History of "Church's theses" and a manifesto on converting physics into a rigorous algorithmic discipline

Warren D. Smith
wds@research.NJ.NEC.COM

date = June 14, 1999

Abstract — **Church's thesis claims that any "reasonable computer" may be simulated by a Turing machine. The "strong" thesis says that the simulation may be performed with only polynomial slowdown. This document is both a history of "Church's thesis" – and particularly of the notion that it is a statement about *physics* – and an opinionated philosophical statement.**

Keywords — Church's thesis, rigorous physics, polynomial time, effective continuous mathematics, computable real numbers, philosophy of science

1 Importance of Church's thesis

1.1 Physics

The task of physics is to predict Nature. Such prediction is an algorithmic task. Therefore, formulations of physics leading to slower algorithms are worse than formulations leading to more efficient algorithms. In fact, if "the" formulation of physics does *not* have an algorithmic formulation *at all*, then I would regard that as a crisis in physics analogous to, and probably even worse than, the crisis in mathematics that came with the realization that there were undecidable mathematical problems [36] [65] [25][1].

But, all this has, so far, been largely ignored by physicists. It is currently unknown which theories of physics are even decidable – the most basic possible computational complexity question. This open question is identical to the question of resolving Church's thesis (CT) for each physical theory.

Extremely basic foundational mathematical and algorithmic questions in physics remain unanswered. For example, consider the Navier-Stokes equations, and Einstein's field equations of general relativity. In both cases, existence and uniqueness[2] of a solution is *not known*. Algorithms with rigorous error bounds are consequently unknown.

1.2 Mathematics

Consider the role of computer science in mathematics. Before the advent of computer science and complexity theory, it was unclear how to compare two mathematical results. Call them theorem A and theorem B. Which is better? The modern answer is, theorem A is better if it has a constructive proof leading to an algorithm (as compared to a nonconstructive proof). It is even better if the algorithm is more efficient. With the aid of complexity theory we can try to find out algorithm *performance measures* and compare theorems. Computer science provides, for the first time in mathematics, a measuring stick.

Also, it provides a way to spot weak theorems likely to be susceptible to improvement.

Claim: computer science can and should play a similar role in theoretical physics.

[1] It was pointed out to me that the realization that quantum physics was inherently only capable of making *statistical* predictions, was a similar crisis in physics. It overturned Einstein's famous religious conviction that "God does not play dice." I cannot resist pointing out that Karl Popper then argued that in fact God *must* play dice, because an all-powerful being with unlimited computational power would see no point in creating a universe whose entire future behavior was already known to him. I think anybody agreeing with that reasoning must now go further and conclude that God would not even see any point in flipping dice (creating universes) whose probability distributions (of possible histories of that universe) were all completely known to him, and hence the ultimate laws of physics also *must* be non-algorithmic! But (as an atheist), while admitting the amusement value of these three religious arguments, I cannot endorse any of them.

[2] A review of what is known about the existence and uniqueness problems for general relativity are in a December 1998 manuscript by S.Klainerman (Princeton math. dept.) and F.Nicolo (Rome), 134 pages; see also [8]. For Navier-Stokes, existence and uniqueness are known under certain conditions involving low "Reynolds numbers." Existence is known under reasonable initial conditions for some nonzero amount of time (it is unknown how much). Nonuniqueness is known at high Reynolds numbers. (I am being purposely vague about the exact conditions, which are expressed as Sobolev norm bounds for the initial conditions and the driving forces.) There might be some natural way to ignore unstable solutions and then (hopefully) regain uniqueness. Phenomena such as shock waves, tornados, and turbulence cause some to doubt eternal existence; probably a poll of experts in the area would get about a 50-50 split on conjectured answers to the existence question. Existence is known for L^2 driving forces if time-invariance is imposed artificially; the question is whether the "natural" (whatever they are) time-evolving solutions will continue to exist. See [23] for a deeper look at the Navier-Stokes question.

1.3 Computer science

CT lies at the foundation of computer science. It is the very essence of what "computability" is.

If it turned out to be possible to build a physical device capable of flouting CT by computing an "uncomputable" function, then most of computer science would lose its relevance.

The "polynomial" or "strong" CT is nearly as important. Without it the whole theory attached to "P and NP" loses relevance. For example, if a device were available to quickly solve "boolean satisfiability" problems (NP-complete), then almost every part of computer science (compilers, debuggers, architecture, cryptography...) would be turned upside down. Mathematicians would no longer be needed, because anyone could simply ask "O device: pray tell me whether there is a proof of Riemann's hypothesis[3] less than 20000 pages long." (Such problems are obviously in NP, since the task of verifying proofs in obviously in P; cf. footnote 11.)[4] "Quantum computers" constitute a strong attack on the strong CT and hence should be viewed with the greatest excitement or trepidation.

2 Origins

Alonzo Church (1903-1995) proposed [9] the idea that his "recursive" functions encapsulated the notion of "computability." He phrased this, in fact, as a definition. K.Gödel (1906-1978) [13] had had almost the same idea earlier. Recursive functions were shortly thereafter shown by A.M.Turing (1912-1954) to be what Turing machines compute [65]. Turing gave arguments showing that an abstract human being could be emulated by a Turing machine. These arguments sound convincing but they ignore the possibility that the human might build a physical computing device, depending for its operation on quantum mechanics, say[5], and then use that device. Emil Post (1897-1954) [43] was the first to express doubt[6] about CT (in his footnote 8): "Church... [views this as more than a] working hypothesis... But to mask this identification under a definition hides the fact that a fundamental discovery on the limitations of the mathematical power of Homo Sapiens has been made and blinds us to the need of its continual verification[7]."

