Topological Views on Computational Complexity

Michael H. Freedman*

1991 Mathematics Subject Classification: 57-XX Manifolds and cell complexes; 68-XX Computer Science; 81-XX Quantum Theory
Keywords and Phrases: computational complexity, topology, quantum field theory

For the pure mathematician the boundary that Gödel delineated between decidable and undecidable, recursive and nonrecursive, has an attractive sharpness that declares itself as a phenomenon of absolutes. In contrast, the complexity classes of computer science, for example P and NP, require an asymptotic formulation, and like the subject of "coarse geometry", demand a bit of patience before their fundamental character is appreciated.

The heart of the matter is to understand which problems can be solved by an algorithm whose "running time" grows only polynomially with the size of the instance. It is interesting to note that in other areas of mathematics things polynomial tend to have excellent limiting behavior: 1. Any polynomial on cardinals: $x \longmapsto \text{poly}(x)$ is continuous at the first infinite cardinal, whereas the power set function $x \longmapsto 2^x$ is not. 2. In complex analysis, polynomials extend conformally over infinity to yield a branch point, whereas exp is essentially discontinuous at infinity. 3. In coarse geometry, groups with polynomial growth, in common with nilpotent Lie groups, have Carnot manifolds as scaling limits (Gromov [G]) in the Gromov-Hausdorf topology. These examples, particularly the last, suggest that polynomial time algorithms might eventually be understood by constructing a more manageable limiting object as polynomial growth groups are understood via nilpotent Lie groups.

In order to make the discussion of algorithms precise, it is necessary to define a computational model. This is more exciting now than it was ten years ago. The "polynomial Church thesis" is up in the air, and there are two robust computational models to sink one's teeth into: the "Turing model" and "Quantum Computing" (QC). (See http://xxx.lanl.gov/abs/quant-ph and [K] for a suggested solid state implementation based on the hyperfine coupling between electron spin and nuclear spin.) Furthermore it is possible that there will be other, perhaps stronger, computational models based on topological quantum field theory [F1].

The thesis of Alonzo Church, propounded in the mid-1940s, asserts that any two definitions of "computable function" will agree. The "polynomial version" of the Church thesis (although I do not know that it was ever endorsed by Church) says that any two physically *reasonable* models of computation will agree on the

*This work was supported by Microsoft Research.

class of polynomial time functions (but not necessarily on the degree of the polynomial, which may in fact be model dependent). "Reasonableness" implies limited accuracy in preparation and measurement of physical states.

It might seem that if one accepts that the universe is fundamentally quantum mechanical (and I am perfectly prepared to neglect the irreversibility of black hole evaporation) that QC is the ultimate model, and no others need be sought. This argument is not entirely convincing, since a solid state system (perhaps one involving global excitation as occurs in the fractional quantum Hall effect) might be governed to considerable accuracy by an effective field theory whose simulation through local QC gates involves exponential inefficiencies. (Note that a preliminary discussion of simulating local Hamiltonians by gates is given in [L].) Topological field theories, because of their discrete character and their connections to NP-hard (actually $\#P$-hard) combinatorial problems, e.g., the evaluation of the Jones polynomials, are the most interesting candidates for further computational models [F1]. The next section contains definitions, but briefly, the class NP, nondeterministic polynomial time, consists of those decision problems where the time to "check" (rather than find) a proposed solution grows only polynomially in the length of the problem instance.

In pure mathematics, problems of fundamental importance occasionally arrive on our doorstep from physics. The only other cases (i.e., origin outside of physics) I can think of are: probability (gambling), crystallographic groups (chemistry), incompleteness (philosophy), and the P/NP problem (computer science). A proof that $P \neq NP$ would be extraordinarily strong, as it would foreclose the possibility of myriad yet-unimagined theories that might connect, say, the colorings of a graph (which is NP complete) and, say, the cohomology of some associated space (which might well be in P as cohomology is essentially linear algebra). These speculations might suggest that the P/NP problem is undecidable. In a platonic world view, where statements of first order arithmetic, such as "$P = NP$", are either true or false, there are two subcases: the very interesting Case (1): undecidable and true, in which case the NP problems *do* admit P-time algorithms, but there is no *documentation* proving they work; and the less interesting Case(2): undecidable and false: there are no P-time algorithms for the NP-complete problems, but there is no proof of this statement.

