasing. The following inequality holds for all suitable[57] complex-valued functions $f(\vec{x})$ defined on $\mathbf{R}^d$, such that f and $\tilde{f}$ are Fourier transform pairs*

$$\left[\int_{\mathbf{R}^d} |f(\vec{x})|^q b(|\vec{x}|) \mathrm{d}\vec{x}\right]^{\frac{1}{q}} \le C_{pqabd} \cdot \left[\int_{\mathbf{R}^d} |\tilde{f}(\vec{y})|^p a(|\vec{y}|) \mathrm{d}\vec{y}\right]^{\frac{1}{p}}, \tag{76}$$

(for some constant $C > 0$ depending on p, q, a, b, d but not on f, $\tilde{f}$) if and only if the weight functions a and b satisfy

$$\sup_{s>0} \left[\int_0^{1/s} b(t^{1/d}) \mathrm{d}t\right]^{1/q} \left[\int_0^s a(t^{1/d})^{1-p'} \mathrm{d}t\right]^{1/p'} < \infty. \tag{77}$$

The freedom to choose the weight functions a, b gives the user of theorem 19 tremendous control and makes it a very powerful tool. The specialization of theorem 19 to the case of power-law weight functions is

Corollary 20 (A power-law uncertainty principle) *Let $1 \le p \le q < \infty$, $b \le 0 \le a$, and let $(1/p) + (1/p') = 1$ so $1 < p' \le \infty$.*

The following inequality holds for all suitable complex-valued functions $f(\vec{x})$ defined on $\mathbf{R}^d$, such that f and $\tilde{f}$ are Fourier transform pairs

$$\left[\int_{\mathbf{R}^d} |f(\vec{x})|^q |\vec{x}|^b \mathrm{d}\vec{x}\right]^{1/q} \le C_{pqabd} \cdot \left[\int_{\mathbf{R}^d} |\tilde{f}(\vec{y})|^p |\vec{y}|^a \mathrm{d}\vec{y}\right]^{1/p}, \tag{78}$$

(for some constant $C > 0$ depending on p, q, a, b, d, but not on f, $\tilde{f}$) if and only if a and b satisfy

$$[(1 - p')a + d]q = [b + d]p'. \tag{79}$$

Lieb's uncertainty principle (§6) arises from the special case $d = 3$, $p' = p = a = 2$, $q = 6$, $b = 0$ of corollary 20 and provides a lower bound on the expected kinetic energy, $\overline{E_K}$, of a particle. If we still demand $p' = p = 2$ and $b = 0$ but instead now allow *any* a with $0 < a < d$ [then $q = 2d/(d-a)$], then we instead get a more powerful thing – a lower bound on $\overline{(E_K)^{a/2}}$:

Theorem 21 (More general uncertainty principle) *Let d be the dimensionality of space. Let p be a constant with $0 < 2p < d$. Then the expectation of the (kinetic energy)^p of a particle with wavefunction $\Psi(\vec{x})$ obeys*

$$\mathrm{Expect}\big((E_k)^p\big) \ge \mathrm{const}_{p,d} \left(\int \rho(\vec{x})^q \mathrm{d}\vec{x}\right)^{1/q} \tag{80}$$

where $\rho = |\Psi|^2$ and $q = 2d/(d - 2p)$.

Consequence: Theorem 21 should enable improving all the results of the present paper to hold under merely the assumption that $\overline{(E_K)^p}$ is bounded (and the QTM algorithm run times should be expressible in terms of this bound), for any fixed $p > 0$ (e.g. $p = 0.001$) instead of $p = 1$.

Second: Lieb himself, in [64] (his eq. 17a and surrounding several pages of discussion), realized that his multi-particle uncertainty principle (our EQ 35) may be strengthened to a form in which the right hand side is proportional to $(\int \rho^3)^{1/3}$.

Consequence: That should enable improving all the results of the present paper to hold for potentials with attractive singularities $-r^{-\gamma}$ for any $\gamma < 2$ (see the argument near EQ 39-41), as opposed to the present restriction $\gamma < 6/5$.

Open problem #2: The restriction for *repulsive* singularities *already* was of the form $r^{-\gamma}$ for any $\gamma < 2$ – and remains unchanged. But can it be weakened? The fact that "hard balls" seem permissible [§7] makes it plausible that QM is well defined, and algorithmic, even in the presence of repulsive singularities far worse than r^{-2}.

However, the proofs of existence and uniqueness of QM by Kato and his successors have, so far, been unable to handle such singularities, so, until that is done, there is no hope to prove algorithmicity in their presence.

Finally, I also mention the appealing-looking, but apparently less useful, "uncertainty principles" that are "logarithmic Sobolev inequalities." Perhaps they will have some future use in quantum mechanics, although right now I know of none. These (extracted from papers by W.Beckner [7] [8]) are obeyed by f with $\|f\|_2 = 1$:

$$\int_{\mathbf{R}^n} |f(\vec{x})|^2 \ln |f(\vec{x})| \mathrm{d}\vec{x} \le \frac{n}{4} \ln \left[\frac{2}{\pi e n} \int_{\mathbf{R}^n} |\vec{\nabla} f|^2 \mathrm{d}\vec{x}\right] \tag{81}$$

and

$$\int_{\mathbf{R}^n} |f(\vec{x})|^2 \ln |f(\vec{x})| \mathrm{d}\vec{x} + \int_{\mathbf{R}^n} |\tilde{f}(\vec{\omega})|^2 \ln |\tilde{f}(\vec{\omega})| \mathrm{d}\vec{\omega} \le \frac{n}{2} \ln \frac{2}{e}. \tag{82}$$

12.6 More open problems

[58]

Open problem #3: Higher order product formulas (§4.1) may be useful in getting algorithmic speedups (as opposed to the simplest – 2-term and 3-term – product formulas, which were all we employed). Work that out. More generally, try to prove better algorithm performance bounds and make the algorithms more practical.

[57] That is, all f such that both f and $\tilde{f}$ may be defined and such that all the integrals that need to converge, do.

[58] Open problems #1 and #2 were in §12.1 and §12.5 respectively.

Stronger assumptions about Nature (§E, §B, §9.2, §9.3) may also be useful for getting speedups.

Open problem #4: Alternatively it might be possible to preserve our results under weaker assumptions. (See also §12.5.) For example, our assumption of a polynomial bound on the size of the box containing our quantum system probably may be weakened.

Our assumption that the potential is bounded above $-Cr_{\min}^{-\gamma}$ could probably be replaced with a weaker looking *norm* condition, such as in Kato's theorem [47] where it sufficed to have "$L^2 + L^\infty$ interparticle potentials." (See also [75] theorem X.19.)

Open problem #5: The techniques used by Lieb and collaborators [63] to extend their proofs of the "stability of matter" to more kinds of quantum mechanics, may also be useful in the quest for generalizing our algorithmic techniques to work in more kinds of quantum mechanics, e.g. allowing vector potentials (magnetic fields; Pauli equation instead of Schrödinger) and/or caricatures of relativity. I believe this should be relatively straightforward, but I have not done it.

Open problem #6: Another important question, indeed perhaps *the* important question as far as the realizability of quantum computers are concerned, is the efficiency of algorithms for simulating *decohering* quantum mechanics [30], e.g. with "master equations" [101][94] instead of Schrödinger's. Although such equations certainly look *more* complicated, I would presume (or guess) that they are actually *easier* to simulate [85].

Open problem #7: Finally, there is the matter of trying to determine the algorithmicity status of more advanced quantum theories such as QED (quantum electrodynamics) EWT (electroweak theory), and QCD (quantum chromodynamics), together comprising the "standard model" thought to describe all nongravitational physics. Unfortunately rigorous mathematical underpinnings (analogues of Kato's work on ordinary QM) are presently nonexistent for these theories, so a necessary first step would be to create them.

A Appendix on notation used in this paper

Our uses of c, α, m_e, e, h, $\hbar$, G, ϵ_0, μ_e, k_B, λ_C to denote certain physical constants follows [19]. $i = \sqrt{-1}$.

$\Box$ denotes the end of a proof. We number all lemmas, claims, assumptions, theorems, etc in the same integer sequence: Claim 1, Lemma 2, Corollary 3, etc. Be warned I have a habit of writing some analysis and discussion and then summarizing it as a "theorem" or "lemma," as opposed to the more formal method of stating the theorem then proving it in a clearly delineated "proof."

"Effective" (as in "effective bound") means that an algorithm is available for computing that bound. An "algorithm" is a computer program guaranteed to terminate. We have on occasion used Ko's [58] notions of "polynomial time computable real numbers" and "polynomial time computable functions $f : \mathbf{R} \to \mathbf{R}$." Roughly speaking (for more precise definitions see chapter 2 of [58]):

1. A real number x is "computable" if there exists an algorithm which, given a positive integer n as input, will output a rational approximation x_n to x such that $|x - x_n| < 2^{-n}$.
2. If the algorithm runs in time at most polynomial(n), then x is "polynomial time computable."
3. If an algorithm exists which, given n and "black box access" to the algorithm for computing some x, will compute f_n with $|f_n - f(x)| < 2^{-n}$, and will do so in at most polynomial(n) steps (where each black-box query takes 1 step plus 1 extra time unit for each bit transmitted or received, and where the black box by convention accepts its input in *unary* form) then $f(x) : [0, 1] \to \mathbf{R}$ is "polynomial time computable on $[0, 1]$."

For algorithms for arithmetic and for approximating functions like $\sin(x)$, see [14] [56].

All time bounds for algorithms are assuming worst case input.