But Post seems to have been the only one in this era expressing any skepticism.

[3] Or any other mathematical statement.

[4] Of course, the impact of a device for solving *uncomputable* problems, e.g. Turing's "halting problem," would be vastly more dramatic. I must admit, though, that if that device solved a comparatively useless undecidable problem, it might not have much impact. But all the undecidable problems currently known are essentially equivalent to the halting problem.

[5] Turing had never considered quantum weirdness in his arguments.

[6] Although Post too was on Church's side; he just apparently felt that more caution was required.

[7] All the papers cited so far this section are reprinted in [14].

3 Connection to physics: initial thoughts

The fact that the strong and weak Church theses are really statements about *physics* seems to have been unnoticed, or anyhow not shouted on the streets, perhaps until the paper of Vergis, Steiglitz, and Dickinson [70] in 1986. This paper, besides pointing this out, showed that the classical mechanics of rigid bodies plus an intuitively appealing "downhill principle" would allow the construction of a machine which would solve NP-complete problems in polynomial time – thus disproving Church's strong thesis if P≠NP. (The downhill principle was later overthrown by Smith [63]; in that manuscript it is shown that the methods of Vergis et al. can never lead to a refutation of Church's strong thesis in either classical or quantal mechanics.)

There had been a philosophical paper by Georg Kreisel (1923-) [33] in 1974 in which Kreisel had explicitly worried that some physical systems might "compute" things that Turing machines could not. Kreisel explicitly mentioned the possibility that some quantum system might have two spectral energies λ_1 and λ_2 whose ratio was an uncomputable real. Although this particular suggestion is false[8], the point – that physical systems might enable super-Turing computational power – had been made. Kreisel also made the important point here (which I agree with totally) that the "reader should not be confused... by doubts about the validity of the theory with regard to the phenomena for which it is intended." That is, Kreisel wanted to know whether CT was true in particular theories of physics, *regardless* of whether that theory was "the" theory of physics.

Kreisel's thoughts about Church+physics began at least as early as 1965. Here is a quote from [32] (section 2.714 and footnote):

> ...(excepting collisions in the 3-body problem, which introduce discontinuities) the theory of partial differential equations shows that the behavior of discretely described (finite) systems of classical mechanics is recursive. But this may not be so in the quantum theory... of large molecules. Footnote: [does] this theory [permit] mechanical approximations?"

This Kreisel quote may either be interpreted as very prescient or totally confused. Examples of confusion: collisions, if treated as elastic bounces, do *not* introduce discontinuities. "The theory of partial differential equations" remains unreconciled with computability theory and has little to do with the n-body problem either. His whole statement about classical mechanics is arguably in conflict with my own later result [59] (although stated intentionally vaguely enough that it might be true).

My personal feeling is that although in these and other writings Kreisel broached the idea that the validity of CT with respect to various physical theories was worth investigating, he never made any such investigation, never pen-

[8] It was ruled out by Pour-El and Richards (see §5), and even if it had been the case, since finite lifetimes cause nonzero "natural line widths," there is doubt this could have permitted defeating CT (although related pathologies conceivably could do so).

etrated deeply, and never got any results, i.e. theorems. Vergis et al (1986), in contrast, did.

Earlier, John Myhill in a US Air Force technical report [Creative computation revisited, 1966] had claimed that certain idealized analog computers could "compute" nonrecursive functions; but I have not been able to obtain a copy of this report. R.Rosen [55] also seemed to realize that CT is about physics, but doesn't go much further than that[9].

S.C.Kleene (1909-1994) in his 1951 book [29] had surveyed the area extensively, but at no point, really, viewed the thesis as a statement about physics; the most physical arguments made up to that point seem to have been Turing's original ones.

Some present day (1999) logicians *still* deny that CT is physical. For example, Martin Davis [private communication] told me he thought CT had nothing to do with physics because "the universe is probably finite... CT is about infinite objects [such as Turing machines][10]." Some others have argued, following Church's approach [9], that CT is merely a *definition* of "computability" and as such remains valid no matter what.

Davis in his book [12] had asked:

> ...how can we ever exclude the possibility of our being presented someday (perhaps by an extraterrestrial visitor) with a (perhaps extremely complex) device or "oracle" that "computes" a noncomputable function?

Davis intended this only as a rhetorical "straw man," since he then demolished it in short order by a recapitulation of Turing's non-quantal arguments. It seems to me, though, that Kreisel's idea for constructing such a device using quantum mechanics might have given Davis pause. Also, in the event that such devices became commonplace, the CT "definition" of computability would become about as interesting as a "definition" of transportation forbidding the use of airplanes and ships.

So I think we are forced to concede that CT really *is* about physics.

4 Polynomial time

The "strong," or "polynomial time" CT claims that any physical system, or reasonable computational device, may be simulated with only "polynomial slowdown."

I would hereby also like to propose the "intermediate strength CT:" a Turing machine may simulate any "computer" with at most polynomial increase in *memory space* consumption.