The assertion that a problem is important to mathematics is usually supported by sketching its relations to other problems and fields. The P/NP problem enjoys a more interesting status. The practice of mathematics is largely the search for proofs of reasonable length (certainly polynomial in statement length) and so is inside NP. Setting aside the constraints of any particular computational model, the creation of a physical device capable of brutally solving NP problems would have the broadest consequences. Among its minor applications it would supersede intelligent, even artificially intelligent, proof finding with an omniscience not possessing or needing understanding. Whether such a device is possible or even in principle consistent with physical law, is a great problem for the next century.

§1. Preliminaries.

The Turing model of computation consists in a bare formulation of a bi-infinite tape, a head which can read/write symbols from a finite alphabet and which is capable itself of being one of finitely-many *internal* states. Its "program" is a finite set of 5-tuples $\{S, q, S', q', M\}$ which say that if it is in state S and reads q, it will assume state S', overwrite q with q', and move right or left according to the indicated motion M. We can absorb knowledge of the last motion into the state q' and so drop the fifth symbol. Without an applicable instruction the machine halts. The internal state, the head position and the contents of the tape together, form the machine's *complete state*. For convenience, one or more additional tapes may be added to the machine, generally decreasing computation time, but by no more than a square root factor. All conventional computers are implementations of the Turing model.

A next step is to allow probabilistic computation where several 4-tuples may begin "S, q", and these will be assigned positive weights p_i summing to one and will be executed with probability p_i, so that the machine now evolves stochastically through a *mixture* of states. Empirically, it is often easier to find probabilistic algorithms that almost always work, than to find traditional exact algorithms.

A further, more radical, innovation is to allow the weights above, now written $w_{\alpha\beta}$, to be complex numbers satisfying $\sum w_{\alpha\beta}\overline{w}_{\beta\gamma} = \delta_{\alpha\gamma}$, where $w_{\alpha\beta}$ is the transition amplitude for $(S,q) = (S,q)_\alpha \to (S,q)_\beta = (S',q')$. The resulting evolution of the computation is now a unitary evolution $U(t)$ in a vector space of complete states. This, briefly, is the model called *quantum computation* or QC. It is an important consequence of this description that the evolution is *local* at any time t: The t^{th} step, or *gate*, in the time evolution $U(t)$ is the identity except on a tensor factor of bounded dimension (typically $\mathbb{C}^4$ or $\mathbb{C}^8$ in detailed specifications).

In the Turing model P represents the class of decision problems $\{D\}$ (answer $\in$ $\{\text{yes, no}\}$) so that there is a program F_D and a polynomial P_D with F_D yielding the answer to each instance I of D in time $\leq P_D(\text{length}\, I)$, where length I is the number of bits required to express I. One says D lies in NP (nondeterministic polynomial time) if there is an *existential* program operating on I plus a number of *guess bits* which correctly answer all instances in polynomial time. The existential program is deemed to answer "yes", if for some setting of the guess bits the machine halts on the symbol 1. The fundamental question of computer science is to show that $P \neq NP$, essentially that it is harder to find a solution than to check a guess.

The problem of the existence of a satisfying assignment for a Boolean formula is the canonical NP-*complete* problem, meaning[1] it lies in NP and if a Turing machine were augmented by an *oracle* capable of (quickly) answering that one problem, then all problems in NP could be solved in polynomial time. (A problem is called "hard" rather than complete if only the second assertion is being made.) The class $\#P$ is the counting analog of NP; computing the number of satisfying assignments of a Boolean formula is the canonical $\#P$-complete problem. In oracle notation $P^{NP} \subset P^{\#P}$, meaning a poly-time machine with access to an $\#P$ oracle is at least as powerful as one with access to an NP oracle.

[1] according to Cook [C]

The model QC is not strictly comparable with Turing, since in QC the output, a *measurement* of a final stationary state, is only probabilistic. However it is believed, because of Shor's QC algorithm [Sh] for factoring integers in polytime, that QC is substantially more powerful than P but perhaps *not* powerful enough to solve NP-complete problems in polynomial time. Computational models that are allowed to handle continuous quantities are almost always absurdly strong (e.g., contain NP), if accuracy is not restricted to poly(log) number of bits. ([Sc], [ADH])