The word "polynomial," when used in discussions of computational complexity, means a function bounded by An^B where A and B are positive *real* (not necessarily integer) constants and n is the number of bits in the input. "Quasi-polynomial" means a function bounded by $An^{B(\log n)^C}$ where A, B, and C are positive real constants. The word "exponential" in the same context will mean $\exp(An^B)$. When used in discussions of physical computers, polynomial and exponential will mean $An_1^{B_1} n_2^{B_2} ... n_k^{B_k}$ and the exp of this, respectively, where the A_k and B_k are positive real constants and the n_k include (besides the number of bits in the input) other physically relevant parameters such as the number of particles, the size of the enclosing box, and the total energy. These words are very convenient because by using them we can (purposely) obscure a vast number of difficult details irrelevant to the question of whether something is polynomial or exponential (such as exactly what polynomial it is – hence the resulting theorem statements can look vague to the uninitiated).

See [44] (and the end of our §10) for discussion of complexity classes such as "#P."

Stars as superscripts, i.e. x^*, denote the complex conjugate if x is scalar, and the complex conjugate transpose (Hermitian adjoint) if x is a matrix or operator.

The "L_p norm," $p > 0$, of a wavefunction $\Psi(\vec{x})$ is $\|\Psi\|_p = (\int |\Psi|^p \mathrm{d}\vec{x})^{1/p}$. If we say $\|\Psi\|$, that means $\|\Psi\|_2$. The "L_∞ norm" of a function is the supremum of its absolute value. The set of functions with finite L_p norm is called "L^p." If $p < q$ then $L^p \subset L^q$. A function is in "$L^2 + L^\infty$" if it may be expressed as the sum of a two functions, one in L^2 and the other in L^∞. $L^1 \subset L^2 + L^\infty$. A function $F(\vec{x}) : \mathbf{R}^3 \to \mathbf{R}$ is "Rollnick" if

$$\int\int \frac{F(\vec{x})\, F(\vec{y})}{|\vec{x} - \vec{y}|^2} \mathrm{d}^3\vec{x}\mathrm{d}^3\vec{y} < \infty. \qquad (83)$$

$L^{3/2}(\mathbf{R}^3) \subset$Rollnick. The L_2 norm of an *operator* M is the supremum of $\|M\Psi\|$ over all Ψ with unit L_2 norm.

"Hölder's inequality" states that if p and q are positive reals with $p^{-1} + q^{-1} = 1$, then

$$\|fg\|_2 \leq \|f\|_{2p}\|g\|_{2q}. \tag{84}$$

Our definitions of O, Ω, θ asymptotic bounding notation follow the recommendations of Knuth [57], nowadays standard in computer science.

We use the same quantum computational model, and the same set of "quantum gates," as are employed by Shor [81]; see that paper for definitions of these notions and related terminology such as "qubit." For a good understanding of what "instructions" are available in a "quantum programming language," see [72]. For example, "no cloning" and unitarity theorems prevent erasure of information, which is commonplace in programming languages for conventional computers. However, this can be gotten around in various ways, for example making variables in the language be "write once" and realizing, e.g., that the time evolution $|a\rangle \to |aa\rangle$ is ok, whereas $|a\rangle \to |a\rangle|a\rangle$ is forbidden (cloning). The "no-cloning theorem" [99] originates from the observation that

$$f|00\rangle + g|11\rangle \neq (f|0\rangle + g|1\rangle)(f|0\rangle + g|1\rangle) = f|00\rangle + g|11\rangle + fg|01\rangle + gf|10\rangle; \tag{85}$$

the left hand side is a state that can be manufactured from $f|0\rangle + g|1\rangle$, but the right hand "cloned" state cannot.

B Appendix: Are intercharge potentials Coulombic?

We claim that interparticle potentials in quantum mechanics should not be modeled as Coulombic. A better model involves a potential free of attractive singularities. (This is undoubtably the case for particles with nonzero sizes such as the proton[59], but we are going to argue this even for particles such as electrons and positrons, which appear to be points[60].)

Really, of course, one should be using deeper theories such as QED, in which "potentials" between particles really are regarded as an obsolete notion (cf. footnote 3). But we have agreed in this paper to restrict ourselves to the garden variety quantum mechanics of the Schrödinger equation with potentials. In that case, it is far from clear a priori what interparticle potential is the best one to use.

Professional simulators of atoms [45][26] generally proceed by solving Schrödinger's or Dirac's equation using exact Coulomb potentials and then adding various perturbative "correction terms." The same result could have been achieved without correcting anything by using a non-Coulomb potential designed to have essentially the same effect as the correction terms. So it is clear that better answers would be obtained, when solving the Schrödinger equation, if non-Coulombic potentials had been used – but which?

B.1 Heuristic semiclassical analysis

In a heuristic attempt to answer this question[61] we'll now demonstrate how to find a potential $U(r)$ which causes the Bohr old-style (pre-Schrödinger) atomic model to give spectral energies $E(n)$ for essentially any desired smooth function $E(n)$.

Bohr[62] had envisioned an electron moving in a classical circular orbit with velocity v and radius r and hence $mv^2/r = U'(r)$ to balance centrifugal and attractive forces. Hence $v = \sqrt{rU'/m}$. He also imposed the de Broglie quantization condition $n\hbar = rmv$. Hence

$$n = r\sqrt{rmU'}/\hbar. \tag{86}$$

The potential plus kinetic energy is

$$E(n) = U + rU'/2. \tag{87}$$

If we now combine the previous two equations and regard n as an arbitrary *real*[63] rather than restricting it to positive integer values, then

$$E(\frac{r\sqrt{rmU'}}{\hbar}) = U + \frac{r}{2}U' \tag{88}$$

which is a differential equation to solve for $U(r)$.

The differential equation (EQ 88) may be solved exactly if $E(n) = an^p$; the result is $U(r) = br^q$ where $p = 2q/(2+q)$, $q = 2p/(2-p)$, and

$$a = \frac{(2+q)b}{2}(\frac{\hbar}{m^2bq})^{q/(q+2)}. \tag{89}$$

In particular if

$$E(n) = \frac{mc^2\alpha^2Z^2}{2n^2} \tag{90}$$

(the Bohr energies) we get $U(r) = -Z\alpha\hbar c/r$ (the Coulomb potential). If $E(n) = an + E_0$ (linear energies) we get $U(r) = E_0 + mr^2a^2/(2\hbar^2)$ (a harmonic oscillator potential).

[59] Nuclei with atomic number Z have radii $\approx Z^{1/3} \times 10^{-15}$ meters.

[60] Electrons in QED are experimentally known to have radius $< 10^{-18}$ meters [35].

[61] There seems to be no alternative to the kind of approach I proffer. For example, if you tried to ask "what is the potential energy in QED of two electrons held at distance r apart?" then you would discover that forcing the two electrons to be *exactly* r apart takes infinite energy due to the $\Delta x \Delta p$ uncertainty principle. The question is ill-posed. One could try to define "time-averaged" QED potentials, but these would give very wrong answers in real problems because the true fields aren't static.

[62] Despite its simplicity, Bohr's model [13] leads to exactly the same energy levels (EQ 90) as the exact solution of Schrödinger's equation.

[63] Such an assumption is necessary to force us to get a unique "natural" $U(r)$ in the end. Otherwise, we could just imagine a potential with "grooves" cut at various places, but otherwise essentially unconstrained.

For a meatier example, consider the known formula [24][34]

$$E(n) = \frac{m_e c^2}{\sqrt{1 + [\frac{\alpha Z}{n - |k| + \sqrt{k^2 - (\alpha Z)^2}}]^2}}, \qquad n, k \text{ integer}, \quad |k| \le \lceil n-1 \rceil. \tag{91}$$

for the exact eigenvalues of Dirac's relativistic equation for an electron in a Coulomb potential with central charge Ze. One may now determine the non-Coulomb potential which would have produced the same results in Bohr's nonrelativistic model.

In (EQ 91) (with any k with $|k| \le \lceil n-1 \rceil$, it does not make much difference which) and $|\alpha Z| < 1$, $E(n)$ is well approximated by the Bohr energies (plus[64] mc^2) for $n \to \infty$, but as $n \to 0$ instead is approximated by a linear expression $an + b$ where

$$a = \frac{mc^2 \alpha Z}{(\sqrt{k^2 - \alpha^2 Z^2} - k)^3 Q^3}, \qquad b = \frac{mc^2}{Q}, \tag{92}$$
$$Q = \sqrt{1 + \frac{\alpha Z}{(\sqrt{k^2 - \alpha^2 Z^2} - k)^2}}.$$

Hence, $U(r)$ as $r \to 0$ approximates a Harmonic oscillator potential. Thus,

Claim 22 (Potentials "more Coulombic than Coulomb") *In order to cause nonrelativistic quantum mechanics to approximate the true relativistic behavior (i.e. Dirac energies) of a Coulomb potential, one should use a potential $U(r)$ which is Coulombic for large r, but for small r is a "Harmonic oscillator" potential (with a certain smooth join, determined by (EQ 88,91), between).*

Note, this is exactly the same kind of behavior as we get by replacing each point charge by a uniform ball of charge as in (EQ 49). The characteristic r length scale at which the behavior changes over is $\approx (\alpha Z)^2$ times the "Bohr radius" r_0 of the ground state, i.e. for hydrogen ($Z = 1$, $r_0 \approx 5.29 \times 10^{-11}$ meter) at about $r \approx 2.82 \times 10^{-15}$ meter. Amazingly enough, this is exactly the "classical electron radius," arising in classical "Thomson scattering," when considering the electron's classical self-energy, etc. (Of course the proton has a RMS radius $r_p \approx 0.8 \times 10^{-15}$ meter so its potential is non-Coulombic below that, but we are considering mystical point protons here for the sake of argument.) The Coulomb singularity has vanished![65]

[64]We will write m instead of m_e in the rest of this section, to make typography easier.

