The computational complexity of the self-avoiding walk on random lattices

1992 J. Phys. A: Math. Gen. 25 521

(http://iopscience.iop.org/0305-4470/25/3/012)

J. Phys. A: Math. Gen. **25** (1992) 521-527. Printed in the UK

The computational complexity of the self-avoiding walk on random lattices

Jonathan Machta†

Department of Physics and Astronomy University of Massachusetts Amherst, MA 01003, USA

Received 25 May 1991

Abstract. It is proved that finding a minimum energy, N-step, self-avoiding walk on a 2D square lattice with assigned site energies is an NP-complete problem. However, if the fraction of lowest energy sites exceeds the percolation threshold then it is argued the problem can almost always be solved in polynomial time.

1. Introduction

This paper is concerned with the computational complexity of finding minimum energy self-avoiding walks (SAWs) in quenched random environments. Several factors have led to the recent interest in SAWs in random environments. Firstly, this model may help us understand the behaviour of polymers in gels and other porous materials. Secondly, this model provides an appealing example of a statistical mechanics problem in which disorder plays an important role. The apparent simplicity of the model is deceptive and despite more than a decade of work by many investigators, considerable disagreement persists. For example, it is of interest to find the scaling of the end-to-end size, R, of the SAW, $\langle R \rangle \sim N^{\zeta}$ where the brackets represent an average over N-step SAWs and the overbar an average over environments. It is not known whether ζ takes the same value as for SAWs in pure enviornments or, if not, whether the value of ζ changes at the percolation threshold. The reader is referred to Le Doussal and Machta (1991) or Lee *et al* (1989) for references to the earlier literature.

Recently it has become clear that the SAW in a random environment is governed by a strong disorder fixed point for low dimensionality or sufficiently strong disorder (Obukhov 1990, Le Doussal and Machta 1991). Problems governed by strong disorder or zero temperature fixed points such as the directed walk in a random environment, random field magnets and spin-glasses, have proved difficult to understand and it is useful to have another clear-cut example to study.

One can often extract useful information about the statistical mechanics of a system governed by a strong disorder fixed point by studying its ground states. Here this means finding a minimum energy or optimal SAW among all N-step SAWs starting from the same origin in a given environment. It is believed that the exponent ζ can also be obtained by computing R for an optimal SAW and then averaging over environments, $R_{\text{opt}} \sim N^{\zeta}$. This identification is justified by a renormalization group analysis which shows that the effective temperature in the problem flows to zero under renormalization.

† Internet address: MACHTA@PHAST.UMASS.EDU

0305-4470/92/030521+07$04.50

Thus, if an efficient algorithm could be found to generate optimal SAWs, it could be used to obtain critical exponents such as ζ. Unfortunately, the results of this paper make it unlikely that such an algorithm exists.

In this paper we investigate the computational difficulty of finding an optimal SAW in a given environment. This is a combinatorial optimization problem similar to other extensively studied problems in computer science such as the travelling salesman (Lawler *et al* 1985). During the last two decades considerable progress has been made in characterizing the computational complexity of various problems. In particular, the property of NP-completeness was introduced and shown to hold for a large number of problems (see Garey and Johnson 1979 for an introduction to the theory of NP-completeness). The class of NP (non-deterministic polynomial time) problems consists of decision problems for which a proposed solution can be checked in polynomial time. The NP-complete class consists of those members of NP which are at least as hard as any other member of NP. If an algorithm exists that could solve one NP-complete problem in polynomial time, it could be used to solve any NP problem in polynomial time. However, it is generally believed, though not proved, that any algorithm which solves all instances of an NP-complete problem must run in exponential time for at least some instances of the problem. The travelling salesman is an example of an NP-complete problem. A comprehensive list of NP-complete problems can be found in Garey and Johnson (1979). It is often difficult to intuit whether a given problem is NP-complete. For example, finding the minimum energy SAW of N or *fewer* steps on a lattice with random positive energies can be solved in polynomial time. On the other hand, the main result of this paper is that finding a minimum energy SAW of *exactly* N steps is an NP-complete problem.