On the topological side, the notion of a topological quantum field theory has emerged through Witten's work. A TQFT is usually understood to be a functor Z from (oriented marked surface, a bounding oriented 3-manifold with link; diffeomorphisms)[2] to (finite-dimensional Hilbert spaces over $\mathbb{C}$, vector; linear maps) which satisfies $Z(\Sigma_1 \cup \Sigma_2) = Z(\Sigma_1) \otimes Z(\Sigma_2)$, $Z(\overline{\Sigma}) = Z(\Sigma)^*$, a gluing axiom (gluing bordism corresponds to composing linear maps), and a unitarity axiom. (See [At] for details.) In particular such a theory assigns scalars to closed three-manifolds containing a link L, and Witten identified one such theory W_k, $SU(2)$-Chern-Simons theory at level k, as a value of the Jones [Jo] polynomial V,

$$W_k(L) = V_L(\zeta) \quad , \quad \zeta = e^{\frac{2\pi i}{k+2}} \quad . \tag{1}$$

Since we will be discussing the utility of this TQFT for solving combinatorial problems such as Boolean satisfiability, it is relevant to observe that counting satisfactions, colorings, and many other combinatorial problems provide by far the simplest examples of systems obeying the TQFT axioms; only the source category must be redefined. (It is tempting to look for the corresponding path-integral interpretations.) To see, for example, how the gluing axiom works for the problem of counting vertex colorings of a graph, let $(G_1; H_1, H_2)$ and $(G_2; H_2, H_3)$ be disjoint finite graphs, each with two preferred disjoint subgraphs where $H_2 \subset G_1$ and $H_2 \subset G_2$ are identified by a fixed isomorphism. Let $G = G_1 \cup_{H_2} G_2$. Let i, j and k index the possible legal colorings of H_1, H_2 and H_3 respectively and let $m_{i,j}$ ($n_{j,k}$) be the number of colorings of G_1 restricting i on H_1 and j on H_2 (j on H_2 and k on H_3). Then the number of colorings $g_{i,k}$ of G which restrict to i on H_1 and to k on H_3 satisfies the composition rule:

$$g_{ik} = \sum_j m_{i,j} n_{j,k}$$

§2. WHAT A TOPOLOGIST MIGHT THINK ABOUT FORMAL SYSTEMS.

A. UNDERSTANDING THE CLASS P:

CONJECTURE: The class of P-time algorithms can be elucidated by constructing some scaling limit as discussed in the introduction. (Also see [F2])

B. GENERAL POSITION IN FORMAL SYSTEMS:

There is an empirical connection between computational complexity of a finite decision problem and the undecidability of an infinitary version [F3]. Although

[2] perhaps with additional structures or labelings

oracle separation results [BGS] show that detailed properties of P must enter the proof, P might be distinguished from NP by finding a translation which carries P into decidable statements. Thus it is natural to ask how common decidable statements are. Let X be a formal system subject to Gödel's second incompleteness theorem, such as Peano arithmetic or ZFC. Let $\{S_i \mid i \in Z^+\}$ be the sentences of X enumerated in some syntactically natural way, e.g., alphabetical order. Consider those sentences which are provable (in X or some fixed finite extension X^+ of X) and let p_i denote the number of these with index $\leq i$ which are provable.

CONJECTURE: "Ubiquity of undecidability" $\operatorname{limsup}(p_i \,/\, \sqrt{i}) = 0$. That is, the "number" of provable statements is less than the square root of the number of statements.

According to Kolmogorov and later Chaitin [Ch] at least half of integers fail to admit short descriptions, but particular true instances of the statement "n has no short description" are always undecidable. This provides a fairly large natural family of undecidable statements, but not the conjectured ubiquity of undecidable statements.

RATIONALE FOR CONJECTURE: The single most useful principle in geometric topology is that submanifolds $P^p, Q^q \subset M^n$ contained in a manifold, generically satisfy $\dim(P \cap Q) = p + q - n$. For finite sets, if P and Q are drawn randomly from M, the same formula holds:

$$\text{expected value of } \log \operatorname{card}(P \cap Q) = \log \operatorname{card}(P) + \log \operatorname{card}(Q) - \log \operatorname{card}(M). \tag{2}$$

In particular two disjoint subsets P and P' of equal cardinality should satisfy:

$$\operatorname{card}(P) = \operatorname{card}(P') < \sqrt{\operatorname{card} M} \tag{3}$$

if their disjointness is simply a matter of chance.