[65] If $Z\alpha \ge 1$ this does not work. However, in that case Dirac's equation is ill posed and its "ground state" has energy $-\infty$. So we will assume $Z \le 137$ in this paper, which since $\alpha \approx 1/137.036$ seems to be the maximum possible number of charge quanta allowable on a point particle in Dirac's relativistic quantum mechanics. Our later discussion makes it plausible that the "Lamb shift" (arising from QED of a higher order than the Dirac equation) would prevent collapse of an electron onto a point nucleus even when $Z \ge 138$, because the appropriate nonCoulombic potential allows a tiny (but nonzero) radius ground state far below Dirac's. However for real nuclei this effect (if it exists) would be irrelevant, because this tiny radius would be considerably smaller than the nuclear radius, so collapse undoubtably *would* occur. Furthermore, when Z exceeds ≈ 170, it becomes energetically feasible for an electron-positron pair to be created, with the electron being sucked into the nucleus and the positron being emitted. These effects should prevent any such highly charged nuclei from existing in real life.

B.1.1 *Lamb shift*

The next most important effect on atomic spectra (beyond the "fine structure," which is handled by Dirac's equation) is the "Lamb shift," which is caused by QED effects unknown to Dirac's equation[66]. The Lamb shift causes the Dirac eigenvalue (at fixed k) to increase [18] [26] by an additive amount

$$\approx \frac{mc^2 Z^4 \alpha^5}{n^3} F(Z\alpha) \tag{93}$$

where $F(Z\alpha)$ is dimensionless, approximately constant for $1 \le Z \le 92$, and *independent of* n. An expression of this form gives a very good approximation to the experimentally observed discrepancies between the $2s_{1/2}$ and $2p_{1/2}$ levels of all hydrogenic ions from hydrogen to uranium [26].

Notice that as $n \to 0$ (or $Z \to \infty$ or $\hbar, 1/\alpha \to 0$), the (positive huge) Lamb shift eventually becomes far larger than the (negative huge) Bohr energy value. This leads (via (EQ 88) and its power law solution) to $U(r)$ behaving proportionally to $r^{-6/5}$ as $r \to 0$. (The characteristic length scale at which the Coulomb potential stops due to this effect is $\approx Z^4\alpha^6$ times the radius of the Bohr ground state, i.e. 8.0×10^{-24} meters for hydrogen.) The singularity has (at least formally) reappeared, but it now is repulsive instead of attractive! Of course, the semiclassical reasoning underlying the Bohr model loses its attraction if the potential is repulsive, so the existence of a repulsive singularity is subject to question. But regardless of that it seems clear that the attractive singularity is gone.

It is believed that at length scales below the "Planck length" $\ell_{\text{Planck}} = \sqrt{\hbar G/c^3} \approx 1.62 \times 10^{-35}$ meter, gravitational effects will become important and will completely alter the picture in some unknown manner. Therefore, arguably, it is pointless to do quantum mechanics below ℓ_{Planck} and so one might as well assume that the Coulomb potential is cut off there. In that case, there is no singularity at all, but the potential reaches a *large* constant value before it stops![67]

[66]For $2 \le Z \le 92$, i.e. the naturally occurring elements, the Lamb shift has a much larger effect on spectral energies than the fact that nuclear sizes are nonzero, incidentally [45].

[67]Computer scientists regard all the reciprocals of these extremely small lengths merely as "constants" which do not affect the polynomiality of simulation algorithms. This point of view is, unfortunately, less justified here than it usually is in computer science, because the constants in the present arguments are a lot larger than the ones usually encountered when analysing algorithms.

B.2 Relativistic replay

Bohr's nonrelativistic atomic model was redone special relativistically by A.Sommerfeld [87][68]. We quickly sketch the derivation analogous to the above but using Sommerfeld's model instead of Bohr's:
Force balance: $mv^2/r = U'(r)\sqrt{1-v^2/c^2}$.
Hence $v^2 = 2rU'c^2/(\sqrt{(rU')^2+4m^2c^4}+rU')$.
Quantization condition: $n\hbar = rmv/\sqrt{1-v^2/c^2}$. Hence

$$n = \frac{rmc\sqrt{2rU'}}{\hbar}\left(\sqrt{(rU')^2+4m^2c^4}+rU'\right)^{1/2} \quad (94)$$

The potential plus kinetic energy is

$$E(n) = U + mc^2/\sqrt{1-v^2/c^2} = \quad (95)$$

$$U + \sqrt{m^2c^4 + (rU')^2/2 + \sqrt{(rU')^2+4m^2c^4}\,rU'/2}$$

Combined differential equation:

$$E\left(\frac{rmc\sqrt{2rU'}}{\hbar}\left[\sqrt{(rU')^2+4m^2c^4}+rU'\right]^{1/2}\right)$$

$$= U + \frac{c^2rU'}{v^2} = \quad (96)$$

$$U + \frac{\sqrt{(rU')^2+4m^2c^4}+rU'}{2}$$

In (EQ 96), if we wish to analyse $U(r) \sim br^q$ for $r \to 0$, then there are two cases. In both cases we get $E(n) \sim an^p$ as $n \to 0^+$, but the values for p are different.

- If $q > 0$ (of course we remove an additive constant from the right hand side by redefining the zero point of our energy scale) (EQ 96) becomes

$$E\left(\frac{r\sqrt{mrU'}}{\hbar\sqrt{2}}\right) \sim U + rU'/2 \quad (97)$$

  so $p = 2q/(2+q)$, $q = 2p/(2-p)$, i.e. the same results as in the nonrelativistic case.

- If $q < 0$, (EQ 96) becomes

$$E\left(\frac{rmc}{\hbar\sqrt{2}}\right) \sim U + rU'. \quad (98)$$

  so $p = q$. This differs from the nonrelativistic case.

In particular, in the case $p = -3$ of interest for the Lamb shift, we have $q = -3$ so that to make Sommerfeld give the right (Lamb perturbed) answers we need a repulsive singularity $U(r) \propto r^{-3}$ for $r \to 0$. We conclude that

Claim 23 (relativistic replay of claim 22; Interparticle potentials are bounded below by a constant) *At least as far as relativistic corrections and/or the simpler QED approximations to the Lamb shift are concerned*[69] *in old-style Bohr atom quantum mechanics, interparticle potentials are bounded below by a constant. The same conclusion arises when considering the Lamb shift from the point of view of Sommerfeld's relativistic Bohr atom.*

B.3 Open question

An idea similar to the above, but which I have not investigated, would be to consider scattering differential *cross sections* instead of energy levels and again try to design the right potential to force the (nonrelativistic) Schrödinger equation to give the right answers, i.e. the Bhabha e^+e^- (relativistic) cross section formula [11] [43] [5] [3].

For 1-dimensional scattering, exact solution methods ("inverse scattering theory") are available to reconstruct the potential from the scattering data [27] [25]. However, I am unaware of any analogue of this theory in ≥ 2 dimensions, and indeed I conjecture the question is ill-posed, i.e. no solution potential exists, in general[70].

There is an extremely cryptic review of forward scattering theory in [26].

C Appendix on higher product formulas

We continue from where §4.1 left off by discussing what is known (including new results) about product formulas with ≥ 4 terms. These include Suzuki's recursive construction of product formulas of arbitrarily high order (§C.2) and his theorem that negative or complex coefficients are required in any formula of order ≥ 3. Also, we explain (§C.1) why these formulas are useful for numerical approximation of the time evolution of classical Hamiltonian systems.

C.0.1 4-term product formula

$$e^{(A+B)t} = e^{[1-\frac{1}{\sqrt{2}}]At}e^{\frac{Bt}{\sqrt{2}}}e^{\frac{At}{\sqrt{2}}}e^{[1-\frac{1}{\sqrt{2}}]Bt}[1+O(t^3)]. \quad (99)$$

(Note $1/\sqrt{2} \approx 0.7071067810$, $1-1/\sqrt{2} \approx 0.2928932190$.) Note, this does not yield an accuracy-order improvement versus the simpler 3-term product (EQ 14). But the errors seemed to be about 3-4$\times$ smaller in a few experiments I tried, suggesting there is a better constant term in the error formula.

[68] Despite its simplicity, Sommerfeld's model leads to exactly the same energy levels (EQ 91) as the exact solution of Dirac's equation.

[69] Conceivably this conclusion would change if the Lamb shift had been handled in very high order QED instead of second-order QED as in (EQ 93). But at third and fourth order the corrections (albeit available only in an incomplete form so far [26]) seem to have the same form as (EQ 93) (just times higher powers of α and with different F's), and hence would *not* alter this conclusion.

[70] The potential is a function of one veriable r. But in ≥ 2 dimensions we want to determine it from a function of two variables (the differential scattering cross section as a function of both angle and of energy). Hence, one would expect it, in general, to be overdetermined.

C.0.2 5-term product formula

There *does not exist* a complex number q such that

$$e^{(A+B)t} = e^{Atq}e^{Bt/2}e^{[1-2q]At}e^{Bt/2}e^{Atq}[1+O(t^4)] \quad (100)$$

so one might as well just choose $q=0$ or $q=1/2$ to get the 3-term product, since any value of q achieves $O(t^3)$ error and the 3-term product involves less work.

To aid discussion of formulae of this type, define a "p-term product formula of order k" to be (as $t \to 0$)

$$e^{(A+B)t} = \underbrace{e^{Atc_1}e^{Btd_1}e^{Atc_2}e^{Btd_2}e^{Atc_3}\dots}_{p \text{ terms}}[1+O(t^{k+1})]. \quad (101)$$

A formula is "palindromic" if the coefficients read left to right are the same numerical sequence as if they are read right to left. It is "all-real" if all the c_j and d_j are real.