The notion of polynomial time equivalence of computational models, and the notion that "polynomial time" meant approximately the same as "tractable," was not due to Church, and seems to have been promulgated by Jack Edmonds in the early 1960s [18]. Others also promulgating this notion included A.Cobham [10] and M.Rabin [53]. The history of this has been surveyed in [58] which also discusses a 1956 letter from K.Gödel to J.von Neumann in which the P=NP question is in essence discussed, and in which Gödel pointed out that the best lower bound he could show on the time complexity of the (NP-complete) problem of "provability in k symbols" is of order k but of course he would presume the problem is much harder[11]. Many examples of polynomial time equivalence were mentioned in Minsky's 1967 textbook [37], for example the fact that a RAM (random access model) is polynomially equivalent to a Turing machine. Thus by 1971 when Stephen Cook wrote his famous paper [11] showing the existence of NP-complete problems and thus reifying the importance of the concept of polynomial time, the notion was already well established.

5 Continuous mathematics meets computability

In order to simulate physics we have to have a theory of *computing with real numbers*. So: which parts of classical analysis may be made "effective?[12]"

Great strides were made in this area during the 1980s. I recommend reading the books by M.B.Pour-El and J.I.Richards [49] and Ker-I Ko [30], in which the foundations of real computation, and polynomial time real computation, are laid.

Ko [30] defines the computable reals as those reals x such that there exists an algorithm which, given an integer $n > 0$, outputs a rational x_n with $|x-x_n| < 2^{-n}$. The polynomial-time computable ("P.T.C.") reals are those for which the computation of x_n takes $\leq$polynomial(n) time. One may similarly define computable functions $f(x)$: given n and "black box" access to a subroutine for computing x_m, an algorithm will output a rational r_n with $|f(x) - r_n| < 2^{-n}$. P.T.C. functions are the same, except that the total amount of time to compute r_n (exclusive of the time spent in the subroutine; i.e. each call to the subroutine "costs" 1 time unit plus the number of symbols transmitted to and from that subroutine. The subroutine takes input in *unary* form.) has to be bounded by a polynomial function of n. Note that polynomial-time computable functions automatically must be continuous. Then we can discuss computable and P.T.C. *functionals*.

1. Consider the $\text{MAX}_{[0,1]}$ functional which maps a function $f : [0, 1]^2 \to \mathbf{R}$ to the function $g : [0, 1] \to \mathbf{R}$ defined by $g(x) = \max\{f(x, y)|0 \leq y \leq 1\}$. P=NP $\Leftrightarrow$ For all P.T.C. functions f, $\text{MAX}_{[0,1]}(f)$ is P.T.C.. Ko also

[9] It has been suggested to me that Robin O. Gandy (1919-1995), a student of Turing's [obituary: Modern Logic 6,1 (1996) 85-87] also made some contributions to physics+Church, but if so, I have been unable to track them down. Similar remarks apply to A.J.Wilkie.

[10] I would counter-argue that Davis's view is uninteresting because of its extreme unproductivity. His view completely trivializes computational physics, our entire lives, the whole universe, etc. as merely a "constant number of states, of no mathematical interest."

[11] Later S.Buss [6] re-proved this, proved NP-completeness, and also showed that even a nondeterministic Turing machine [which always "guesses right"] would require order $k/\log k$ steps. Later work on the same topic was in [3].

[12] "Effective" is a standard term in logic meaning "comes with an algorithm." For example, "$f(x)$ is continuous" and "there is a last prime of the form $2^k + k^2 + 5$" would be ineffective statements, but if there were an algorithm to compute an upper bound on that k, the second statement would become effective.

showed that there is a P.T.C. function $f : [0,1] \to \mathbf{R}$ on $[0,1]$ such that f's maximum need not even be computable and there can be an infinite number of maxima. All this built on a construction originally due to E.Specker[13] (1920-) [64].

2. Let INT be the operator mapping a function $f : [0,1] \to \mathbf{R}$ to $g(x) = \int_0^x f(t)\mathrm{d}t$. Then P=#P $\Leftrightarrow$ For all P.T.C. functions f, INT(f) is P.T.C.. (The "counting-complexity" class "#P" was defined in [67] [68].) This result is due to Harvey Friedman [22].

3. There exists a P.T.C. function $f : [0,1] \to \mathbf{R}$ such that $f''(x)$ exists, but $f'(x)$ but is not *computable*. But if $f''(x)$ exists and is *continuous* on $[0,1]$, then $f'(x)$ is P.T.C.. Thus if f is C^∞ on $[0,1]$, then $f^{(m)}(x)$ is P.T.C. for any fixed m. But there exist C^∞ P.T.C. f so that $f^{(m)}(x)$ is not computable to accuracy 2^{-n} in polynomial(n,m) steps. But if f is *analytic* on $[0,1]$, then $f^{(m)}(x)$ is computable to accuracy 2^{-n} in polynomial(n,m) steps. (Such results are also discussed in [49]. The first realization of the uncomputable nature of differentiation was by Myhill [41].)

4. If $f : [0,1] \to [0,1]$ is 1-to-1 and P.T.C. then f^{-1} is also. But P=LOGSPACE $\Leftrightarrow$ for all such f, f^{-1} is LOGSPACE-computable.

5. Let $f : [0,1]^2 \to [0,1]^2$ be 1-to-1 and P.T.C.. P=NP $\Rightarrow$ f^{-1} is P.T.C. $\Rightarrow$ P=UP. (For the "unique-P" complexity class UP, see [69].)

Some other results discussed in [30] include:

1. Effective versions of Weierstrass's polynomial approximation theorem (see also [49]). Example:
A function $f(\vec{x})$ mapping real vectors to reals is computable on some computable hyperrectangle if and only if there exists a computable sequence of polynomials converging *effectively* to it in the L_∞ norm on that hyperrectangle.