If X is consistent, then provable statements Q and their negations Q' are disjoint. We believe that in a system complex enough to be incomplete, the global structure of Q inside all statements is essentially random and so expect Q to be asymptotically of less than square root size. This is analogous to thinking that the primes are "randomly" distributed in the integers according to the density $\frac{1}{\log n}$, a model with considerable predictive power.[3]

C. $P \neq NP$ HAS PREDICTIVE POWER IN LINK THEORY:

In computer science the notions of width arise in identifying subclasses of NP-hard problems which are actually solvable in P-time. Among these are problems of constant width, or even polylog width problems. (Compare page 95 [We].) For the present purpose define the width of a link L to be the $\inf_{\pi} \sup_{r} \left| L \cap \pi^{-1} r \right|$, where π is a smooth product projection $\mathbb{R}^3 \to \mathbb{R}$, $r \in \mathbb{R}$. Let $\mathcal{L}$ be the set of finite links. Call a mapping $i : \mathcal{L} \to \mathcal{L}$ *information preserving*, if some $\#P$-hard data

[3]For example, the density of double primes seems to be correctly predicted, up to a small multiplicative constant, from this assumption.

about $L \in \mathcal{L}$ can be quickly computed from data about $i(L)$, e.g., if $V_L(e^{2\pi i/5})$ is quickly (P-time) calculable from $V_{i(L)}(e^{2\pi i/5})$.

CONJECTURE: The image of an information-preserving map i, $\{i(L)\}$, cannot have constant width (or even width $\leq$ poly ($\log$ crossing $\#(L)$).

This conjecture is implied by the conjecture $P \neq NP$ if we also make the modest assumptions that i can be computed on a link L in time $\leq$ poly(#crossing) and that the crossing number obeys: #crossing($i(L)$) $<$ poly(#crossing(L)).

The Jones polynomial at $e^{2\pi i/5}$, according to Witten [W], Reshetikhin and Turaev [RT], is the scalar output of a TQFT. Bounded width implies a fixed bound on the dimension of the Hilbert spaces which arises as the link is sliced into elementary bordisms.

Thus the calculation time for the composition of the elementary bordisms in TQFT is linear in the number of compositions. Since the dimension of Witten's Hilbert space grows (only) exponentially with width, poly log width is an adequate assumption for the entire calculation to grow at a polynomial rate.

D. "FINITE TYPE INVARIANTS" IN COMBINATORICS:

Vassiliev's book [V] contained implicitly a notion of "finite type" link invariant, clarified by Birman-Lin [BL] and Bar Natan [BN1], who showed that the perturbative invariants associated to the Witten-Chern-Simons theory are finite type. The fundamental idea of a finite-type invariant can be reproducted in any combinatorial setting where a notion of an (oriented) *elementary difference* can be defined. In oriented link theory the formal difference between two link diagrams, where a positive crossing in the first has been replaced by a negative crossing in the second, is the notion of elementary difference.

We give two examples in graph theory. In both cases the fundamental theorem [BN2] that the finite-type invariants of link theory can be computed in polynomial time continues to hold. One finds for invariants of type $= n$, a bound on computation time $\leq O(\#^n)$, where $\#$ is the number of edges in the graph. Analogous to the Witten-Chern-Simon theory where the l.h.s. is a $\#P$-hard nonperturbative invariant and the r.h.s. is an asymptotic expansion with finite-type coefficients, we find that (in the two cases respectively) after suitable change of variables, the chromatic and flow polynomials of a graph, which in their totality are $\#P$-hard to calculate, can be expressed as a polynomial whose k^{th} coefficient is of type $= k$.

EXAMPLE 1: Define an elementary difference on finite graphs modulo isomorphism to be an ordered pair consisting of a finite graph followed by the graph with one edge deleted, $(G, G \setminus e)$. A (real valued) invariant on finite graphs $f : \{\text{graphs}\} \to \mathbb{R}$ is *type* n, if given any collection of $n+1$ edges $\{e_1, \dots e_{n+1}\}$ of G, all $(n+1)^{\text{st}}$ order differences given by a sum over subsets vanishes:

$$\sum_{S \subset 2^{\{e_1, \dots, e_{n+1}\}}} (-1)^{|S|} f(G \setminus S) = 0 \tag{4}$$