Despite the nonexistence result of (EQ 100), there *is* a 5-term product of order 3 if we allow *complex* c_j and d_j. There are exactly two solutions:

$$(c_1, d_1) = (\frac{1}{4}, \frac{1}{2}) \pm Ri, \quad c_2 = \frac{1}{2}, \quad (d_2, c_3) = (\frac{1}{2}, \frac{1}{4}) \mp Ri, \quad (102)$$

where $R = \sqrt{3}/12 \approx 0.1443375673$. Note the "complex conjugate palindromicity" $c_1 = c_3^*$, $d_1 = d_2^*$. This suggests that using $q = 1/4$ in (EQ 100) (thus getting the closest all-real formula to the complex formula (EQ 102)) might be a good idea for the purpose of reducing the error constant, although not its order. Of course, this is equivalent to using the 3-term product twice (on two half-size time intervals), and thus it will indeed reduce the constant – by a factor of 4.

C.0.3 6-term product formula

There are exactly four palindromic 6-term product formulae of order 3, and of these exactly two are all-real. Let $Q \approx 0.9196615233$ be one[71] of the two real roots of

$$12x^4 - 24x^2 + 16x - 3. \quad (103)$$

Then the formula arises from

$$\begin{aligned} c_1 = d_3 &= -Q^3 + 3Q/2 - 1/3 \approx 0.2683300956, \\ d_1 = c_3 &= Q \approx 0.9196615233, \quad (104) \\ c_2 = d_2 &= Q^3 - 5Q/2 + 4/3 \approx -0.187991619. \end{aligned}$$

These are the 4 palindromic special cases of a 1-parameter family of formulae indicated by Ruth [76] [31]. (Actually, the full solution space, i.e. with no palindromicity restriction, consists of exactly two 1-parameter families. No member of either of these families achieves order 4.) The fact that c_2 is negative is the first occurrence of a bad trend. In fact, Suzuki [92] showed (his "theorem 3") that

[71] The other is $R \approx -1.695156345$, but this is less desirable; the two complex roots $0.3877474108 \pm 0.1000711207i$ are presumably even less so.

Theorem 24 (Suzuki nonexistence theorem) *There are no finite length product formulae of any order ≥ 3 in which all the coefficients are positive reals.*

C.0.4 7-term product formula

The unique all-real palindromic 7-term order-4 formula is

$$\begin{aligned} c_1 = c_4 &= \frac{R}{2} \approx 0.675603596, \\ d_1 = d_3 &= R = 1/(2-2^{1/3}) \approx 1.351207192, \\ c_2 = c_3 &= \frac{1-R}{2} \approx -0.175603596, \\ d_2 &= 1 - 2R \approx -1.702414384. \quad (105) \end{aligned}$$

This formula was known to previous workers, and in particular Suzuki [90] (eq. 30 and 31) claims to have found a 1-parameter family of 7-term order-4 formulae, of which this is the palindromic special case. However, my computer's Gröbner basis (described later) indicates that the solution space (with no palindromicity restriction, and allowing complex c_j and d_j) is finite, so Suzuki's family must have been invalid. (This has been confirmed both by `MAPLE`'s Gröbner algorithm, and also by Henry Cejtin's independently written code.) Indeed, there are exactly 5 complex solutions, namely

$$\begin{aligned} c_1 = 1/4 - Q/2, \quad d_1 = 1/2 - Q, c_2 = 1/2 - Q/2, \quad (106) \\ d_2 = 1/2, c_3 = 1/4 + Q/2, \quad d_3 = Q, c_4 = Q/2 \end{aligned}$$

where $Q = 1/4 \pm i\sqrt{15}/12$, and

$$\begin{aligned} c_1 = c_4 &= 1/4 - C/4, \\ d_1 = d_3 &= 1/2 - C/2, \\ c_2 = c_3 &= 1/4 + C/4, \\ d_2 &= C \quad (107) \end{aligned}$$

where $3C^3 + 3C^2 - 3C + 1 = 0$, i.e. $C \approx 0.3512071921 \pm 0.2691725452i$ and $C \approx -1.702414384$. These last 3 solutions are each palindromic. None of these 5 solutions achieves order 5. The last of these 5 solutions is the only all-real solution and gives (EQ 105). Again (the second example of the bad trend) the middle three terms in (EQ 105) involve negative coefficients.

C.0.5 (8-13)-term products

No palindromic product with ≤ 13 terms exists with order 5. Also, no products with either 8 or 9 terms (i.e. with no palindromicity restriction) exist of order 5. Surprisingly, in view of the 7-term order 4 product (EQ 105), there does not even exist a palindromic 8-term product of order 4, although of course non-palindromic ones exist.

C.0.6 How I proved all this

Verifying all these formulae is straightforward but tedious; one simply expands the exponentials as Taylor series out to whatever order one needs, keeping in mind the noncommutative nature of multiplication among A and B.

For example, to derive (and prove) the 4-term product formula, which I will write as

$$e^{(A+B)t} = e^{Atc_1}e^{Btd_1}e^{Atc_2}e^{Btd_2}[1+O(t^3)] \quad (108)$$

we need to check agreement on terms of degree 1 (i.e. we must have $c_1 + c_2 = 1$ so that $A = A$ and $d_1 + d_2 = 1$ so that $B = B$) and on terms of degree 2. The degree-2 terms are $AB/2!$, $BA/2!$, $AA/2!$ and $BB/2!$. These respectively correspond to

$$c_1d_1 + c_1d_2 + c_2d_2 = 1/2, \quad (109)$$
$$d_1c_2 = 1/2, \quad (110)$$
$$c_1c_2 + c_1^2/2 + c_2^2/2 = 1/2, \quad (111)$$
$$d_1d_2 + d_1^2/2 + d_2^2/2 = 1/2. \quad (112)$$

These 6 equations have a 1-parameter family of solutions

$$c_1 = 1-x, \quad d_1 = \frac{1}{2x}, \quad c_2 = x, \quad d_2 = 1 - \frac{1}{2x} \quad (113)$$

and if we also impose the palindromic conditions $c_1 = d_2$, $c_2 = d_1$, then there are exactly two solutions

$$c_1 = d_2 = 1 \mp \frac{1}{\sqrt{2}}, \quad c_2 = d_1 = \pm\frac{1}{\sqrt{2}}. \quad (114)$$

Of course, we prefer the upper signs so that all the coefficients are positive.

More generally, checking that the error term is $O(t^{k+1})$ requires verifying $2^{k+1} - 2$ equations.

The 3-7 term product formulae above (and the nonexistence proofs for the {5,8,9,10,11,12}-term products) were found by a systematic computer search. All of the error terms are tight, as was also proved by computer. The method was simply to generate all the equations involved in verifying an error term of order t^k, and then find a "Gröbner basis" [4] [22] for them. For example, the 6 polynomials arising in the previously described search for the 4-term order-3 product, have Gröbner basis

$$c_1 - d_2, \quad c_2 + d_2 - 1, \quad d_1 + d_2 - 1, \quad 1 - 4d_2 + 2d_2^2. \quad (115)$$

Once this has been done, there is a trivial algorithm to test solvability of the polynomial system – specifically: if the Gröbner basis consists of the single element $\{1\}$, then the system is not solvable! A slightly more complicated algorithm will test finiteness of the solution space, and (assuming solubility and finiteness) will actually find all solutions. Specifically, Gröbner bases generate the same polynomial ideal as the original set of polynomials but are "upper triangular" in the sense that one may solve the last equation (which is univariate), substitute it into the other equations, and then solve the second to last (which is then univariate), substitute, et cetera, to solve everything. In the present case, we solve the last equation

$$1 - 4d_2 + 2d_2^2 = 0, \quad (116)$$

finding $d_2 = 1 \pm 1/\sqrt{2}$, and then the other equations solve instantly. The commercial symbolic manipulation system `MAPLE` works, although a Gröbner package written by Henry Cejtin runs faster[72].

C.1 The connection to "symplectic integrators"

After this computer search, I found out that in the case of the 3, 6, and 7-term products, my computer had merely rediscovered previous work (cited above). Its other results seem new.

The previous authors were investigating these formulae for a different reason: they lead to "symplectic numerical integration" schemes for "classical Hamiltonian systems" of differential equations. Here is the explanation of that (patterned after [100]).

A "Hamiltonian system" is a system of ordinary differential equations of the form

$$\dot{\vec{q}} = \frac{\partial H}{\partial \vec{p}}, \quad \dot{\vec{p}} = -\frac{\partial H}{\partial \vec{q}}. \quad (117)$$

where $H(\vec{q}, \vec{p})$ (called the "Hamiltonian") is a scalar valued function and $\vec{q}(t)$ and $\vec{p}(t)$ are n-dimensional vectors. Letting $\vec{z}$ be the $2n$-vector whose elements are $\vec{q}(t)$ and $\vec{p}(t)$, (EQ 117) may be rewritten, with the aid [59] of the "Poisson bracket" notation

$$\{F, G\} = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}\frac{\partial G}{\partial p_j} - \frac{\partial F}{\partial p_j}\frac{\partial G}{\partial q_j} \quad (118)$$

as

$$\dot{\vec{z}} = \{\vec{z}, H(\vec{z})\}. \quad (119)$$

If any quantity I obeys $\{I, H\} = 0$ then I is an invariant (preserved by time evolution)[73]. If we introduce the differential operator "D_G" by $D_G F = \{F, G\}$ then (EQ 119) may be further rewritten

$$\dot{\vec{z}} = D_H \vec{z}. \quad (120)$$

It now becomes apparent that the solution is

$$\vec{z}(t) = \exp(tD_H)\vec{z}(0) \quad (121)$$

[72] All results except two were obtained by both packages. These two were (i) the nonexistence proof for the 9-term order 5 product without palindromicity restriction, which was obtained by Cejtin's software only, after running 4.9 CPU days. (Arithmetic on integers 15000 digits long was required during the Gröbnerization.) Also (ii) the nonexistence proof for the 12-term order 5 palindromic product was shown by Cejtin's code after 98 CPU days, and involved integers 55000 digits long and the consumption of several hundred megabytes of RAM. Cejtin and I have reason to believe that this calculation would have been speeded up by a factor of 200 by use of faster bignum arithmetic subroutines such as the gnu bignum package (which employs Karatsuba's algorithm [56]). Additional large speedup factors beyond this (I would guess a factor of 10) should be obtainable with bignum algorithms of Schönhage-Strassen type [56]. Further large speedups might be obtainable by modular methods to avoid bignums entirely (perhaps especially for nonexistence proofs).