2. Results on ordinary differential equations. (See also [59].) For example Aberth [2] exhibited a computable function $f(x,y)$ which is uniformly continuous (but not Lipshitzian) in a rectangle containing $(0,0)$. (Ko improved this by providing a P.T.C. f.) *But* for this f the solution $y(t)$ of $\dot{y} = f(t,y)$ with $y(0) = 0$ is not a computable function.

3. The 2-dimensional variational problem of finding

$$\underset{g,h \in C[0,1]}{\text{infimum}} \int_0^1 \int_0^1 f(x,y,g(x),h(y))\mathrm{d}x\mathrm{d}y. \qquad (1)$$

For known P.T.C. f this number is computable in NEXPTIME (exponential time on a nondeterministic Turing machine) but is not computable in DEXPTIME unless certain complexity classes collapse [42].

Smith [61] recently solved an open problem of Ko's about "whether differentiability helps" in computing the maximum of a P.T.C. function on an interval. Smith argued that the answer is "no," but in order to get a theorem he had to *extend* the Ko theory of computable reals to include notions of codelength and/or transformation time for computer codes. The idea is that computable real numbers are actually represented, of course, by chunks of computer code. Example: If two P.T.C. reals are added, we get another P.T.C. real whose codelength is at most the codelength of the two summands, plus a constant. Smith's theorem was that either (1) The maximum value M of a P.T.C. C^∞ bounded function f on $[0,1]$ is not a P.T.C. real, or (2) it is P.T.C. but no polynomial time algorithm exists to transform P.T. code for f to get P.T. code for M, or (3) P=NP.

Pour-El and Richards [49] defined "computability structures on Banach spaces." Banach spaces are complete linear normed spaces. By means of "effective generating sets" (computable sequences of elements in the Banach space whose computable linear combinations are dense in the space) this notion is made computational. Then their most impressive result is that a closed linear operator (whose domain includes the effective generating set) is "computable" (meaning: maps computable sequences of elements in the Banach space to computable sequences of elements) if and only if the transformation is bounded. Myhill's result [41] that differentiation is not a computable operator follows instantly. They also gave as another example the wave equation in 3 dimensions, for which time-evolution is not a bounded operator in the L_2 norm, and hence maps initial data which is a computable function in the unit cube to an uncomputable function. But in an "energy norm," (Sobolev$_{2,2}$ norm) however, it is bounded and hence preserves computability.

Pour-El and Richards [49] also concerned themselves with the spectral structure of linear operators. They showed that each eigenvalue of an "effectively determined self-adjoint operator" (the Schrödinger Hamiltonians of physically realizable quantum systems qualify [27]) is a computable real number, i.e. for each eigenvalue there exists an algorithm to compute it to any demanded accuracy.

The entire *sequence* of spectral eigenvalues, however, is not necessarily computable, i.e. no finite-length algorithm necessarily exists to list *all* the eigenvalues. Nor is an eigenfunction corresponding to any given eigenvalue necessarily a computable function. Pour-El and Richards [49] exhibited operators for which both of these things were, in fact, uncomputable.

But Smith's "Church's thesis meets quantum mechanics" paper [60] shows that operators corresponding to *physically realizable* Hamiltonians (at least, in the quantum mechanics he permits) are nicer: a finite-length algorithm *does* exist to list all the eigenvalues. Some eigenfunction might still be uncomputable, though (this question remains open).

[13] Specker, incidentally, was a co-author on a remarkable paper [31] showing that no "noncontextual hidden variable theory" could emulate the predictions of quantum mechanics. Specker was familiar with, and indeed first popularized, the notion that quantum "logic gate" devices could exist which would implement a logic wider than the usual boolean logic. That thinking could be viewed as a precursor to the notion of quantum computers. An annotated bibliography of Specker's work 1949-1979 may be found in Enseign. Math. (ser. 2) 27,1-2 (1981) 85-98.

6 The development of rigorous physics

"Rigor" is an approach in which the goal is to make statements, called "theorems," in which every term used has an unambiguous definition, and in which an explicit list of assumptions implies an explicit claim. This has a "proof" which is an explicit gapless chain of logical implications, each of which themselves is a theorem, an assumption, or a logical axiom.

Rigor is necessary in order to merge computational complexity theory with physics.

Over the course of the 20th century, great strides have been made in making quantum mechanics a rigorous part of mathematics. Probably the greatest hero in this battle has been Tosio Kato (1917-). Most physicists have never heard of Kato. Nevertheless, during the 1950s Kato proved, essentially, existence and uniqueness of the solutions of Schrödinger's equation in a wide class of potentials (including all the ones that seem physically realizable) – and more [27] [28]. Kato's are undoubtably the biggest theorems in quantum mechanics.

These feats were made possible by the gradual development of more and more tools in the theory of linear operators, which has also been called "functional analysis" and "the analysis of the 1970s."

The late 20th century also saw the enshrining of this theory as applied to quantum mechanics in the 4-volume tome [54] by Reed and Simon, and the establishment of the journals *Journal of Mathematical Physics* and *Communications in Mathematical Physics* as showcases for work of this nature.

I have found the work of E.H.Lieb (1932-) [35] to be particularly helpful. Lieb has throughout his career set high standards of rigor and simplicity in the doing of physics. One of his research themes has been proving the "stability of matter." Although mentioning that to many physicists tends to result in a somewhat blank or disgusted look as a response (!), really this question strikes at the heart of the rigorous underpinnings of the subject, and needs to be understood before much further progress becomes possible.

