Entanglement complexity of self-avoiding walks

1992 J. Phys. A: Math. Gen. 25 6557

(http://iopscience.iop.org/0305-4470/25/24/010)

J. Phys. A: Math. Gen. 25 (1992) 6557–6566. Printed in the UK

Entanglement complexity of self-avoiding walks

E J Janse van Rensburg†, D A W Sumners‡, E Wasserman§ and S G Whittington¶

† Department of Mathematics and Statistics, York University, Downsview, Ontario, Canada M3J 1P3
‡ Department of Mathematics, Florida State University, Tallahassee, FL 32306-3027, USA
§ E I Du Pont de Nemours & Co., Wilmington, DE 19880-0328, USA
¶ Department of Chemistry, University of Toronto, Toronto, Canada M5S 1A1

Received 11 June 1992

Abstract. Self-avoiding walks on three-dimensional lattices are flexible linear objects which can be self-entangled. We discuss several ways to measure entanglement complexity for n-step walks, and prove that these complexity measures tend to infinity with n. For small n, we use Monte Carlo methods to estimate and compare the n-dependence of two of these complexity measures.

1. Introduction

Linear polymer molecules in dilute solution are highly flexible and can be both self-entangled and entangled with other molecules. This paper is concerned with characterizing the degree of self-entanglement, and the dependence of self-entanglement on polymer length. The problem is of some practical importance because of the influence of polymer entanglement on crystallization behaviour (de Gennes 1984) and on rheological properties (Edwards 1967). Moreover, topological entanglement of flow lines is related to the magnetic energy of flows in incompressible perfectly conducting fluids (Freedman 1988, Freedman and He 1991).

If the molecule undergoes a ring closure reaction, the resulting ring polymer can be knotted, and questions about entanglement complexity can then be asked and (to some extent) answered using standard topological ideas from knot theory. A question (Frisch and Wasserman 1961, Delbruck 1962) which has attracted attention for thirty years is: what is the probability of knot formation, as a function of the number n of monomers in the ring polymer?

A convenient model for a ring polymer in dilute solution in a good solvent is an n-step self-avoiding polygon on a regular lattice, such as the simple cubic lattice Z^3. The knot probability for small n in this system has been studied using Monte Carlo methods (Vologodskii *et al* 1974, Janse van Rensburg and Whittington 1990), and the corresponding problem for a continuum model has also been investigated (Michels and Wiegel 1986, Koniaris and Muthukumar 1991). The consensus from these calculations is that, for moderate values of n, the knot probability is rather small.

0305-4470/92/246557+10$07.50

It can be shown rigorously (Hammersley 1961) that the number p_n of self-avoiding polygons behaves as

$$p_n = \mathrm{e}^{\kappa n + o(n)} \tag{1.1}$$

and (Sumners and Whittington 1988, Pippenger 1989) that the number p_n^0 of unknotted polygons behaves as

$$p_n^0 = \mathrm{e}^{\kappa_0 n + o(n)} \tag{1.2}$$

with $0 < \kappa_0 < \kappa$ so that the knot probability $P(n)$ behaves as

$$P(n) = 1 - \mathrm{e}^{-\alpha n + o(n)} \tag{1.3}$$

for some positive constant α. Hence the knot probability tends to unity (exponentially rapidly) as n goes to infinity. Similar asymptotic results hold for various continuum models of piecewise-linear polygons in R^3 (Diao 1990, Diao *et al* 1992). See also Frisch and Klempner (1970) and Kendall (1979).

For polygons there are many possible measures of entanglement complexity, including crossing number, genus, unknotting number, span of a knot polynomial (such as the Alexander or Jones polynomial), the value of the Alexander polynomial, $\Delta(t)$, at $t = -1$, etc. Soteros *et al* (1992) have shown that each of these measures of entanglement complexity diverges as n goes to infinity. The proof of this statement relies on 'tight' trefoils appearing on long polygons with positive density, and the fact that these quantities add (or multiply) for trefoils.