[73] A theorem of Poisson is that if I and J are invariants, so is $\{I, J\}$.

and if D_H is a sum of known operators (which happens if H is a function of $\vec{q}$ plus a function of $\vec{p}$) whose Hamiltonian flows are known, then we may apply our splitting formulae to get approximate solutions. All of these solutions will be exactly "symplectic." That is, the time evolution operator $\exp(tD_H)$, which is the exponential of an antisymmetric real operator (or more generally: the exponential of i times a Hermitian operator) i.e. *unitary*, will be being approximated by an exponential of another exactly unitary operator.

C.2 Suzuki's recursive construction of splitting formulas of every even order

Suzuki [90] [89] showed how to generate product formulae of arbitrarily high order. We won't discuss Suzuki's maximally general construction (in [91]). Specifically let a formula be called "of kth order" if the error term is $O(t^{k+1})$. Then let

$$S_2(t) \;=\; \exp(\frac{At}{2})\exp(Bt)\exp(\frac{At}{2}) \qquad (122)$$

be the order-2 approximant (3-term product) above. Then a recursive construction of formulae of all even orders $2m$ arises via

$$S_{2m}(t) \;=\; S_{2m-2}(R_m t)S_{2m-2}([1-2R_m]t)S_{2m-2}(R_m t) \qquad (123)$$

where R_m is the real root of

$$2R_m^{2m-1} + (1-2R_m)^{2m-1} = 0. \qquad (124)$$

This yields the 7-term product above as S_4 (i.e. when $m = 2$) with $R_2 = 1/(2-2^{1/3}) \approx 1.351207192$. Similarly $R_3 \approx 1.17467176$, $R_4 \approx 1.1161829$, $R_5 \approx 1.0870271$, $R_6 \approx 1.0695657$; and

$$R_m = 1 + \frac{\ln 2}{2m} + O(m^{-2}) \text{ as } m \to \infty. \qquad (125)$$

In general when $m \geq 1$ we get an expression for S_{2m} (of order $2m$) as a $(2 \cdot 3^{m-1} + 1)$-term product. (Each product would have exactly 3 times as many terms as the previous one, except that 2 pairs of terms each coalesce.) This formula always has real coefficients, although some of them will be negative if $m \geq 2$. Fortunately, the maximum absolute value among the coefficients is

$$\prod_{j=2}^{m}\left(1 + \frac{\ln 2}{j} + O(j^{-2})\right) \;=\; \exp\sum_{j=2}^{m}\left(\frac{\ln 2}{j} + O(j^{-2})\right)$$
$$=\; e^{(\ln 2)(\ln m)+O(1)} \;=\; O(m^{\ln(2)}) \;=\; O(m^{0.69315}) \qquad (126)$$

Thus, the "oscillation" isn't too severe, and carrying $O(m^{\ln(2)})$ times more decimal places in one's calculations, than one would have normally needed, ought to suffice to take care of the effects of roundoff error. Since the product has $2 \cdot 3^{m-1} + 1$ terms, this extra work factor is in comparison not a great burden. On the other hand, for fixed matrices A and B, it appears that the infinite sequence of product formulas arising from the Suzuki recursion will *not* in general converge to a limit[74].

There is no reason to believe Suzuki's recursive construction is "optimal." Indeed Yoshida [100] found[75] an order-6 formula with only 15 (beating Suzuki's 19) terms, but still (of course) with a negative coefficient.

It is easily seen (by considering the same sort of brute force Taylor series expansion method as before) to be impossible for a k-term product of the form above to approximate better than $O(t^{k+1})$. Counting degrees of freedom suggests the conjecture that eventually for large k, the approximation order w for the best k-term product cannot be better than about $(\log_2 k) \pm O(1)$, since there are only order k coefficients to vary to try to make the $\approx 2^{w+1}$ words formed from the w letters A and B cancel. Suzuki's construction therefore comes within a constant factor of the minimum possible number of terms in the product, if this conjecture is believed.

D Appendix on Rayleigh-Ritz method

There is a conceptually simpler simulation algorithm for quantum mechanics which seems also to work to yield Church's weak thesis. Indeed, it even yields the intermediate-strength version if some clever but extremely impractical algorithmics[76] are employed to save memory in the linear algebra. This algorithm seems of less interest than ours because it apparently does not lead to Church's strong thesis on a quantum computer.

On the other hand, it does show

Theorem 25 *Let H be a multi-particle Hamiltonian operator in a computable cubical box, arising from a Rollnick ♣ interparticle potentials, and computable [58] except for in the neighborhoods of singularities over which its integral is known*[77]*. Then the* sequence *of all eigenvalues of the Hamiltonian is a computable real sequence. In other words there exists an algorithm which, given positive integers n and m, will output a rational approximation to the nth eigenvalue (for some ordering which roughly corresponds to, but may* not *precisely correspond to, increasing order*[78]*) accurate to 2^{-m}.*

[74] Let E_m denote the relative error arising from using the order-m formula, and find a recurrence something like $E_m \approx (3E_{m-1}/3^{2m})^{1+(\ln 2)/m}$, whose solution diverges to ∞.

[75] Numerically, i.e. this is not a rigorous existence proof, but is almost certainly correct.

[76] MATRIX INVERSION was shown by Csanky [23] to be in the computational complexity class "NC." It then follows that linear algebra operations such as matrix-vector multiplication and linear and least squares system solving are in PSPACE *even* if these matrices and vectors have exponentially many elements and the jk elements of the matrices and vectors are only available through "black box" function calls with arguments j and k. Indeed, one may readily verify the (perhaps stronger, and definitely not weaker) statement that all the linear algebra operations just mentioned are in $P^{\#P}$.

[77] It will suffice if the interparticle potentials V obey $|V| < Cr^{-\gamma}$ for some computable real constant C and some $\gamma < 2$.

[78] Quite possibly it is undecidable which of two eigenvalues is greater. This issue is irrelevant if we only are interested in approximate eigenvalues. We will get the ordering correct up to numerical

This is interesting because Pour-El and Richards [73] had shown that:

1. Each eigenvalue of an "effectively ♣ determined self-adjoint operator" (the Schrödinger Hamiltonians of physically realizable quantum systems qualify [47]) is a computable real number, i.e. for each eigenvalue there exists an algorithm to compute it to any demanded accuracy.

2. *But* the entire *sequence* of spectral eigenvalues is not necessarily computable, i.e. no finite-length algorithm necessarily exists to compute *any* eigenvalue. Nor is an eigenfunction corresponding to any given eigenvalue necessarily a computable function. Operators are exhibited in [73] for which both of these things were, in fact, uncomputable.

Thus, considering point 2, we now see that physically realizable Hamiltonian operators [47] are "nicer" than just any old effectively determined self-adjoint operator.

Here is a sketch of the algorithm:

1. Find all the eigenvectors of the Hamiltonian below some energy κE (where E is the energy bound on the initial state, and κ is a number we can make arbitrarily large). Due to eigenvalue counting bounds in [84] we know the number of such eigenvectors is finite and the process of finding them will terminate[79]. The eigenvectors Ψ_1, Ψ_2,... we find will be approximate. But they will be good approximations in the sense that

1. They will be exactly orthonormal (well, up to exponentially small roundoff error arising from finite precision arithmetic carried to a polynomially large number of decimal places)

2. $e^{itH}\Psi$ will differ from $e^{it\lambda}\Psi$ (where λ is some approximate eigenvalue) by an arbitrarily small amount in the L_2 norm (if t is bounded by the time the simulation is supposed to cover)[80].

The process of finding these eigenfunctions Ψ_j is the "Rayleigh Ritz method" in which we express each Ψ_j as a linear combination of Fourier modes for the box the physical system is contained in (up to some high energy) and then solve for the coefficients to minimize $\|H\Psi_j\|$ subject to the constraints of normalization and orthogonality to all previous eigenfunctions. This solve is a finite dimensional quadratically and linearly constrained least squares problem [12][62]. The entries in the matrix defining that least squares problem are each computable reals – assuming H is nice enough, i.e. the potential is in L^1 and, except for in the neighborhoods of singularities in which the integrals are known, computable (see chapter 5 of [58]). The validity of the Rayleigh Ritz method for physically realizable Hamiltonians (with $L^2 + L^\infty$ potentials ♣) was shown by Kato [47][81]. This was later improved (see theorem X.19 of [75] vol. 2) to allow Rollnick+$L^\infty(\mathbf{R}^3)$ interparticle potentials ♣. The resulting Ψ_j are not the true eigenfunctions of H both because they do not incorporate high-frequency Fourier modes, and also because two eigenvalues λ could be very close to equal, in which case the 2-space spanned by the two corresponding eigenvectors would be effectively indistinguishable from an eigenspace. But since it isn't one, our "eigenfunction" might be anywhere in this 2-space, perhaps even orthogonal to where it should be. However, thanks to lemma 13 of §6 relating kinetic to total energy, we realize that any eigenfunction of energy E cannot have a sum of absolute squares of its Fourier coefficients for modes above energy $\kappa^2 E$ being more than κ^{-2}. By making κ large enough we can force the Ψ_j to obey the two approximation properties above. Neither two close λs, nor mode cutoffs, bother us for the purpose of obtaining these approximation properties.