One of Lieb's accomplishments has been to make Thomas-Fermi theory (which one can think of as an inspired approximate treatment of quantum mechanics) rigorous. What was previously just a heuristic approximation with no theoretical understanding whatever attached saying when, why, and how well it worked, is now known to provide rigorous lower bounds on spectral energies and to (in certain common asymptotic situations) provide answers that are asymptotically correct. This is an example of the difference that theorems can make. Before theorems appear: there is no way to compare nonrigorous ideas that float around except by methods such as the "reputations" of their authors, and computational experience (all of which, really, is just anecdotal). After theorems appear: true understanding dawns.

Despite all this progress, still, very little of physics has much rigor attached to it. Even in 1999, most physicists are unaware of and/or disparage rigorous reasoning[14]. For example, I work for the 10-year old NEC Research Institute, which employs about 65 physicists, both theoretical and experimental. Among the many physics papers and technical reports they authored during these years, I do not think there has been a single theorem/proof pair in them, nor a single citation of any paper in *J.Math.Phys.* or *Commun.Math.Phys.* If one looks through those papers (which are by no means atypical) it is rare that one sees either an explicit list of assumptions, or a concise self-contained statement summarizing the results of their reasoning; these are the two most important ingredients of a theorem.

Let us now abandon socio-cultural issues and consider physical theories themselves. Consider the BCS theory (as physicists call it; I would prefer the word "model") which, to great acclaim, "explained" superconductivity. While I can only applaud whatever empirical success that theory has had, I am unable to find any theorems there[15].

The issues of "measurement" and "decoherence" lie at the foundation of quantum mechanics, but in my opinion all treatments of them so far have not managed to be simultaneously rigorous and useful.

The rigor situation in theories of physics which go beyond garden-variety quantum mechanics (i.e. roughly corresponding to all the topics taught to physics students after the second year in the usual undergraduate curricula as of 1999) is far worse and stands in considerable contrast to some of the striking empirical successes of some such theories, such as QED (quantum electrodynamics) [26] [7] [52].

QED, is, at the core, an astoundingly simple theory [20]. It also yields, in some cases, predictions that agree with experiment to very high accuracy. For example: Consider the ionization energy of Helium-4 (expressed as a frequency). The experimental values are
5945204356 ± 48 MHz [5] and 5945204236 ± 45 MHz [19]. The theoretical value, based on a numerical solution of Schrödinger's equation, plus corrections for the non-infinite mass of the nucleus of order 143 GHz, plus relativistic corrections of order 17 GHz, plus QED ("Lamb shift") corrections (-41232.8 ± 91.1 MHz), plus corrections for the nonzero RMS radius ($1.673 \pm .001$fm) of the nuclear charge distribution ($+29.593 \pm .018$ MHz) is

[14] An experience of my own is, while not terribly important, rather illuminating. I was looking for a rigorous formulation of the energy-time uncertainty principle. Although every physicist "knows" this principle in some form such as "$\Delta E \Delta t \gtrsim \hbar/2$," I was unable to find even one who knew what Δt and ΔE meant exactly, and who could state (or even knew of anybody who had stated) the principle as an *inequality* in which the quantities on both sides of the inequality were *precisely defined.* This was 1998. Finally I did the job myself [62].

[15] Actually, although nobody currently seems capable of making a theorem of the form "the laws of quantum mechanics imply X" where X is some statement about superconductivity, I would claim people *are* capable of making theorems of the form "the laws of quantum mechanics, plus assumptions A, B, C,... imply X." Indeed, many papers are (in essence) doing just that. In that case, the physicists writing that paper should have given an explicit theorem statement and an explicit list of assumptions. But the usual method they adopt is *not* to give a precise statement nor to list assumptions explicitly (and indeed disguise the fact that they even exist). I believe this policy impedes further progress, while the policy I am suggesting would aid it.

5945204225.9 ± 91.1 MHz [16].
The disagreement between theory and experiment is comparable to the disagreements between the two experiments and to the theoretical estimated error. The total accuracy is of order 2 parts in 10^8.

Despite the high accuracies often achieved by such series of smaller and smaller perturbative corrections, a brilliantly simple argument of Dyson[16] made it extremely plausible [17] that the usual prescription (Feynman diagrams and integrals) for calculating quantities in QED always yields *divergent* series. But another mysterious prescription called "renormalization" will (it is now generally agreed) always succeed in producing a unique finite value for each term in each such series. Renormalization is a technique for "removing infinities" by simply replacing them with the known values of the electron mass and charge (quantities which, if calculated in unrenormalized QED, would both be ∞). Thus QED is a semi-empirical theory and is so in a very peculiar sense. Some physicists, such as Dirac, never accepted QED, but (to a good approximation) they all died [56].

QCD [72] [52] [7] is mathematically very similar to QED, but it is more complicated and the calculations do not converge as quickly (or rather, diverge more quickly) making it extremely hard to get numbers out of it. Experimental confirmation of QCD is (so far) therefore much less dramatic than it is for QED.

Anyway, to cut a long story brutally short, as far as I can tell, QED, QCD and general relativity all are very beautiful theories but also all suffer from a severe absence of theorems. In the case of QED and QCD, well known foundational problems exist. In the case of QCD, the theory also suffers from an absence of algorithms to compute predictions about nature (actually all of them suffer from this, but the problem is most dramatic for QCD).

What is to be done about this? In my opinion there is no easy way out. Proving theorems is hard work. But I think the situation would be considerably improved if physicist-authors did not make it a common *policy* to try to delude their readers into thinking they have proved a theorem, when in fact they have not[17]. I have no objection to plausibility arguments and conjectures in physics, and nor should anyone – but I object when an effort is made to mislabel them as theorems. Such mislabeling misleads people and slows and prevents further progress in the field. If instead, gaps in rigorous reasoning were *highlighted* by their authors as such, this would *aid* further progress.