From a strictly topological point of view linear polymers are unknotted, and the arguments and numerical approaches used for the ring case cannot be directly applied to the linear case. This is because entanglement is a property of a pair of spaces, the three-dimensional ambient space and the one-dimensional subspace. If the subspace is a circle, the ambient space can be taken to be all of R^3. On the other hand, if the subspace is an arc, the ambient 3-space used to define entanglement must be either carefully chosen or canonically defined by the walk itself (Sumners and Whittington 1990). In section 2 we discuss several ways in which entanglement complexity can be defined for a self-avoiding walk model of a linear polymer, point out advantages and disadvantages of these schemes, and prove asymptotic results about them. In section 3 we apply some numerical tests to two of these, and then compare them by calculating the average measures of entanglement complexity for self-avoiding walks with n steps, as a function of n, using Monte Carlo methods.

2. Measures of entanglement complexity

We begin this section by discussing some results which bear on the entanglement complexity of self-avoiding walks, and which can be proved rigorously. Sumners and Whittington (1988) introduced the idea of a *knotted arc*. The lattice Z^3 consists of a vertex set, which is the set of integer points in R^3, and an edge set which is the set of edges joining pairs of vertices which are unit distance apart. If we consider any finite self-avoiding walk on Z^3, which we shall refer to as a *pattern*, then each vertex of the pattern has an associated dual 3-cell (a unit cube) with the occupied vertex as its barycentre. The union of these dual 3-cells forms a canonical (uniquely determined by the pattern) neighbourhood of the pattern in 3-space. This neighbourhood may be homeomorphic to a 3-ball, and we restrict our attention to patterns with this property. By extending the pattern by two *half-edges* so that they intersect the boundary of the neighbourhood (which is homeomorphic to a 2-sphere), we obtain a (3,1) ball-pair in which the oriented 1-ball (the pattern together with the two additional half-edges) is properly embedded in the 3-ball. Two (3,1) ball-pairs are the same *oriented knot type* if there is a homeomorphism of pairs from one to the other, such that the homeomorphism preserves the orientation of both the ambient 3-ball and the arc. The standard unknotted ball-pair is the intersection of the x-axis with the usual unit ball ($x^2+y^2+z^2 \leqslant 1$) in R^3. A (3,1) ball-pair is *unknotted* if there is a homeomorphism of pairs of the ball-pair onto the unknotted ball-pair. If not, we call the pattern a *knotted arc*, and the (3,1) ball-pair determined by the pattern is a *tight knot*. If, in addition, there exists a self-avoiding walk on which (translates of) the pattern occur three times then the pattern is a *K-pattern*. Figure 1 shows an explicit example of a K-pattern which corresponds to a trefoil knotted arc under the above neighbourhood construction. In fact (Soteros *et al* 1992), every knot type can be represented by a K-pattern which forms a tight knot, so Kesten's theorem (Kesten 1963) can be used to prove the following theorem:

Theorem 2.1. Every (3,1) ball-pair knot type appears as a knotted arc with positive density in all except exponentially few sufficiently long self-avoiding walks on Z^3.

If we now compute the Alexander polynomial $\Delta(t)$, say, of the polygons obtained by joining up the ends of n-step self-avoiding walks, with a curve which does not enter any of the 3-balls associated with the knotted arcs, and average $\Delta(-1)$ over all such walks, this implies that

$$\langle \log \Delta(-1) \rangle_n \to \infty \tag{2.1}$$

as $n \to \infty$. In fact the divergence is at least linear in n.

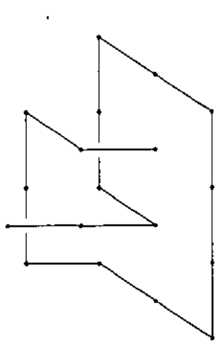

Figure 1. A knotted arc which is also a K-pattern.

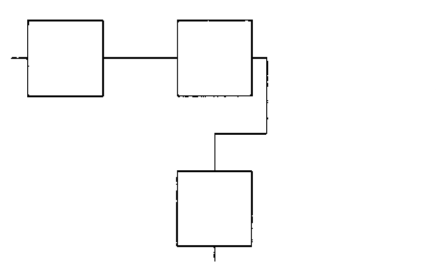

Figure 2. A pattern consisting of three knotted arcs. The squares represent the 3-balls defined by the arcs.

This suggests that knotted arcs might give a useful measure of entanglement complexity in self-avoiding walks. The first problem is that finding such arcs is computationally very difficult. In addition, if one could find them, the entanglement complexity would be an 'underestimate', in some rather loose sense, since the walk will most likely contain other contributions to its entanglement.