2. Express the initial state as a linear combination of these Ψ_j. The coefficients may be obtained by numerical integration and using orthonormality. We use the fact [73] that the integrals of "computable continuous functions" (which ours are, assuming the initial state wavefunction is[82]) are computable reals. Indeed they are computable in #P ♣ if the integrand is a "polynomial time computable" real function [58].

Eigen-wavefunctions arising from Coulombic interparticle potentials have bounded gradient [47] [48] and are differentiable everywhere except at the Coulomb singularities, and hence in view of our algorithm for computing them, are computable continuous functions in the sense of [73].

3. Time-evolve the Ψ_j by multiplying them by the appropriate complex exponential phase factors.

The error introduced by this algorithm will be arbitrarily small in norm because the expression of the initial state as a sum of approximate eigenfunctions has arbitrarily small norm error, and each approximate eigenfunction's time evolution (since time evolution, being unitary, does not magnify error norms) can be made arbitrarily more close to being given by the complex exponential phase factor multiplication we recommended.

(Every time we say "very close" or "arbitrarily small" above, we of course have an effective, i.e. computable,

imprecisions in calculating those eigenvalues.

[79] Actually, the bounds in [84] are asymptotic rather than effective ♣. This still suffices to force termination provided there is some way to detect when we have exceeded the cutoff energy – and there is. Anyway, effective (although worse) upper bounds for eigenvalue counts for bound systems (i.e. every system confined to a box) are also available [63] *provided* the potential is in L^{3N} for an N-particle 3D system (which is the case if we regularize the potentials beforehand as in §6.1, and this is justifiable, in the sense that the approximate eigenvalues E in our sense will be negligibly perturbed by the regularization, if the bounds in that section tell us that time-evolution for time h/E will result in negligible perturbation of the wavefunction in L^2 distance.)

[80] Our approximate eigenfunctions need not be close to any genuine eigenfunction, but that does not matter for our purposes.

[81] Indeed, Kato claims the eigenfunctions are bounded everywhere, continuous, and have bounded derivatives everywhere except at the Coulomb singularities. (This need not be true for non-eigenfunctions.)

[82] And this is pretty much a necessary assumption — since how can one even "read the input," otherwise?

bound. Our requirement that the quantum system live in a finite box possibly can be dispensed with by using known exponential falloff results about Schrödinger eigenfunctions [20][82][79].)

E Appendix: The boundedness of an exponentially growing function of the energy

The assumption that the expected value of the energy is bounded suffices for the purposes of this paper[83]. But at first, I had thought the following much stronger assumption (not as strong as "boundedness" of all operators, but conferring nearly the same benefits) about the energy of the quantum system being simulated, was going to be required. It may in the future allow considerably stronger results than mine to be obtained (in circumstances in which the Assumption is justified), and it is highly interesting for its own sake.

Assumption 26 (Exponential boundedness of reasonable physical systems) *For some fixed $C > 1$, assume that the expectation value of C^E, where E is the energy of any "reasonable" physical system (measured in some agreed-upon units) is bounded by $\mathcal{E}$.*

Obviously Schrödinger time evolution preserves this class (if the Hamiltonian operator is self-adjoint [75], which we assume it is, for "reasonable" physical systems [47]).

The assumption that some exponential function C^E of a system's energy has bounded expectation value $\mathcal{E}$ is an extremely powerful and useful assumption, and also (we would like to argue) can nevertheless be a very realistic assumption. I'll now describe why this assumption can drastically simplify a great deal of rigorous quantum mechanics, and then indicate why it may correspond[84] to physical "reasonableness."

In books [95] [75] on rigorous quantum mechanics, the majority of the book is generally devoted to telling us how horrible linear operators, especially unbounded operators, are, as compared to the comparatively simple and pleasant world of finite dimensional matrices. The whole story of the advance of rigorous quantum mechanics has been the story of the accumulation of more and more mathematical results enabling one to handle various classes of such operators in various ways.

Here are 4 standard examples of ways in which operators are not as nice as finite dimensional matrices.

1. Matrix exponential. If M is a finite dimensional matrix, then one may define the exponential function simply by using the everywhere converging Taylor series $\exp(M) = I + M + M^2/2! + M^3/3! + \ldots$. But: this series need not converge if M isn't finite dimensional. A counterexample: let M be diagonal with $M_{jj} = j$ for all $j \geq 1$.
2. Left and right inverses. If A and B are finite dimensional matrices, then $AB = I \Rightarrow BA = I$. But this is not the case for infinite dimensional matrices; a counterexample are the right and left shift operators acting on a space indexed by the positive integers.
3. Balls in finite dimensional spaces are "compact," allowing one to make statements about limits easily. Not so in infinite dimensional spaces.
4. If $M = M^*$ (M equals its Hermitian conjugate) then M is self-adjoint. This is true for finite dimensional matrices, but not infinite ones (a fact that has rarely, if ever, worried a physicist).

Now: using this hall of horrors as a testing ground, let us contrast this with the situation in which all our operators are quantum mechanical and we are only interested in what happens when they are applied to "physically reasonable" wavefunctions Ψ satisfying assumption 26.

1. Although the Taylor series for $\exp(M)$ still needn't converge, the series for $\exp(M)\Psi$ does, in many cases, in particular if M is k times the Hamiltonian operator for any complex k with $|k|$ sufficiently small.
2. Although $AB\Psi = \Psi$ for all Ψ in our class still does not imply $BA\Psi = \Psi$ in general, for operators A, B with complete orthonormal sets of eigenfunctions, i.e. the ones of primary concern in quantum mechanics, $AB = I \Rightarrow BA = I$, so this issue is ignorable, and in fact was already ignorable without the need for our exponential assumption.
3. We are now most commonly dealing, not with balls such as $\sum_{j\geq 0} |x_j|^2 \leq 1$, but rather with sets such as $\sum_{j\geq 0} 10^j |x_j|^2 \leq 1$. (If 10^j is the "exponential function" being considered.) These sets *are* compact.
4. If we restrict our attention to Ψ satisfying our assumption, then if $M = M^*$, then M will behave exactly like a self-adjoint operator *on the set of permissible Ψ only.* This is because no issues can arise about a difference between the domains of M and M^*, since both have domain that includes the entire set of interest.

So, the reader now has some idea why wavefunctions with bounded expected exponential of energy, are a very pleasant class, and why assumption 26 has the power to make rigorous quantum mechanics much easier if and when it applies. It also has the power to make the operator analysis required to get rigorous error bounds for quantum simulation algorithms, much easier.

But why should assumption 26 correspond to "physically reasonable" wavefunctions?

[83] Indeed, it appears possible to weaken this assumption and instead merely assume that the expectation of E^p, for some constant p with $0 < p < 1$, is bounded. See §12.5.

[84] But see footnote 55.

Well, to begin with, it seems reasonable that the expectation value of energy itself, should be bounded. But mathematicians can produce plenty of probability densities (e.g. proportional to $(1+E^2)^{-1.1}$) in which an expectation value $\overline{E}$ exists but a variance does not – much less this (far more divergent) exponential quantity.

So: here is a physico-philosophical argument that the expectation value of C^E should be bounded, for some $C > 1$.

Suppose you take a physical system and lift it onto your lab table. If you can't put it on your lab table, it was not a "reasonable" system to simulate[85]. You know you cannot lift 1000Kg. So when you lifted it, you were performing a Copenhagen *measurement* which collapsed the wavefunction down to a state with *zero* probability of energy$>$ 1000Kg$\times c^2$. So, you can now work with bounded operators, indeed even with finite dimensional matrices, if there are only a finite number of energy eigenmodes below this energy. And every increasing function of energy, in particular 10^x, must have bounded expectation value.

A more refined version of the same argument: Really you and the lab system constitute a bigger combined system. Before turning on your experiment, you first isolated it. But before isolating it, you lifted it onto the lab table. So you were not *really* performing a "measurement" since we now (being more sophisticated than the Copenhagen interpretation) realize that there was nothing "classical" here performing a measurement – you and the system were, together, just some quantum system. However, still, it seems reasonable to assume it was exponentially improbable that you managed to lift 1000Kg – maybe not 0 probability, but exponentially small.

Specifically, assume a Boltzmann max-entropy thermal energy distribution governs such fluctuations. This falls off like $\exp(-E/[k_B T])$ at a fixed temperature T. The probability that, by some astonishing statistical fluke, you happened to be so far in the tail of the Boltzmann distribution that you managed to lift 1000Kg up 1 meter, which was, say, 9000 Joules above your normal ability, is about

$$\exp \frac{-9000 \text{ Joules}}{4 \times 10^{-21} \text{ Joules}} \approx 10^{-10^{24}} \tag{127}$$

at room temperature.

Under such conditions, some exponentially increasing function of system energy will indeed have bounded expectation value!

A different reason I think assumption 26 is profoundly interesting is that it seems to have some relation to *analyticity*[86]. To see what I mean, consider the following[87] lemmas

Lemma 27 *The function*

$$f(x) = \sum_j a_j \, e^{ijz} \tag{128}$$

is an analytic function of z for all real z, and indeed for all z in a strip of nonzero width centered on the real axis, if and only if

$$\sum_j |a_j| \, C^{|j|} \tag{129}$$

converges for some $C > 1$.