Closely related to the SAW in a random environment is the directed self-avoiding walk (DSAW) in a random environment. The DSAW differs from the SAW in that one coordinate along the path is required to be non-decreasing. Like the SAW, the DSAW is governed by a strong disorder fixed point for low dimensionality or sufficiently strong disorder (Kardar 1985, Cook and Derrida 1989). Though the DSAW and the SAW share many features in common (Le Doussal and Machta 1991) optimal DSAWs in a random environment can be found using an efficient polynomial time algorithm (Huse and Henley 1985). On the other hand, the fact that the SAW in a random environment is NP-complete suggests that obtaining ground states will be computationally difficult. Another statistical mechanics model with quenched disorder for which finding the ground state is known to be NP-complete is the spin-glass in more that two dimensions (Barahona 1982, Bachas 1984). On the other hand, the random field Ising model is a polynomial time problem in any dimension (Barahona 1985). It is of general interest to develop connections between statistical mechanics and the theory of computational complexity (see Mezard *et al* 1987).

2. Proof of NP-completeness for SAWs

The SAW in a random environment is defined as follows. Consider an $L \times L$ square lattice. To each site, α, of the lattice assign an energy cost, E_α, for visiting that site. An N-step SAW on the lattice is defined as a non-repeating sequence of nearest-neighbour sites, $\Gamma = \langle \alpha_1, \alpha_2, \ldots, \alpha_N \rangle$. The first site, α_1, is fixed at the origin which is taken as the centre of the lattice. The energy of a SAW, $E(\Gamma)$, is the sum of the energies of the sites along Γ. A particular instance of the optimization problem is an $L \times L$

lattice with assigned energies and a given length N such that $N < L^2$. The goal is to find a least energy SAW of N steps. We refer to such a SAW as optimal.

A decision problem associated with finding an optimal SAW can be posed as follows: 'Is there an N-step SAW starting at the origin on a given realization of an $L \times L$ lattice with an energy less than or equal to a bound?'. In the theory of NP-completeness the size of the problem is measured in terms of the number of bits required to specify an instance of the problem so it is necessary to restrict each site energy to a finite number of bits. In fact we shall prove NP-completeness when the site energies are either 0 or 1. Henceforth we refer to the decision problem with energies 0 or 1 simply as SAW. Since the energy of a SAW can be computed in $O(N)$ time, SAW is in the class NP.

The general technique for proving that a new problem is NP-complete is based upon constructing a 'polynomial transformation' from a known NP-complete problem to the new problem. The idea is to show that the existence of a polynomial time algorithm for the new problem would imply that the known problem (and hence all NP-complete problems) could be solved in polynomial time. For the proof that SAW is NP-complete, the known problem is deciding whether there is a Hamiltonian path on a given cubic, planar graph. We refer to this problem as HP. An instance of HP is a cubic (each vertex has three edges), planar graph, $G = (E, V)$ (see Berge (1985) for an introduction to graph theory). A Hamiltonian path is a simple path (self-avoiding walk) which visits each vertex of the graph. Garey *et al* (1976) proved that HP is NP-complete.

The strategy for showing that SAW is NP-complete is to embed an arbitrary planar cubic graph G in an $L \times L$ lattice. The sites of the lattice which represent vertices or edges of G are assigned energy 0 while all other sites are assigned energy 1. A Hamiltonian path exists on G if and only if there is a zero energy SAW with a length determined by the embedding of G. A crucial ingredient in the proof is checking that the required embedding can be accomplished in polynomial time and that the size of the required lattice grows as a polynomial in the size of G.

Let M be the order (number of vertices) of the graph G. The first step is to obtain a plane drawing of the graph. Hopcroft and Tarjan (1974) show that there is an algorithm which accomplishes this in $O(M)$ time. The next step is to construct a first embedding of the graph in an $L_1 \times L_1$ lattice. An embedding consists of a map from the set of vertices, V, into the sites of the lattice and from the set of edges, E, to non-intersecting simple paths on the lattice. A practical way to visualize an embedding is to imagine that each vertex represents an electrical circuit element and that each edge represents a connecting wire. Then an embedding is a layout of the circuit on a circuit board which has no wire crossings.