Instead, one needs a property of the complete walk, rather than a collection of terms corresponding to sub-walks. Suppose that we only consider *unfolded* walks (Hammersley and Welsh 1962). We shall call a walk x-unfolded if the x-component of the first vertex is strictly less than that of any other vertex and if no vertex of the walk has x-component greater than that of the last vertex. That is, if we number the vertices $i = 0, 1, 2, \ldots n$ and write (x_i, y_i, z_i) for the coordinates of the ith vertex then a walk is x-unfolded if $x_0 < x_i \leqslant x_n$, for all $i > 0$. The walk shown in figure 1 is x-unfolded.) Unfolded walks can be uniquely completed to (3,1) ball-pairs which capture all of the entanglement complexity of the unfolded walk. We add half-edges in the negative and positive x-direction to the first and last vertices of the walk and construct a 3-ball with two parallel faces perpendicular to the x-axis, and four more faces (perpendicular to the remaining coordinate directions) so that the walk (and the added half-edges) is properly embedded in the 3-ball. We say that the unfolded walk is *knotted* if the knot type of the associated (3,1) ball-pair is non-trivial. Since unfolded walks are not exponentially rare in the set of self-avoiding walks (Hammersley and Welsh 1962), it is easy to prove:

Theorem 2.2. Every (3,1) ball-pair knot type appears as a knotted arc with positive density in all except exponentially few sufficiently long unfolded self-avoiding walks on Z^3. Hence, all except exponentially few unfolded walks on Z^3 are knotted.

It is easy to establish that (2.1) is valid for unfolded walks and, again, that the divergence of $\langle \log \Delta(-1) \rangle_n$ is at least linear in n.

In this case it is straightforward (computationally) to check whether a walk is unfolded and, if it is, to compute the above measure of entanglement complexity. The problem is that most walks are not unfolded. Of course, they could be converted into unfolded walks by successive reflections, but this operation can create or destroy entanglements.

There is one measure of entanglement complexity which can be defined for any arc in R^3, and which does not require the careful construction of an ambient 3-ball to define the entanglement. One of the most appealing visual measures of knot complexity is *crossing number*. For the knot type of an embedded circle in 3-space, the crossing number is the minimum number of crossings possible in any projection, minimized over all projections of all representatives of that knot type. Since every arc in R^3 is unknotted (i.e. ambient isotopic to a subarc of the x-axis), any non-trivial complexity measure we define using R^3 as the ambient space will fail to be a knot type invariant. Nevertheless, for a fixed arc in R^3, we can define the *crossing number* of that arc to be the average value of the number of crossings in a projection, averaged over all projections (i.e. directions on the 2-sphere). We define the *crossing number* χ of a self-avoiding walk as follows: For a given projection direction $\hat{x}$ which gives rise to a regular projection (all crossings are transverse crossings of a pair of strands), we let $\chi_1(\hat{x})$ denote the number of crossings in that projection. If this arc is contained in a flow line of a perfectly conducting incompressible fluid, the crossing number defined earlier is related to the magnetic energy of that flow (Freedman 1988, Freedman and He 1991). We define the *crossing number* of the arc as

$$\chi = \frac{\int \chi_1(\hat{x})\mathrm{d}\hat{x}}{\int \mathrm{d}\hat{x}}. \tag{2.2}$$

The crossing number of an arc will be some non-negative real number, in general not an integer. For a given length n, we define the *average crossing number* $\langle\chi\rangle_n$ to be the average value of χ, averaged over all n-step self-avoiding walks. We now prove a basic result about the average crossing number:

Theorem 2.3. $\langle\chi\rangle_n \to \infty$ with n, and the divergence is at least linear in n.

Proof. The proof relies on Kesten's pattern theorem (Kesten 1963). We construct a K-pattern consisting of three tight trefoils, such that no straight line hits more than two of the three-balls which are defined by the tight trefoils, and such that every projection of this pattern contains at least 3 crossings. Such a pattern is sketched in figure 2. Kesten's theorem asserts that there exists a positive number ϵ such that this pattern occurs at least ϵn times on all except exponentially few n-step self-avoiding walks, for sufficiently large values of n. For any walk in which this pattern appears ϵn times, the crossing number of any projection of that walk will be at least $3\epsilon n$, which proves the theorem. □

Searching for knotted arcs contained in a long self-avoiding walk is difficult for two reasons:

(1) Detecting that the union of the dual 3-cells to some sub-walk is a 3-ball is difficult.