In other words, wavefunctions with bounded expected exponential of *kinetic* energy are analytic. Furthermore

Lemma 28 *If there is some point x at which $f(x)$ is non-analytic (for example, it does not have a kth derivative there, for some k) then (EQ 129) diverges.*

F Acknowledgements & disparagements

Henry Cejtin wrote and ran the Gröbner basis package mentioned in §C.0.5. Dominic Mayers, Leonid Gurvits, and Nick Littlestone gave me useful comments on preliminary drafts. Abrams and Lloyd [67][1] had earlier claimed that quantum computers were universal simulators for quantum systems, but the "systems" they were considering were discrete rather than continuous, which made their job comparatively trivial.

References

[1] Daniel S. Abrams & Seth Lloyd: Simulation of many body Fermi systems on a universal quantum computer, Phys. Rev. Lett. 79,13 (1997) 2586-2589.

[2] L.Adleman, J.DeMarrais, M.Huang: Quantum computability, SIAM J.Computing 26,5 (1997) 1524-1540.

[3] T. Arima et al.: Precise measurement of Bhabha scattering at a center-of-mass energy of 57.77 GeV, Physical Review D 55 (1997) 19-39.

[4] T.Becker & V. Weispfenning: Grobner bases: a computational approach to commutative algebra, Springer-Verlag 1993.

[5] W. Beenakker, F.B. Berends, S.C. van der Marck: Large-Angle Bhabha Scattering, Nucl. Phys. B 349 (1991) 323-368.

[85] Provided "Church's thesis" only requires the computer to have the ability to simulate experiments one could actually have run in a lab.

[86] Incidentally, more than one physics paper contains an unsupported statement of the form "because the interparticle potentials are analytic functions, the wavefunction must be analytic." Unfortunately, such statements, although perhaps true for some class of initial conditions, have never been justified. The best that seems currently known [47] [48] is that, for example, eigen-wavefunctions arising from Coulombic interparticle potentials have bounded gradient, except at the Coulomb singularities, and (Kato claims the following, but without providing a proof, in [47]) have second partial derivatives almost everywhere. No proof of any differentiability beyond this, is known in general, although for the 1-particle *one-dimensional* Schrödinger equation, there is an analyticity result [40].

[87] These are easy consequences of well known results in complex analysis [54] [55].

[6] V.B. Berestetskii, E.M. Lifshitz, and L.P. Pitaevskii: Quantum electrodynamics, (2nd ed.) Pergamon Press (Oxford ; New York) 1982.

[7] W.Beckner: Geometric proof of Nash's inequality, Int'l. Math. Res. Notices #2 (1998) 67-72.

[8] W.Beckner: Pitt's inequality and the uncertainty principle, Proc. Amer. Math. Soic. 123 (1995) 1897-1905.

[9] J.Benedetto & H.Heinig: Weighted Hardy spaces and the Laplace transform, Springer Lecture Notes in Math. #992 (1983) 240-277.

[10] E.Bernstein & U.Vazirani: Quantum complexity theory, SIAM J. Comput. 26,5 (1997) 1411-1473.

[11] H.J.Bhabha: The scattering of positrons by electrons with exchange on Dirac's theory of the positron, Proc. Royal Soc. London A154 (1936) 195-206.

[12] Ake Bjorck: Numerical methods for least squares problems, SIAM 1996.

[13] N.Bohr: Philosophical magazine 26 (1913) 1-25, 476-502, 857-875; Nature 95 (23 Oct 1913) 231-232.

[14] J.M.Borwein & P.B.Borwein: Pi and the AGM, a study in analytic number theory and computational complexity, Wiley 1987.

[15] Michel Boyer, Gilles Brassard, Peter Hoeyer & Alain Tapp: Tight bounds on quantum searching, Fortschritte der Physik 46, 4-5 (1998) 493-505.

[16] Yu.A. Brychkov, H.J.Glaeske, A.P. Prudnikov, & Vu Kim Tuan: Multidimensional Integral Transformations, Gordon and Breach 1992.

[17] Harry Buhrman, Richard Cleve, John Watrous, & Ronald de Wolf: Quantum fingerprinting, to appear in Physical Review Letters, quant-ph/0102001.

[18] K.T.Cheng & W.R.Johnson: Self-energy corrections to the K-electron binding in heavy and superheavy atoms, Phys. Rev. A 14,6 (1976) 1943-1948.

[19] E.R.Cohen & B.R.Taylor: The 1986 CODATA Recommended Values of the Fundamental Physical Constants, J.Res.Nat'l Bur. of Standards USA 92,2 (Mar-Apr 1987).

[20] J.Combes, L.Thomas: Asymptotic behavior of eigenfunctions for multi particle Schrödinger operators, Commun.Math.Phys. 34 (1973) 251-270.

[21] Handbook of Convex Geometry (2 vols, ed. P.M.Grüber and J.M.Wills) Elsevier 1993.

[22] D.A.Cox, J.Little, D. O'Shea: Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra (2nd ed.), Springer 1997.

[23] L.Csanky: Fast parallel matrix inversion algorithms, SIAM J. Comput. 5,4, (1976) 618-623.

[24] Charles G. Darwin: Proc. Royal Soc. London A 118 (1928) 654-

[25] P.Deift & E.Trubowitz: Inverse scattering on the line, Commun. Pure & Applied Math. 32 (1979) 121-251.

[26] G.W.F. Drake (ed.): Atomic molecular and optical physics handbook, Amer. Inst. of Physics 1996. Contains the following survey articles relevant to us:
G.W.F. Drake: High precision calculations of Helium 154-177.
J.Sapirstein: Quantum electrodynamics 327-340.
P.J.Mohr: Tests of fundamental physics 341-351.
M.R.Flannery: Elastic scattering, classical, quantal and semiclassical 499-525.

[27] L.D.Fadeev: The inverse problem in the quantum theory of scattering (trans. from the Russian by B.Seckler) J.Math.Phys. 4 (1963) 72-104.

[28] R.P.Feynman: Space-time approach to non-relativistic quantum mechanics, Reviews in Modern Physics 20 (1948) 367-387.

[29] R.P.Feynman & A.R.Hibbs: Quantum mechanics and path integrals, New York, McGraw-Hill 1965.

[30] R.P. Feynman & Frank Vernon: The theory of a general quantum system interacting with a linear dissipative system, Annals of Physics 24 (1963) 118-173.

[31] E.Forest & R.D.Ruth: Fourth-order symplectic integration, Physica D 43 (1990) 105-117. (See eq 4.3.)

[32] L.Fortnow, J.Rogers: Complexity limitations on quantum computation, J. of Computer and System Sciences, to appear.

[33] Z.Ge & J.E.Marsden: Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson operators, Phys.Lett. A 133 (1988) 134-139.

[34] Walter Gordon: Z. Phys. 48 (1928) 11-

[35] Kurt Gottfried & Victor F. Weisskopf: Concepts of particle physics (2 vols), Oxford University Press 1984-1986.

[36] R.L.Graham, D.E.Knuth, O.Patashnik: Concrete mathematics, Addison-Wesley (2nd ed.) 1994.

[37] W.Greiner & J.Reinhart: Quantum electrodynamics (2nd ed.) Springer 1994.

[38] Lov K. Grover: A fast quantum mechanical algorithm for database search, Proceedings Sympos. Theory of Comput. (1996) 212-219.

[39] F.Haake: Quantum signatures of chaos, Springer 1991.

[40] Nakao Hayashi & Saburou Saitoh: Analyticity and smoothing effect for the Schrödinger equation, Annales de l'Institut Henru Poincaré 52,2 (1990) 163-173.

[41] H.Heinig: Weighted norm inequalities for classes of operators, Indiana Univ. Math. J. (4) 33 (1984) 573-583. (See also C.Carton-Lebrun & H.P.Heinig: SIAM J. Math. Anal. 23,3 (1992) 785-798 and Math. Rev. 93e:42012.)

[42] . J.O.Hirschfelder, C.F.Curtis, R.B.Bird: Molecular theory of gases and liquids, Wiley 1954, reprinted with additions & corrections 1964.

[43] J.M.Jauch & F.Rohrlich: The theory of photons and electrons, Springer (expanded 2nd ed.) 1976. (12-2 covers Bhabha scattering.)

[44] D.S.Johnson: A catalogue of complexity classes, pp. 67-162 in Handbook of theoretical computer science volume A, (J.van Leeuwen ed.) MIT Press & Elsevier 1990.

[45] W.R. Johnson & G. Soff: The Lamb shift in hydrogen like atoms $1 \leq Z \leq 110$, Atomic Data & Nucl. Data Tables 33 (1985) 405-446.

[46] W.Jurkat & G.Sampson: On rearrangement and weight inequalities for the Fourier transform, Indiana Univ. Math. J. (4) 33 (1984) 257-270; On maximal rearrangement inequalities for the Fourier transform, Trans. Amer. Math. Soc. (2) 282 (1984) 625-643.

[47] Tosio Kato: Fundamental properties of Hamiltonian operators of Schrödinger type, Trans. Amer. Math. Soc. 70 (1951) 195-211.

[48] Tosio Kato: On the eigenfunctions of many particle systems in quantum mechanics, Commun. on Pure & Applied Math. 10 (1957) 151-177.

[49] Tosio Kato: Schrödinger operators with singular potentials, Israel J. Math. 13 (1972) 135-148.

[50] Tosio Kato: On the Trotter-Lie product formula, Proc. Japan Acad. 50 (1974) 694-697.