7 Results emerge about physics and complexity

It is my view that in some sets of "laws of physics," and with explicit definitions of vague terms such as "simulate," "reasonable physical system," and so forth, CT need no longer be a thesis; it can be a theorem. CT can be proved true in some physical theories, whereas in others it can be proved false; and investigating which, is a matter of considerable interest. Disproofs can be gotten by exhibiting physical systems which do uncomputable things. Proofs can be gotten by exhibiting simulation algorithms.

I showed [59] in 1993 that there was a sense in which CT was *false* for N point masses obeying Newton's laws of gravity and motion. Specifically, initial states of the masses, having bounded energy and mass, exist such that uncountably many topologically distinct trajectories are possible in 1 minute. In contrast, a Turing machine can only perform a finite amount of computation in finite time, (or even in infinite time only a countably infinite amount). Hence, in these laws of physics, unsimulable physical systems exist. This demonstration leaned heavily on work of J.Gerver [24] on "noncollision singularities." On the other hand with slightly modified more realistic versions of Newton's laws (more closely approximating the demands of special and general relativity – speed limits and Schwarzschild radii are imposed) I proved Church's *strong* thesis by explicitly exhibiting a simulation algorithm with polynomial slowdown. The polynomial has as arguments the timespan of the physics being simulated, the masses, their energies, and the number of decimal places of accuracy demanded. (Indeed it was shown in [59] that a wide class of differential equations are simulable with "polynomial slowdown.")

My Church disproof shows the existence of unsimulable systems, but it arguably does *not* show how to *build* a physical system (assuming these laws of physics held) that can do something uncomputable. That is because apparently (although this has never been proved) the initial conditions leading to unsimulability are of measure zero in the space of all possible initial conditions. In that case, achieving them in practice is impossible.

The moral of this is that the intuitive notion of "simulability" is not identical to the intuitive notion of "computational equivalence." In unsimulable cases, any Turing machine trying to simulate a finite timespan of the motion would have to keep coming back forever to its input tapes for more decimals of precision about the initial conditions. (We assume the input tapes contain the entire infi-

[16] Freeman J. Dyson (1923-), one of the premier rigorous physicists of this century, and a major player in the invention [56] of QED.

[17] The most usual way in which this happens, is this. A physics paper, posing a *purely mathematical* problem appears; and in its abstract it says "We show [apparent solution to the problem]." But no theorem is present, the authors do not know how to make one, and their solution may or may not be correct – it is often very difficult to decide. In such cases it would have been better to say "We conjecture [solution] and support this conjecture with some reasoning which contains gaps, which we highlight, and which we fill with heuristic plausibility arguments." I won't give examples of that pathology, because it would require several pages of analysis of somebody's paper to make my point, none of it relevant to the present paper. But I will give two examples in which the original authors themselves (both Nobelists) were (comparatively) innocent since the mislabeling was performed by subsequent authors. Many authors [51] have referred to "Onsager's reciprocal relations" in statistical mechanics as "Onsager's theorem." It is no theorem. Nor did Onsager ever claim it was. (Indeed, I have recently found counterexamples!) Many authors [15] have also referred to "Anderson's theorem." Anderson himself has confirmed to me that he never claimed it was a theorem, and nobody has made it one, either. In my opinion, Anderson's original paper [4] would have been improved if he had tried to make his result more theorem-like; for example, if he had made an explicit list of his assumptions and if he had highlighted the gaps in his reasoning which prevent the final result from being a theorem.

nite decimal expansion of the initial masses, velocities, and positions.) But in simulable cases, only a finite number of decimals would need to be read off the input tapes before it would become possible to answer any typical question such as "does body 5 pass within the unit circle?" about the motion in the next minute.

Christopher Moore [40] showed that a point particle bouncing between a finite number of perfect parabolic and planar mirrors can simulate a Turing machine, and hence has undecidable long term behavior. Moore then observed that the mirrors could be "softened" so that the long term behavior of a particle moving in a certain smooth and finitely describable potential in 3D, is undecidable.

Peter Shor, in an astounding development [57], showed that a "quantum computer" could factor N-digit integers into primes and solve discrete logarithm problems modulo N-digit primes, in a number of "quantum gate operations" only polynomial in N. This strongly suggested that Church's "strong" (polynomial time) thesis was false in any model of physics permitting such quantum computers. (It still remains unproved that quantum polynomial time is a larger set than conventional polynomial time, but Shor's result makes it likely.) The notion that such quantum computers might in fact be buildable was made much more plausible by the subsequent discovery (in which Shor also played a large role) of quantum error correcting codes and quantum fault-tolerance techniques. (These are reviewed in [50].)

In my opinion, though, the question of whether quantum computers are buildable remains open both experimentally and also in a theoretical sense. Experimentally, so far the best that has been accomplished is (roughly) to build a low reliability "quantum logic gate" with about 1.5 inputs (2 inputs, but one of the inputs was partially simulated) and to cool two quantum bits to the $|00\rangle$ state [38] [39]. Theoretically, I believe that one should analyse quantum computers in models of *open* quantum mechanics which include "decoherence" [73]. Although this has been done in some ultra-simplified toy models of decoherence, I would contend that really, true understanding has not been reached about the true nature of and about the ultimate limits on, decoherence. (For example, in some models of decoherence [66], correlated qubits decohere faster the further apart they are [at a rate proportional to the square of the separation, in fact]. Decoherence with this faster-further property causes all quantum computer schemes to break.) So far no theorems have been proved within the maximally realistic decoherence models [21] [71] to say that quantum computers are indeed buildable.