(2) Given that the union of the dual 3-cells is indeed a 3-ball, computing the knot type of the resulting ball-pair is difficult, because the ambient 3-ball is itself distorted, and must be straightened out before a projection of the arc in the ball can be determined.

Searching for unfolded walks is also a fruitless task; they are not exponentially rare in the set of all walks, but they are rare indeed. For these reasons, neither of these approaches (searching for knotted arcs or searching for unfolded walks) is satisfactory from a computational point of view. However, they do suggest some alternative schemes which retain some of their advantages, and we describe two of these schemes which close up the walk to form a polygon.

For any given self-avoiding walk on Z^3 we can choose a direction at random and construct two parallel rays, whose origins are the two end-points of the walk and which are parallel to the prescribed direction. Almost all such rays will have irrational direction cosines and so will not pass through any of the vertices of Z^3. If we regard these two rays as meeting at the point at infinity we have a closed curve which will almost always be simple. Hence we can compute any of the usual measures of self-entanglement of a simple closed curve (such as $\Delta(-1)$) for this associated simple closed curve, and use this as a measure of the entanglement complexity of the walk itself. In general, the value will depend on the chosen direction so it will be convenient to average over all directions. Once again we have the problem that this construction can create or destroy entanglements. For instance, a knotted arc which occurs in the walk may not be retained as a knot in the resulting polygon, since one of the rays might pass through the 3-ball associated with the knotted arc and lead to an unknotted polygon.

Let F be any good measure of knot complexity, as defined in Soteros *et al* (1992). Such measures include number of prime factors, genus, bridge number minus one, crossing number, unknotting number, span of any non-trivial Laurent polynomial, log(order), minor index, braid index minus one, etc. Given any such measure F, we define a related measure ϕ_1 of a self-avoiding walk as follows. For a given direction $\hat{x}$, if we regard the rays as meeting at infinity, we obtain a polygon which we call $\mathcal{P}(\hat{x})$. We compute $F(\mathcal{P}(\hat{x}))$, and average this over all directions $\hat{x}$. More precisely,

$$\phi_1 = \frac{\int F(\mathcal{P}(\hat{x}))\mathrm{d}\hat{x}}{\int \mathrm{d}\hat{x}}. \tag{2.3}$$

We can now average over all n-step self-avoiding walks to produce $\langle\phi_1\rangle_n$. We now prove a basic result about these entanglement measures:

Theorem 2.4. $\langle\phi_1\rangle_n \to \infty$ with n, and the divergence is at least linear in n.

Proof. Kesten's pattern theorem asserts that the pattern sketched in figure 2 occurs with positive density on all but exponentially few sufficiently long self-avoiding walks. For each such walk, no matter which rays are added to the walk to form a polygon, the polygon will contain, with positive density, knotted arcs (of arbitrarily prescribed knot type) whose defining 3-balls do not intersect either of the added rays. The polygon is therefore always badly knotted, and divergence follows as in Soteros *et al* (1992). □

We define an *entanglement number* ζ_1 of a self-avoiding walk as follows. For a given direction $\hat{x}$, if we regard the rays as meeting at infinity, we obtain a polygon which we call $\mathcal{P}(\hat{x})$ and we associate an indicator function $\chi(\hat{x})$ with this polygon which is 1 if the polygon is knotted, and zero otherwise. We define the entanglement number of the walk as

$$\zeta_1 = \frac{\int \chi(\hat{x})\mathrm{d}\hat{x}}{\int \mathrm{d}\hat{x}}. \tag{2.4}$$

Theorem 2.5. All except exponentially few sufficiently long self-avoiding walks have entanglement number ζ_1 equal to one.

Proof. The result follows from theorem (2.4) by taking F to be, for instance, $\log(\Delta(-1))$. □

This result also gives information about the rate at which the average entanglement number $\langle\zeta_1\rangle_n$ approaches 1 as n goes to infinity and we state this result in the next theorem:

Theorem 2.6. The average entanglement number approaches unity exponentially rapidly as n goes to infinity.