The chromatic polynomial of a finite graph, $P_G(\lambda)$, has degree $= V$, the number of vertices, and satisfies the "contraction-deletion" recursion relation:

$$P_G(\lambda) - P_{G \setminus e}(\lambda) = -P_{G/e}(\lambda), \tag{5}$$

where $G \smallsetminus e$ is G with e deleted, and G/e is G with e contracted. Let $\overline{P}_G$ be the polynomial with "reversed" coefficients, $\overline{P}_G = \lambda^V P_G(\lambda^{-1})$. The recursion relation on $\overline{P}$ becomes:

$$\overline{P}_G(\lambda) - \overline{P}_{G\smallsetminus e}(\lambda) = -\lambda \overline{P}_{G/e}(\lambda). \tag{6}$$

Line (6) implies the constant term of $\overline{P}(\lambda)$ applied to an elementary difference (l.h.s. (6)) is zero. Inductively it is easy to see that the coefficient of a degree $= n$ term of $\overline{P}$ will vanish any formal differences of order $= n + 1$. Comparing with line (4) we have:

OBSERVATION 1: The coefficient of degree n in $\overline{P}$ is a type $= n$ invariant w.r.t. deletion of finite graphs.

EXAMPLE 2: Dual to the chromatic polynomial is the *flow polynomial* $F_G(\theta)$. It has degree $= E$, the number of edges of G, and satisfies the recursion relation

$$F_G(\theta) - F_{G/e}(\theta) = -F_{G\smallsetminus e}(\theta) \tag{7}$$

Let $\overline{F}_G(\theta) = \theta^E F_G(\theta^{-1})$ so that

$$\overline{F}_G(\theta) - \overline{F}_{G/e}(\theta) = -\theta F_{G\smallsetminus e}(\theta) \tag{8}$$

Now if we define the ordered pair $(G, G/e)$ to be the *elementary difference*, then we obtain a dual notion finite-type graph invariant and have the:

OBSERVATION 2: The coefficient of degree n in $\overline{F}$ is a type $= n$ invariant (w.r.t. contraction) of finite graphs.

A general principle seems to be that if the associated graded objects to the finite type invariants (dual cord diagrams in Vassiliev's theory) span a finite-dimensional space, then calculating finite-type invariants should be polynomial time in the complexity of the instance (eg., link, graph, etc., ...). (Compare with [BN2].)

In the case of graphs, for either of the two preceding notions of elementary difference, the graded object at level n is only 1-dimensional, being spanned by the general "n-singular" graph. In the two cases, the general n-singular graph is a formal signed difference $\sum_{S\subset G_0} (-1)^{|S|}(G \smallsetminus G_0)$ or $\sum_{S\subset G_0} (-1)^{|S|} G/\text{components } G_0$, respectively, where $G_0 \subset G$ is a subgraph of n edges. Thus the only finite-type invariants are polynomials in the coefficients of $\overline{P}$ and $\overline{F}$ respectively.

OBSERVATION 3: For type $= n$ invariants, w.r.t. deletion (or contraction), the time to compute is bounded by $O(E^n)$.

PROOF. Consider deletion; the contraction case is similar. If f is type $= n$, f is zero on graphs with $k + 1-$ singular graph edges and therefore constant on k-singular graphs with isomorphic singular sets. Given, as in the Vassiliev theory, a system of "integration" constants, it takes no more than E steps to evaluate the function f on graphs once f is known on 1-singular graphs. Each of these steps requires at most E preliminary steps to integrate a function on 2-singular graphs to obtain the evaluation of f on 1-singular graphs. Proceeding in this way, the result follows by induction.

For the chromatic polynomial, there is a subgraph sum formula for the coefficients, which gives the same growth in complexity we just obtained. It is also known that the linear coefficient of P is $\#P$ hard to compute. I presume the same is true for the flow polynomial. It is intriguing that there is a general approach to filtering $\#P$-hard information by polytime "approximations" of increasing degree. The art to finding useful approximations, less trivial than the two examples presented here, seems to be in choosing the "elementary differences". The situation is parallel to the Witten-Chern-Simon theory where there is a $\#P$-hard nonperturbative l.h.s. and an asymptotic expansion on the r.h.s. where the individual coefficients are finite type, and therefore polynomial time invariants.

From group theory we give a final example of an *unoriented* difference motivated by the formal structure of Wertinger presentations.

EXAMPLE 3: An elementary difference (G, G') is defined to be an unordered pair of groups where G and G' admit presentations which agree except for a single relation in which the literals (generators and generator inverses) read *backwards* in G' as compared to G. The consequence of the difference being unordered is that all finite type invariants defined from it are ambiguous up to sign. I have not yet made any investigation of this algebraic version of the "crossing change" in link theory.