In 1999 I formulated and proved [60] a version of Church's weak thesis in the simplest kind of quantum mechanics (nonrelativistic, no quantum fields, Schrödinger's multiparticle equation, not-too-singular scalar potentials only). Remarkably, this proof yields the strong CT if the simulator is a quantum computer, and hence the "intermediate" CT if the simulator is a conventional computer.

References

[1] Oliver Aberth: Analysis in the computable real number field, J. Assoc. Comput. Machinery 15 (1968) 275-290 (MR 38 #5626).

[2] O. Aberth: The failure in computable analysis of a classical existence theorem for Ordinary Differential Equations, Proc. Amer. Math. Soc. 30 (1971) 151-156.

[3] Misha Alehknovich, Samuel R. Buss, Shlomo Moran, Toniann Pitassi: Minimum Propositional Proof Length is NP-Hard to Linearly Approximate (Extended Abstract), in *Mathematical Foundations of Computer Science* (1998) 176-184.

[4] P.W.Anderson: Theory of dirty superconductors, J.Phys.Chem.Solids 11,1/2 (1959) 26-30.

[5] S.D.Bergeron et al.: Measurement of the He Ground state Lamb shift via the two-photon $1^1S - 2^1S$ transition, Phys.Rev.Lett. 80,16 (1998) 3475-3478.

[6] Sam R. Buss: On Godel's Theorems on Lengths of Proofs II: Lower Bounds for Recognizing k Symbol Provability, pp.57-90 in (P.Clote & J.Remmel eds.) Feasible mathematics II. Birkhauser, 1995.

[7] Ta-Pei Cheng & Ling-Fong Li: Gauge theory of elementary particle physics, Oxford University Press, 1984; reprinted 1988, 1991.

[8] Demetri Christodoulou & Sergei Klainerman: The global nonlinear stability of the Minkowski metric, Princeton Univ. Press 1993.

[9] Alonzo Church: An unsolvable problem of elementary number theory, Amer. J. Math. 58 (1936) 345-363; A note on the Entscheidungsproblem, J. Symbolic Logic 1 (1936) 40-41 & Correction 101-102.

[10] A.Cobham: pp. 24-30 in Proc. 1964 int'l conf. for logic, methodology, and philosophy of science (Y.Bar-Hillel ed.) North-Holland 1965.

[11] S.Cook: The complexity of theorem proving procedures, Proc. 3rd ACM Sympos. on Computing (1971) 151-158.

[12] Martin Davis: Computability & Unsolvability, McGraw-Hill 1958; Dover reprint.

[13] Martin Davis: Why Gödel Didn't Have Church's Thesis, Information & Control 54 (1982) 3-24.

[14] M.Davis (ed.): The Undecidable, Raven Press 1965.

[15] P-G. de Gennes: Superconductivity of metals and alloys (translated by P.A. Pincus), Addison-Wesley 1989.

[16] G.W.F.Drake & W.C.Martin: Ionization energies and quantum electrodynamics effects in the lower levels of neutral Helium (He-4) Canad.J.Phys. 76,9 (1998) 679-698.

[17] F.J.Dyson: Divergence of perturbation theory in quantum electrodynamics, Phys. Rev. 85,4 (1952) 631-632.

[18] J.Edmonds: Paths, trees, and flowers, Canadian J. Math. 17 (1965) 449-467.

[19] K.S.E.Eikema et al.: Phys.Rev. A 55,3 (1997) 1866-1884

[20] Richard P. Feynman: QED, the strange theory of light and matter, Princeton University Press, 1985. (Popular exposition.)

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

[22] Harvey Friedman: On the computational complexity of maximization and integration, Advances in Math. 53 (1984) 80-98.

[23] G.Galdi: An introduction to the mathematical theory of the Navier-Stokes equations (vols. 1-4), Springer 1984-.

[24] Joseph L. Gerver: The existence of pseudocollisions in the plane, J. Differential Equations 89 (1991) 1-68.

[25] Kurt Godel: über formal unentscheidare sätze der Principia Mathematica und verwandter Systeme I, Monatshefte für Math. und Physik 38 (1931) 173-198.

[26] J. M. Jauch & F. Rohrlich: The theory of photons and electrons, Springer-Verlag (2d expanded ed.) 1976.

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

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

[29] S.C.Kleene: Introduction to metamathematics, Van Nostrand 1952, North Holland 1951.

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

[31] S. Kochen & E.P. Specker: The problem of hidden variables in quantum mechanics, J. Math. Mechan. 17 (1967) 59-87. A simplification of this proof is in A.Peres: Two simple proofs of the Kochen-Specker theorem, J.Phys. A (math. gen.) 24,4 (1991) L175-L178.

[32] Georg Kreisel: Mathematical logic, pp. 95-195 in "Lectures in modern Mathematics" (T.Saaty, ed., vol 3) 1965.

[33] Georg Kreisel: A notion of mechanistic theory, Synthese 29 (1974) 11-26.

[34] P.Odifreddi (ed.): Kreseliana: About and Around Georg Kreisel, A.K.Peters, Wellesly MA, 1996.

[35] The stability of matter, Selecta of Elliott H.Lieb (ed. W.Thirring) Springer (2nd ed.) 1997.