Given a plane drawing of a planar graph of degree four or less, Valiant (1981) showed that an embedding on an $L_1 \times L_1$ square lattice can be obtained with $L_1 = O(M)$ and that the embedding algorithm requires $O(M^2)$ time. For the present argument one needs an embedding with the additional property that all edges are represented by paths of same length. The following construction converts the first embedding into one which has equal path lengths. The length, Q, of the longest path in the first embedding is determined. Note that $Q = O(M)$. Next the $L_1 \times L_1$ square lattice and the associated representation of G are embedded in a refined, $L_2 \times L_2$, square lattice such that $L_2 = 3KL_1$ where K is to be determined. Each nearest-neighbour bond of the first lattice is mapped onto a straight segment of $3K$ bonds in the second lattice so that the length of the longest path representing an edge in G is now $3KQ$. In order to lengthen shorter paths, 'switchbacks' are added as shown in figure 1. If P switchbacks

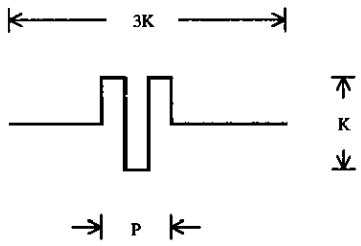

Figure 1. Introducing switchbacks to increase the path length by PK.

are added to a straight path of length $3K$ then the length of the path becomes $3K + PK$. If K is taken as $3(Q-1)$ then every path can be modified so that its length is $9Q(Q-1)$. For example suppose that some edge is represented by a single bond in the first embedding and thus by a straight path of length $3K$ in the refined embedding. Then, inserting $P = K = 3(Q-1)$ switchbacks yields a new path of length $9Q(Q-1)$ as desired. Notice that the switchbacks are inserted in the middle third of the original bonds insuring that path crossings are not created. Finally the switchback are 'insulated' from each other by an additional refinement in which the number of the lattice points is doubled as shown in figure 2. The final size of the lattice is $L \times L$ with $L = 18(Q-1)L_1 = O(M^2)$. The embedding of G in this lattice can be performed in time $O(M^3)$ and requires specifying $O(M^3)$ sites for the representation of G. The sequence of transformations from the plane drawing to the equal length, insulated embedding is shown in figure 2 for an example with $M = 2$ and $Q = 3$.

The transformation of HP to SAW is completed by assigning energy 0 to each site on the $L \times L$ lattice which represents a vertex of G or is part of a path representing an edge of G. All other sites are assigned an energy 1. If a Hamiltonian path exist in G then it is represented by a SAW of exactly $18Q(Q-1)M$ steps. Thus the length of the SAW is chosen as $N = 18Q(Q-1)M$. If there is a SAW of this length with energy zero it represents a Hamiltonian path since it traverses exactly M edges of G, visiting each vertex once. Conversely, if there is a Hamiltonian path in G, it is represented by an $N = 18Q(Q-1)M$-step SAW with energy zero. Note that insulating the switchbacks (see figure 3) insures that zero energy paths do not stray from the embedded edges.

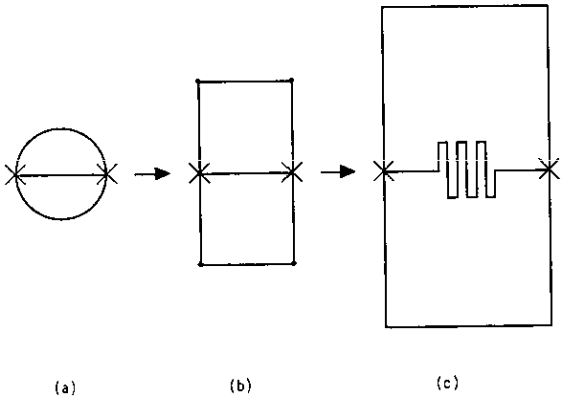

Figure 2. Transformations of a planar graph: (*a*) the plane drawing, (*b*) the first lattice embedding and (*c*) the equal path length, insulated embedding. The individual lattice sites are not shown in (*c*) but the lattice size is 36×72. Vertices are marked $\times$.

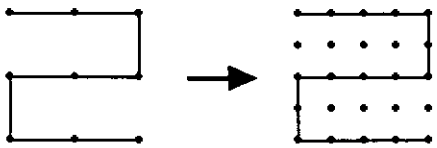

Figure 3. Doubling the number of lattice sites so that switchbacks are insulated from each other.

The size of the required instance of SAW is $O(L^2) = O(M^4)$ and the time required to transform HP to a SAW is $O(M^3)$. Thus we have exhibited a polynomial transformation from HP to SAW and thereby proved that SAW is NP-complete.