§3. THE PHYSICS OF COMPUTATIONAL MODELS.

We should generally be interested in physical systems–even rather hypothetical ones—whose preparation may specify an instance of a problem and whose measurement can be (quickly) deconvolved to give the answer to that instance. A standard pitfall is to expect to make measurements to too great an accuracy, or at too low a temperature, or in some similar way to disregard the presence of some exponentially growing difficulty. At a fundamental level, any device is "analog". The distinction between analog and digital can be expressed as whether the coarse graining occurs later (analog) or earlier (digital). The success of digital over analog in the first 50 years of computers can be explained by realizing that the usual analog representations of a number, e.g., as a voltage, amounts to storing the number in unary and therefore exponentially less efficient than binary notation. On the other hand, it has been known for some time, that physically measurable quantities of some idealized systems are $\#P$-hard to compute. This makes one wonder if it is not worth the price of working in analog long enough to allow nature to make a truly difficult computation, rather than simply executing a gate, before measuring.

The Ising model for vertex spins on a graph with edge interactions has, in the ferromagnetic case, a Hamiltonian $H = -\sum_{\text{edges } e_{ij}} \sigma_i \sigma_j$, $\sigma_i \in \{-1, 1\}$. The partition function $Z(\beta) = \sum_{\text{spin states } \sigma} e^{-\beta H(\sigma)}$, $\beta = \frac{1}{kT}$, when written in a hightemperature expansion, becomes:

$$Z = e^{\beta|E|}P\left((e^2)^{-\beta}\right) \quad \text{where} \tag{9}$$

$P(x)$ is the generating function $\sum_{k=0}^{|E|} b_k x^k$ with $b_k = \#$ (bipartite subgraphs of k-edges). (See, e.g., [JS].)

For the purposes of this article let us pretend that $Z(\beta)$ is a measurable quantity. To be more realistic one might consider specific heat $= \frac{\partial^2 \log Z}{\partial \beta^2}$, or some correlation function such as $\left(\sum_\sigma \sigma_i\sigma_j e^{-\beta H(\sigma)}\right) \Big/ Z(\beta)$, but to illustrate our point we take the partition function as our measurable quantity, since the tie in to the graph theory is most convenient. The following analysis owes much to conversations with Christian Borgs and Jennifer Chayes.

Using the standard methods, the coefficients b_k take time $\leq O(E^k)$ to compute, so the low coefficients are easy; and it is further known [JS] that the highest non-zero b_k is $\#P$-hard. Our goal in building a "statistical mechanical computer" would therefore be to input a graph G and then tease out the leading coefficient $b_{\max}$ from measurements of $Z(\beta)$ at various temperatures. The problem is essentially to recover the coefficients of a polynomial from measurements of its values at $\{e^{-2\beta_i}\}$ for some collection of positive values of temperature T. This is done by inverting the linear system $(e^{-2j\beta_i})$. Since the coefficients are a priori integers, only some threshold accuracy is needed for an exact determination. Unfortunately numerical instabilities are encountered in the inversion. The essential point is that to determine the leading coefficient of a polynomial P, most information is gained by evaluating P at a large number (so that the low-order contribution is negligible). Unfortunately the physical requirement that temperatures be positive restricts $\beta_i > 0$ and therefore $0 < e^{-2\beta_i} < 1$; this forces P to be evaluated only at small values.

One way out of this numerical problem is to study an anti-ferromagnet on graphs with Hamiltonian $H = \sum_{\text{edges}} \sigma_i\sigma_j$; this allows P to be sampled in the range $1 < x < \infty$ where very low (positive) temperatures will be most revealing of $b_{\max}$. This resolves the numerical instability but ushers in a different problem: An antiferromagnet is a highly frustrated system and only approaches its Gibbs measure with exponential slowness: time to equilibrium $\approx O(e^{\frac{1}{kT}})$ as temperature approaches zero. So the "antiferromagnet computer" would take exponentially long to be initialized to the graph G whose $G_{\max}$ it was computing. Essentially, the antiferromagnet is not qualitatively more efficient at finding its equilibrium than the presently available numerical algorithm, the Metropolis method (see page 124 [We]), and might in fact be rather close to a highly parallel implementation of that algorithm.