[36] Yuri V. Matiyasevich: Hilbert's tenth problem, MIT press 1993.

[37] Marvin Lee Minsky: Computation: finite and infinite machines, Prentice-Hall 1967.

[38] C.Monroe, D.M.Meekhof, B.E.King, W.M.Itano, D.J.Wineland: Demonstration of a fundamental quantum logic gate, Phys. Rev. Lett. 75,25 (1995) 4714-4717

[39] C. Monroe, D.M. Meekhof, B.E. King, S.R. Jefferts, W.M. Itano, D.J. Wineland, and P. Gould: Resolved-sideband Raman cooling of a bound atom to the 3D zero-point energy, Phys. Rev. Lett. 75 (1995) 4011-4014.

[40] C.Moore: Unpredictability and undecidability in dynamical systems, Physical Review Letters, 64,20 (1990) 2354-2357.

[41] J. R. Myhill: A recursive function defined on a compact interval and having a continuous derivative that is not recursive, Michigan Math. J. 18 (1971) 97-98.

[42] C.H.Papadimitriou & J.Tsitsiklis: Intractable problems in control theory, SIAM J. Contr. Optim. 24 (1986) 639-654.

[43] Emil Post: Finite combinatory processes, J.Symbolic Logic 1 (1936) 103-105.

[44] M.B.Pour-El: Abstract computability and its relation to the general purpose analog computer, Trans. AMS 199 (1974) 1-28.

[45] M.B.Pour-El & J.I.Richards: A computable ordinary differential equation which possesses no computable solution, Annals of Mathematical Logic 17 (1979) 61-90.

[46] Marian Boykan Pour-El & J.Ian Richards: The Wave Equation with Computable Initial Data Such That Its Unique Solution Is Not Computable, Advances in Mathematics 39 (1981) 215-239.

[47] M.B.Pour-El & J.I.Richards: Noncomputability in analysis and physics: a complete determination of the class of noncomputable linear operators, Advances in Mathematics 48 (1983) 44-73.

[48] M.B.Pour-El & J.I.Richards: Computability and Noncomputability in classical analysis, Trans. Amer. Math. Soc. 275 (1983) 539-560.

[49] 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. Verien. 94 (1992) 53-55.)

[50] 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.")

[51] I.Prigogine: Introduction to Thermodynamics of irreversible processes, Wiley 1967.

[52] C.Quigg: Gauge theories of the strong weak and electromagnetic interactions, Benjamin/Cummings 1983.

[53] M.O.Rabin: Degree of difficulty of computing a function and a partial ordering of recursive sets, TR#2 Hebrew University 1960; Mathematical theory of automata, Proc. 19th Sympos. Applied Math, (1966) 153-175.

[54] Michael Reed & Barry Simon: Methods of modern mathematical physics, (4 vols.) Academic Press 1972-9; rev. & enl. ed. 1980-

[55] R.Rosen: Church's thesis and its relation to the concept of realizability on biology and physics, Bull. Math. Biol. 24 (1962) 375-394.

[56] S.S.Schweber: QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga, Princeton University Press, 1994.

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

[58] M.Sipser: The history & status of the P versus NP question, Symp. Th. of Comput. Sci. 24 (1992) 603-618.

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

[60] Warren D. Smith: Church's thesis meets quantum mechanics, NECI tech report dated 02/26/99.

[61] Warren D. Smith: Finding the maximum of a polynomial time computable bounded smooth function on an interval is NP-complete, NECI tech report dated 02/05/99.

[62] Warren D. Smith: The energy time uncertainty principle, NECI tech report dated 09/02/98. All these NECI tech. reports available electronically at `http://www.neci.nj.nec.com/homepages/wds/works.html`.

[63] Warren D. Smith: Plane mechanisms and the "downhill principle" Lecture notes 1998. Available electronically at `www.neci.nj.nec.com/homepages/wds/pu-planelink.ps` and `pu-hctable.ps`.

[64] Ernst P. Specker: Der satz vom Maximum in der rekursiven Analysis, pp. 254-267 in "Constructivity in Math." (ed. A.Heyting) Stud.Logic Found. Math. North-Holland Amsterdam 1959.

[65] Alan Turing: On computable numbers, with an application to the Entscheidungsproblem, Proceedings of the London Mathematical Society, 42,2 (Dec 1936) 230-265; A correction, 43 (1937) 544-546.

[66] W.G. Unruh and W.H. Zurek: Reduction of a wave packet in quantum brownian motion, Phys. Rev. D 40,4 (1989) 1071-1094

[67] Leslie G. Valiant: The complexity of enumeration and reliability problems, SIAM J. Comput. 8,3 (1979) 410-421.

[68] Leslie G. Valiant: The complexity of computing the permanent, Theor. Comput. Sci. 8,2 (1979) 189-201.

[69] Leslie G. Valiant & Vijay V. Vazirani: NP is as easy as detecting unique solutions, Theor. Comput. Sci. 47 (1986) 85-93.

[70] A. Vergis, K. Steiglitz, and B. Dickinson: The complexity of analog computation, Math & Computers in Simulation 28 (1986) 91-113.

[71] U.Weiss: Quantum dissipative systems, World Scientific 1993. Series title: Series in modern condensed matter physics ; vol. 2.

[72] F.J.Yndurain: The theory of quark and gluon interactions, Springer (2nd rev. & enl. ed.) 1983.

[73] W.H.Zurek: Decoherence and the transition from quantum to classical, Physics Today (October 1991) 36-44; also see April 1993 issue, 82-90.