Proof. Let c_n^0 be the number of n-step self-avoiding walks with entanglement number strictly less than 1. We write $\langle\zeta_1\rangle_n$ as

$$\langle\zeta_1\rangle_n = \frac{(c_n - c_n^0) + \sum_w' \zeta_1(w)}{c_n} \tag{2.5}$$

where the sum is over those walks w which have entanglement number strictly less than 1. Defining $\langle\zeta_1\rangle_n^0$ as the average value of ζ_1 *over those n-step walks with* $\zeta_1 < 1$, (2.5) can be written as

$$\langle\zeta_1\rangle_n = 1 - (1 - \langle\zeta_1\rangle_n^0)c_n^0/c_n\,. \tag{2.6}$$

Clearly $0 \leqslant \langle\zeta_1\rangle_n^0 < 1$ and, from theorem 2.5,

$$c_n^0/c_n = \mathrm{e}^{-\beta n + o(n)} \tag{2.7}$$

for some $\beta > 0$. This suffices to prove the theorem. □

An alternative and apparently simpler closing scheme is to join the ends of the walk with a line segment. There is an immediate problem, in that this line segment will in general pass through vertices of the lattice so that the resulting polygon might not be self-avoiding (and therefore not a simple closed curve). To avoid this problem we first add to the ends of the walk two parallel line segments (of equal length, less than unity) in a randomly chosen direction, and then join up the ends of the figure, forming (almost always) a simple closed polygon. Once we have obtained this polygon derived from the walk, given any good measure of knot complexity F, we can obtain an associated measure of entanglement complexity ϕ_2 by averaging F over all possible directions. As before, we have that $\langle\phi_2\rangle_n \to \infty$ with n. One can also define an entanglement number ζ_2 and prove that for n sufficiently large ζ_2 is almost surely equal to one. Moreover, the average entanglement number approaches 1 exponentially rapidly as n goes to infinity. The proofs are exactly analogous to those given for theorems 2.5 and 2.6.

With this scheme one can also prove a useful theorem which deals with the relative degree of entanglement of walks and polygons with the same value of n.

Theorem 2.7. If c_n^0 is the number of n-step self-avoiding walks with $\zeta_2 < 1$, then

$$c_n^0/c_n \geqslant \mathrm{e}^{-(\kappa-\kappa_0)n + o(n)}. \tag{2.8}$$

Proof. The set of n-step walks with $\zeta_2 < 1$ includes the set of n-step walks with $\zeta_2 = 0$ and for which the two vertices of degree 1 are unit distance apart. Every (rooted, directed) unknotted polygon with $n+1$ edges can be converted to a walk of this type by deleting the last edge. Hence

$$c_n^0 \geqslant 2(n+1)p_{n+1}^0 = \mathrm{e}^{\kappa_0 n + o(n)} \tag{2.9}$$

where we have made use of (1.2). Since $c_n = \mathrm{e}^{\kappa n + o(n)}$, (2.9) implies (2.8). □

3. Numerical results

We have compared the two approaches of *parallel rays* and *joined displaced end-points* by generating a sample of self-avoiding walks (with n edges, $n \leqslant 2000$), using a pivot algorithm (Lal 1969, Madras and Sokal 1987), and estimating the value of $\Delta(-1)$, the order of the knot in the resulting polygon. In figure 3 we plot $\langle\log\Delta(-1)\rangle_n$ against n for each of the two approaches, where the angular brackets indicate averages over