The joint failure of the ferro- and anti-ferromagnet to lead (even in principle) to an analog computer for $\#P$ problems, suggests a generic weakness of classical statistical mechanical systems for computation. They sample states sequentially in time and hence have limited information-processing capacity. While a classical system, such as an Ising magnet, explores its state space sequentially in time, quantum mechanics offers another possibility.

The Feynman path integral computes the evolution operator as a coherent superposition of (infinitely many) states. The resulting evolution incorporates a vast amount of information very quickly; to be useful for computing, this evolution must be guided to answer discrete combinatorial problems. The discrete character of topological quantum field theories and their interpretation in terms of the $\#P$-hard Jones polynomial make these an attractive candidate for a new computational model [F1]. In the physics literature the Abelian Chern-Simons functional occurs in the Lagrangian for certain nonclassical surface layer conductivities governed by the integral quantum Hall effect [Ko]. The Abelian CS functional is known to compute linking number [S].

The $SU(2)$-Chern-Simons functional appears to enter into solid state physics via the fractional quantum Hall effect, a phenomenon of "quasi-particle" conductivity [Wi], [TL]. More abstractly, the TQFT with that functional as Lagrangian is known to compute $\#P$-hard values of the Jones polynomial [W].

Here is a notional sketch of how $SU(2)$-Chern-Simons theory might be implemented as a general computational model. A logical problem X, such as satisfiability of a Boolean formula, would be coded as a link $L = \text{code}(X)$. (See [J] and [JVW] for one way in which this may be done). The link would be described as a braid and implemented in a $(2+1)$-dimensional space time by forcing the motion of charge defects, i.e., "quasi particles", in a very cold surface layer of silicon. This is the preparation or "input" phase. If $SU(2)$-Witten-Chern-Simons is really physically important in this situation, one should expect some detectable "observable" consequence of the particular input braid, containing information on its Jones polynomial, as "output". A key point is whether the observable is a real number, e.g., a measured conductivity, in which case it is the analog version of a number expressed in unary. In contrast, if the observable can itself be some configuration or state of an ancillary collection of quasi-particles, then this is the analog version of a binary number, with addressable information, and much more efficient.

The choice of the translation to links raises a topological issue. For a link L of n crossings, an elementary estimate from the skein relations is that if c is a coefficient of the Jones polynomial V_L, then $|c| < (2\sqrt{2})^n$. Very crude statistical considerations—thinking of the coefficient as the result of a random walk as the contributions of various signs accumulate—suggest typically $|c| < (2\sqrt{2})^{n/2}$. On the other hand, in many cases these are overestimates, for torus links $|c| = 0$ or 1. If the observables, say $V_L(e^{2\pi i/p})$, must be read in "unary", it may be essential, given limited accuracy, to have the ensemble of links, image(code), more like the torus links than the generic link. Is it possible to encode the general Boolean formula into links in such a way that (1) from the evaluation of $V_{\text{code}(X)}, \text{sat}(X)$ may be quickly determined, and (2) so that $|c| \leq \text{poly}(\text{length } X)$ for the coefficients c of $V_{\text{code}(X)}$? Here is a separate question, but with the same motivation: Are there TQFTs which yield information about $V_{(p)L}$, the Jones polynomial $\in Z_p[\mathbb{C}]$ with coefficients reduced modulo a prime p? Positive answers to either question would ease the problem of identifying V_L from observations of limited accuracy, and allow a Chern-Simons theory even with an essentially unary output, to form the basis of a powerful, if still theoretical, model of computation.

Computer science is driving an interaction between logic, physics, and mathematics, which will explore the ability of the physical world to process information. I have tried to convey the excitement and scope of this endeavor and to point to paths that mathematicians, particularly topologists, might penetrate.

References

[ADH] L. Adelman, J. Demarrias and M. Huang *Quantum computability*, Siam J. Computing 26 (1997), 1524–1540.

[At] M. Atiyah, *Geometry and Physics of Knots*, Lezioni Lincee [Lincei Lectures], Cambridge University Press, Cambridge, 1990.

[BGS] T. Baker, J. Gill, and R. Solovay, *Relativizations of the $\mathcal{P} = ?\mathcal{NP}$ question*, SIAM J. Comput. 4 (1975), 431-442.

[BL] J. Birman and X.-S. Lin, *Knot polynomials and Vassiliev's invariants*, Invent. Math. 111 (1993), 225–270.