It is a simple consequence of this theorem that several related optimization problems are at least as hard as SAW and thus 'NP-hard'. Examples of NP-hard problems are finding optimal SAWs on lattices of dimension two and greater and with real number site energies.

3. Cases where SAW may be solved in polynomial time

The fact that SAW is NP-complete strongly suggests that any algorithm capable of solving it will run for exponential time for *some* instances of the problem. Are these instances generic or exceptional? A discussion of this question requires imposing a probability distribution on the assignment of site energies. Suppose the site energies are independent, identically distributed (i.i.d.) random variables taking the value 0 with probability p and 1 with probability $1-p$. Suppose that $N < L/2$ so that no SAW reaches the perimeter of the lattice and that $p > p_c$ where $p_c = 0.592\,75\ldots$ is the percolation threshold for two-dimensional site percolation (see Stauffer 1985). For $p > p_c$ the general situation is that there is a spanning or 'infinite' cluster of lowest energy sites which either includes the origin or encloses the origin within a cavity whose size scales as the percolation correlation length.

Despite the fact that SAW is NP-complete, above the percolation threshold there is a polynomial time algorithm which almost always finds an optimal N-step SAW. This algorithm is based upon a standard shortest path algorithm, the Bellman-Ford method (see Lawler 1976). The Bellman-Ford algorithm finds a minimum energy simple path of K steps or less between a given point and all other points on an $L \times L$ lattice with non-negative site energies. We call the lowest energy simple path of K steps or less between two sites a K-optimal path. The Bellman-Ford algorithm runs in time $O(L^6)$.

Define S as the set of sites which are connected by zero energy simple paths to the perimeter. Note that S includes the infinite cluster of zero energy sites. The first step is constructing an optimal SAW is to use the Bellman-Ford method to find S and a zero energy path from the perimeter to every point in S. If the origin is in S we are done; if not the next step is to find an N-optimal path from the origin to every site in S which can be connected to the origin in N steps or less. Thus we have two simple paths to every site $\alpha \in S$ which is sufficiently close to the origin; one from the origin to α and a second, with zero energy, from α to the perimeter. Among the sites in S choose a site having the lowest energy N-optimal path to the origin. Call this path Γ' and its terminus $\beta \in S$. If this path has fewer than N steps, a segment of the zero energy path from β to the perimeter is appended to Γ' to obtain an optimal N-step SAW, Γ, whose energy is $E(\Gamma) = E(\Gamma')$. If Γ' intersects the path from β to the perimeter,

an optimal SAW can be obtained by truncating Γ' at the last intersection with the path to the perimeter and then adding the required length of the path to the perimeter. The running time of this algorithm is $O(L^7)$.

If any optimal SAW visits S then the above algorithm is guaranteed to find an optimal SAW. To see this, suppose Γ'' is optimal and visits $\gamma \in S$. Then the algorithm will have identified a N-optimal path from the origin to some $\beta \in S$ which has an energy less than or equal to $E(\Gamma'')$. This path is then extended to N steps with no further increase in energy thus constructing the SAW, Γ, with energy less than or equal to $E(\Gamma'')$. Since Γ'' is optimal, $E(\Gamma) = E(\Gamma')$ and the algorithm has produced an optimal walk. On the other hand, if no optimal SAW visits S, the algorithm may fail to produce an optimal SAW.

In the limit that N and L are taken to infinity holding $p > p_c$ fixed, the algorithm given above is expected to find an optimal SAW with probability one. The argument for this, though non-rigorous, follows from accepted ideas in percolation theory. The key idea is that the incipient infinite cluster and thus S typically enclose the origin within a region of the size of the correlation length, ξ, or less. More specifically, consider the set of sites, A, which are connected to the origin by simple paths which do not intersect the incipient infinite cluster. We expect that the probability density for the size of A falls off exponentially in $|A|/\xi^2$. As N becomes large the probability that all SAWs starting at the origin intersect the incipient infinite cluster approaches one exponentially in N. Note that self-avoidance prevents a SAW from occupying an area less than N. Thus, in the limit that N and L go to infinity, the above algorithm almost always yields a provably optimal SAW. The conclusion is that finding optimal SAWs is *generically* a polynomial time problem above the percolation threshold in two dimensions. It is tempting to conjecture that finding optimal SAWs is generically NP-hard below the threshold for percolation of the lowest energy sites.