[51] T.Kato: A second look at the essential self-adjointness of Schrödinger operators, pp. 193-201 in "Physical reality and mathematical description", Reidel Dordrecht 1974.

[52] Quantum electrodynamics (T.Kinoshita ed.) World Scientific 1990.

[53] T.Kinoshita: Fine structure constant obtained through an improved calculation of the electron $g-2$, IEEE Trans.Instrum.Meas. 46,2 (1997) 108-111.

[54] K.Knopp: Elements of the theory of functions, Dover Publications 1953.

[55] K.Knopp: Infinite sequences and series (translated by Frederick Bagemihl) Dover Publications 1956.

[56] D.E.Knuth: Seminumerical methods, (2nd ed.) Addison-Wesley 1973.

[57] D. E. Knuth: Big omicron and big omega and big theta, SIGACT News 8,2 (Apr.-June 1976) 18-24.

[58] Ker-I Ko: Complexity Theory of Real Functions, Birkhauser, 1991.

[59] L. D. Landau & E. M. Lifshitz: Mechanics, Pergamon 1976.

[60] L. D. Landau & E. M. Lifshitz: Quantum Mechanics (non-relativistic theory), (3rd ed.) Pergamon 1977.

[61] S.Laporta & A.Remiddi: The electron $g-2$ and the value of α: a check of QED at 1ppb, Acta Phys. Polonica B 28,3-4 (1997) 959-977.

[62] C.L.Lawson & R. J. Hanson: Solving least squares problems, Prentice-Hall, Englewood Cliffs NJ, 1974.

[63] E.H.Lieb: The stability of matter from atoms to stars, selecta of E.H.Lieb, Springer 1997. Papers included in this volume especially relevant to us are:
Eigenvalue counting:
#III.2: 269-303 in "Studies in mathematical physics" Princeton University press 1976 (with W.Thirring),
#III.4: Proc.Sympos.Pure Math. 36 (1980) 241-252,
Uncertainty principle:
#V.1: Phys.Rev.Lett. 35 (1975) 687-689 & 1116 (with W.Thirring) and
#VI.1: Rev.Mod.Phys. 48 (1976) 553-569.
Issues probably relevant to extending the present work to allow magnetic fields and/or relativity:
#V.3: Commun.Math.Phys.104 (1986) 251-270 (with J.Fröhlich & M.Loss),
#V.4: Commun.Math.Phys.104 (1986) 271-282 (with M.Loss),
#V.5: Phys.Rev.Lett. 75 (1995) 985-989 (with M.Loss & J.P.Solovej),
#V.9: Phys.Rev.Lett. 61 (1988) 1695-1697 (with H-T.Yau),
#V.10: Commun.Math.Phys. 118 (1988) 177-213 (with H-T.Yau),
#V.11: Helv.Phys.Acta 69, 5/6 (1996) (with M.Loss & H.Siedentop),
#V.12: J.Stat.Phys. 89 (1997) 37-59 and Phys.Rev.Lett. 79 (1997) 1785-1788 (with H.Siedentop & J.P.Solovej).

[64] E.H.Lieb: Kinetic energy bounds and their application to the stability of matter, pp. 371-382 in *Schrödinger operators*, ed. H.Holden & A.Jensen, Springer Lect.Notes Phys. #345 (1989). Reprinted 191-202 in [63].

[65] E.H.Lieb: Uncertainty principle, short article in Encyclopedia of Physics, ed. R.Lerner & G.Trigg, Addison-Wesley 1991.

[66] E.H.Lieb, M.Loss & J.P.Solovej: Stability of matter in magnetic fields, Phys.Rev.Lett. 75 (1995) 985-989. (Reprinted in [63].)

[67] Seth Lloyd: Universal quantum simulators, Science 273 (23 Aug 1996) 1073-1078.

[68] B.Muckenhoupt: Weighted norm inequalities for the Fourier transform, Trans. Amer. Math. Soc. 276 (1983) 729-742; also see Proc. Amer. Math. Soc. (1) 88 (1983) 97-100.

[69] Heide Narnhofer: Quantum theory for $1/r^2$ potentials, Acta Physica Austriaca 40 (1974) 306-322.

[70] A.W.Naylor & G.R.Sell: Linear operator theory in engineering and science, (1st ed.) Holt Rinehart Winston 1971; (2nd ed.) Springer 1982.

[71] Edward Nelson: Feynman integrals and the Schrödinger equation, J. Math. Phys. 5 (1964) 332-343.

[72] Bernhard Ömer: A Procedural Formalism for Quantum Computing, Department of Theoretical Physics Technical University of Vienna, 1998 `http://tph.tuwien.ac.at/~oemer/qc/qcl/qcl.html`.

[73] Marian Boykan Pour-El and J.Ian Richards: Computability in analysis and physics, Springer-Verlag 1988. (Reviewed: Bull. Amer. Math. Soc. 24 (1) (1991) 216-228; Jahres Deutsch. Math. Verein. 94 (1992) 53-55.)

[74] John Preskill: *Reliable quantum computers*, Proc. Roy. Soc. Lond. A 452, 1969 (8 Jan 1998) 385-410. (Special issue of PRS A on "Quantum Coherence and Decoherence.")

[75] M.Reed & B.Simon: Methods of modern mathematical physics (4 vols, revised & enl. ed.), Academic Press 1980.

[76] Ronald D. Ruth: A canonical integration technique, IEEE Trans. Nucl. Sci. NS-30,4 (1983) 2669-2671. (See eq 30.)

[77] Michael D. Scadron: Advanced quantum theory and its applications through Feynman diagrams, Springer (2nd ed.) 1991.

[78] L.I.Schiff: Quantum mechanics, (3d ed.) New York, McGraw-Hill 1968.

[79] I.E.Schnoll: On the behavior of eigenfunctions of the Schrödinger equation, Math.Sbornik 42 (1957) 273-286.

[80] L.Schulman: Techniques and applications of path integration, Wiley 1981.

[81] P.W.Shor: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. 26,5 (1997) 1484-1509.

[82] B.Simon: Pointwise bounds on eigenfunctions and wave packets in N-body quantum systems,
I: Proc. Amer. Math. Soc. 42 (1974) 395-401;
II: Proc. Amer. Math. Soc. 45 (1974) 454-456;
III: Trans.Amer.Math.Soc. 208 (1975) 317-329;
IV: (with Deift, Hunziker, Vock) Commun. Math'l. Phys. 64,1 (1964) 1-34;
V: (with Carmona) Commun. Math'l. Phys. 80,1 (1981) 59-98.

[83] W.D.Smith: Church's thesis meets the N-body problem, NECI tech report dated 09/09/93.

[84] W.D.Smith: Fundamental physical limits on information storage , NECI tech report dated 07/31/98.

[85] W.D.Smith: Argument against quantum computers (or against certain decoherence models), report dated March 2003. Presented at Simons Conference Quantum information & Computation, SUNY Stony Brook, May 2003.

[86] W.D.Smith: History of "Church's theses" and a manifesto on converting physics into a rigorous algorithmic discipline, NECI tech report dated Feb. 1999. All these NECI tech reports are available electronically at `http://math.temple.edu/~wds/homepage/works.html`

[87] A.Sommerfeld: Zur quantentheorie der Spektrallinien, Ann.Physik 51 (1916) 1-94; 125-167. The relativistic attack is pages 44-94. Sommerfeld later rederived the same results using a relativistic version of the Hamilton-Jacobi *partial* differential equation formulation of classical mechanics in his book "Atomic structure and spectral lines," Methuen London 1923 (translated from the 3rd German edition, Vieweg Braunschweig 1919, by H.L.Brose) pp. 467-496; 608-611. That surely served as inspiration for E.Schrödinger in deriving *his* equation.

[88] Masuo Suzuki: On the convergence of exponential operators – the Zassenhaus formula, BCH formula, and systematic approximants, Commun. Math. Phys. 57 (1977) 193-200.

[89] Masuo Suzuki: Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics, J.Math.Phys. 26,4 (1985) 601-612.

[90] Masuo Suzuki: Fractal decomposition of exponential operators with applications to many body theories and Monte Carlo simulations, Physics Letters A 146,6 (June 1990) 319-323. (See eq 30 and 31).

[91] Masuo Suzuki: General theory of fractal path integrals..., J. Math'l. Phys. 32,2 (Feb 1991) 400-407.

[92] Masuo Suzuki: Convergence of exponential formulas for unbounded operators, Reviews Math. Phys. 8 (1996) 487-502.

[93] G.Talenti: Best constant in Sobolev inequality, Ann. di Matematica Pura ed Applicato (Ser. 4) 110 (1976) 353-372.

[94] Max Tegmark: Apparent wave function collapse caused by scattering, Found. Phys. Lett. 6 (1993) 571-590.

[95] W.Thirring: A course in mathematical physics, vol 3: quantum mechanics of atoms and molecules, (2nd ed.) Springer 1992.

[96] Edward A. Uehling: Polarization effects in the positron theory, Phys. Rev. 48 (1935) 55-63.

[97] Ken Umeno: Simulating quantum non-integrable systems with quantum computers, 319-322 in Fourth workshop on physics and computation, Boston University (22-24 Nov 1996).

[98] S.Varadarajan: Lie groups, Lie algebras, and representations, Springer (GTM 102) 1984.

[99] W.K. Wootters & W.H.Zurek: A single quantum cannot be cloned, Nature 279 (1982) 802-803.

[100] H.Yoshida: Construction of higher order symplectic integrators, Physics Letters A 150 (1990) 262-268.

[101] Wojciech H. Zurek: Decoherence and the transition from quantum to classical, Physics Today 44 (October 1991) 36-44. [See also the Letters section, pages 13-15 and 81-90 of the April 1993 issue.]