[BN1] D. Bar-Natan, *Perturbative aspects of the Chern-Simons topological quantum field theory*, Ph.D. Thesis, Princeton Univeristy, 1991; *On the Vassiliev knot invariants*, Topology 34 (1995), 423–472.

[BN2] D. Bar-Natan, *Polynomial invariants are polynomial*, Math. Res. Lett. 2 (1995), 239–246.

[C] S. Cook, *The complexity of theorem-proving procedures*, Proc. 3rd ACM Symp. on Theory of Computing, Association for Computing Machinery, New York, 1971, pp. 151–158.

[Ch] G. Chaitin, *Information-theoretic limitations of formal systems*, J. Assoc. Comput. Mach. 21 (1974), 403–424.

[F1] M. H. Freedman, *P/NP, and the quantum field computer*, Proc. Natl. Acad. Sci. USA 95 (1998), 98–101.

[F2] M. H. Freedman [1998] *Limit, logic, and computation*, Proc. Natl. Acad. Sci U.S.A. 95, 95–97.

[F3] M. H. Freedman, k-sat on groups and undecidability, 1998 ACM Symp. on Theory of Comp. (STOC '98) (to appear).

[G] M. Gromov, *Groups of polynomial growth and expanding maps*, Inst. Hautes tudes Sci. Publ. Math. 53 (1981), 53–73; *Carnot-Carathéodory spaces seen from within*, Progr. Math. 144, Sub-Riemannian geometry (1996), 79–323.

[J] F. Jaeger, *Tutte polynomials and link polynomials*, Proc. Amer. Math. Soc. 103 (1988), 647–654.

[Jo] V. Jones *A polynomial invariant for knots via von Neumann algebras* Bull. Amer. Math. Soc. 12 (1985), 103–111.

[JS] M. Jerrum and A. Sinclair, *Polynomial-Time Approximation Algorithms for Ising Model*, Proc. 17th ICALP, EATCS, 1990, 462-475.

[JVW] F. Jaeger, D. L. Vertigan, D. J. A. Welsh, *On the computational complexity of the Jones and Tutte polynomials*, Math. Proc. Cambridge Philos. Soc. 108 (1990), 35–53.

[K] B. Kane, *A Silicon-based nuclear spin quantum computer*, Nature 393 (1998), 133-137.

[Ko] M. Kohmoto, *Topological invariants and the quantization of the Hall conductance*, Ann. Physics 160 (1985), 343–354.

[L] S. Lloyd, Universal quantum simulators, Science 273 (1996), 1073–1078.

[P] J. Preskill, *Fault tolerant quantum computation*, http://xxx.lanl.gov/ps/quant-ph/9712048, 19 Dec. 1997 (*Introduction to Quantum Computation* (H.-K. Lo, S. Popescu and T. P. Spiller, eds.) (to appear)).

[RT] N. Reshetikhin and V. Turaev, *Invariants of 3-manifolds via link polynomials and quantum groups*, Invent. Math. 103 (1991), 547–597.

[S] A. S. Schwarz, *The partition function of degenerate quadratic functional and Ray-Singer invariants*, Lett. Math. Phys. 2 (1977/78), 247–252.

[Sc] A. Schöhage, Proc. 6th ICALP Lect. Notes in Comp. Sci. (Springer, NY) 71 (1974), 520–529.

[Sh] P. Shor, *Algorithms for quantum computation: discrete logarithms and factoring*, 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, 1994, IEEE Comput. Soc. Press, Los Alamitos, CA, 1994, pp. 124–134.

[TL] C. Ting, and C. H. Lai, *Spinning braid-group representation and the fractional quantum Hall effect* Nuclear Phys. B 396 (1993), 429–464.

[V] A. Vassiliev, *Cohomology of knot spaces. Theory of singularities and its applications*, Adv. Soviet Math. 1 (V. I. Arnold, ed.), Amer. Math. Soc. (Providence, RI), 1990.

[W] E. Witten *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. 121 (1989), 351–399.

[We] Welsh, D. J. A. *Complexity: knots, colourings and counting* London Mathematical Society Lecture Note Series 186, Cambridge University Press, Cambridge, 1993.

[Wi] F. Wilczek, *Fractional statistics and anyon superconductivity* World Scientific Publishing Co., Inc., (Teaneck, NJ) 1990.

Michael H. Freedman
Microsoft Research
1 Microsoft Way
Redmond, WA 98052-6399