Unfortunately, finding a *single* optimal SAW on lattices where the 'good' (lowest allowed energy) sites percolate does not yield useful information about the statistical mechanics of SAWs in random environments. The reason for this is that there are too many optimal SAWs on lattices with percolating good sites so that entropy plays an important role and the full ensemble of low energy SAWs must be investigated. Thus the above polynomial time algorithm cannot be used to study the controversial issue of SAWs on percolation clusters (see Lee *et al* 1989). The abundance of exact ground states is probably also the reason why it is computationally easy to find *some* optimal walk. Within a renormalization group analysis (Le Doussal and Machta 1991), one finds that the zero temperature recursion relations, corresponding to the optimization problem, flow to the pure system fixed point when the fraction of lowest energy sites exceeds the percolation threshold whereas the finite temperature recursion relations flow to the strong disorder fixed point. On the other hand, if the good sites do not percolate (as is always the case when the site energies are chosen from a continuous distribution) then the zero temperature and finite temperature recursion relations both flow to the strong disorder fixed point implying that the optimization problem and the statistical mechanics problem are equivalent.

4. Discussion

From the point of view of understanding the statistical mechanics of SAWs in random environments our conclusions are discouraging. Since SAW is NP-complete, there is

little hope of finding a polynomial time algorithm of the kind that has allowed so much progress to be made on the directed walk problem. The property of NP-completeness also leads us to suspect that the range of analytic and numerical subtleties which characterize spin-glasses may also appear in the disordered SAW problem. Is there, for example, replica symmetry breaking?

The suspicion that NP-completeness implies spin-glass like features leads us to inquire into the connection, if any, between the computational complexity and statistical mechanics of a model system? There are few hard results in this area (see, however, Mezard *et al* 1987, Baum 1986). Comparing the disordered SAW and DSAW may prove useful in addressing this issue. Though these models have a number of similar features they differ in their computational complexity; there is a polynomial time algorithm for finding an optimal DSAW while finding an optimal SAW is NP-hard. It would be interesting to uncover a qualitative physical difference between these models which could be traced to their differing computational complexity.

Acknowledgments

I thank Pierre Le Doussal and Yonathan Shapir for useful discussions. I am grateful to Seth Malitz for carefully reading the manuscript and bringing relevant articles in the computer science literature to my attention. This work was supported in part by NSF Grant DMR 9014366.

References

Bachas C P 1984 *J. Phys. A: Math. Gen.* **17** L709
Barahona F 1982 *J. Phys. A: Math. Gen.* **15** 3241
—— 1985 *J. Phys. A: Math. Gen.* **18** L673
Baum E B 1986 *Phys. Rev. Lett.* **57** 2764
Berge C 1985 *Graphs* (New York: Elsevier)
Cook J and Derrida B 1989 *J. Stat. Phys.* **57** 89
Garey M R and Johnson D S 1979 *Computers and Intractability* (San Francisco: Freeman)
Garey M R, Johnson D S and Tarjan R E 1976 *SIAM J. Comput.* **5** 704
Hopcroft J and Tarjan R 1974 *J. Assoc. Comput. Mach.* **21** 549
Huse D A and Henley C L 1985 *Phys. Rev. Lett.* **54** 2708
Kardar M 1985 *Phys. Rev. Lett.* **55** 2923
Lawler E L, Lenstra J K, Rinnooy Kan A H G and Shmoys D B 1985 *The Travelling Salesman Problem: A Guided Tour of Combinatorial Optimization* (New York: Wiley)
Le Doussal P and Machta J 1991 *J. Stat. Phys.* to appear
Lee S B, Nakanishi H and Kim Y 1989 *Phys. Rev.* B **39** 9561
Mezard M, Parisi G and Virasoro M A 1987 *Spin-Glass Theory and Beyond* (Singapore: World Scientific)
Obukhov S P 1990 *Phys. Rev.* A **42** 2015
Stauffer D 1985 *Introduction to Percolation Theory* (London: Taylor and Francis)
Valiant L G 1981 *IEEE Trans. Comput.* **C30** 135