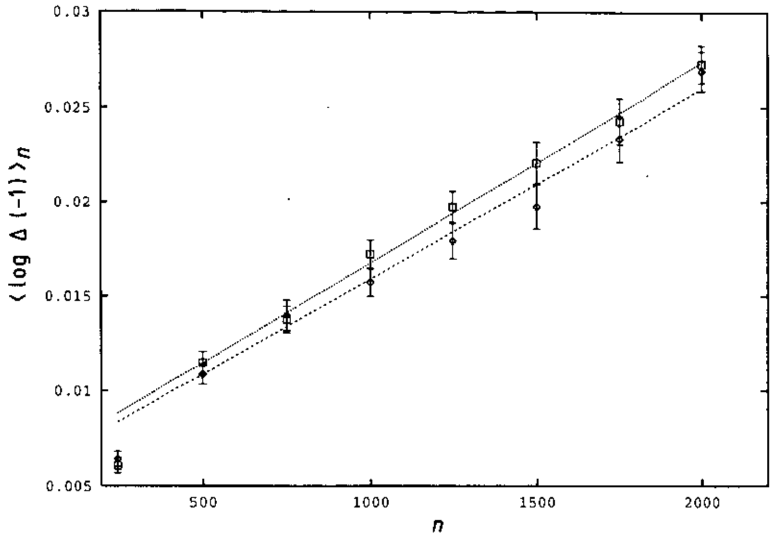

Figure 3. Monte Carlo estimates of $\langle\log\Delta(-1)\rangle_n$ as a function of the length of the self-avoiding walk. $\Delta(-1)$ is calculated for the polygon obtained by either the 'parallel rays' (◇) or the 'joined displaced end-points' (□) construction. The lines are least-squares fits to the data.

angles and over the walks in the sample. The data are consistent with a linear increase, though we are not able to rule out a faster than linear divergence. The lines represent weighted least-squares fits (assuming linear behaviour), neglecting the data point at $n = 250$. The estimates of the slopes are $(1.01 \pm 0.07) \times 10^{-5}$ and $(1.06 \pm 0.07) \times 10^{-5}$ so that the error bars overlap and there is no evidence that the two measures of complexity are diverging at different rates.

For the same samples of walks we have also estimated $\langle\zeta_1\rangle_n$ and $\langle\zeta_2\rangle_n$ by estimating ζ_1 and ζ_2 for each walk in the sample and averaging over the set of walks. We plot $\log(1 - \langle\zeta_1\rangle_n)$ and $\log(1 - \langle\zeta_2\rangle_n)$ against n in figure 4. The lines represent weighted least-squares fits to the data, not including the points at $n = 250$. The estimated values of the slopes are essentially identical, and correspond to

$$\langle\zeta\rangle_n = 1 - C\exp[(-8.9 \pm 0.5) \times 10^{-6} n] \tag{3.1}$$

where the constant C has different values for the two cases.

These methods can be used to characterize the dependence of the entanglement complexity on the 'solvent quality'. To mimic the effect of solvent quality we introduce a contact potential between neighbouring pairs of vertices as follows. For a given walk we count the number of pairs of vertices in the walk which are unit distance apart but not incident on a common edge. Let this number be m. We associate a (reduced) energy $-m\phi$ with the walk so that the walk has a weight proportional to $\exp(m\phi)$. Increasingly positive values of ϕ correspond to decreasing solvent quality. For a fixed value of n we have estimated $\langle\Delta(-1)\rangle_n$ as a function of ϕ and the results are shown (for $n = 1000$) in figure 5. The order increases rapidly as ϕ increases, suggesting that the entanglement complexity of linear polymers increases as the solvent becomes worse.

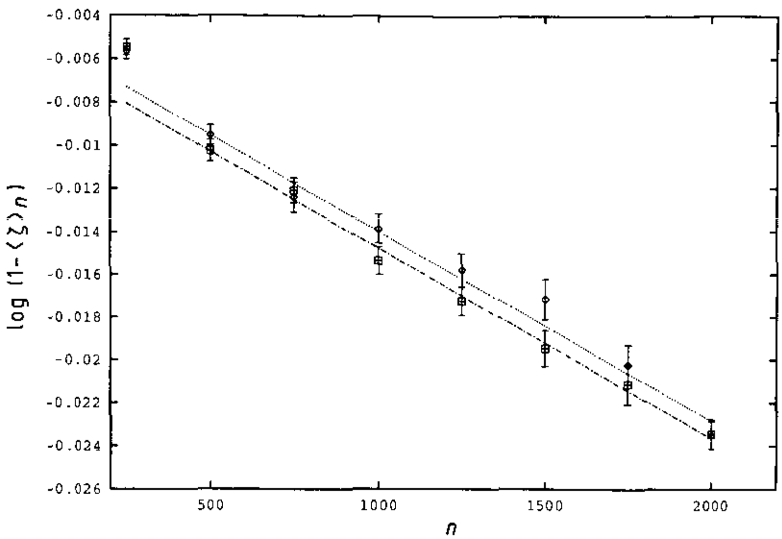

Figure 4. Monte Carlo estimates of $\log(1 - \langle\zeta_1\rangle_n)$ (◇) and $\log(1 - \langle\zeta_2\rangle_n)$ (□) for self-avoiding walks, as a function of the number of edges in the walk. The lines are least-squares fits to the data.

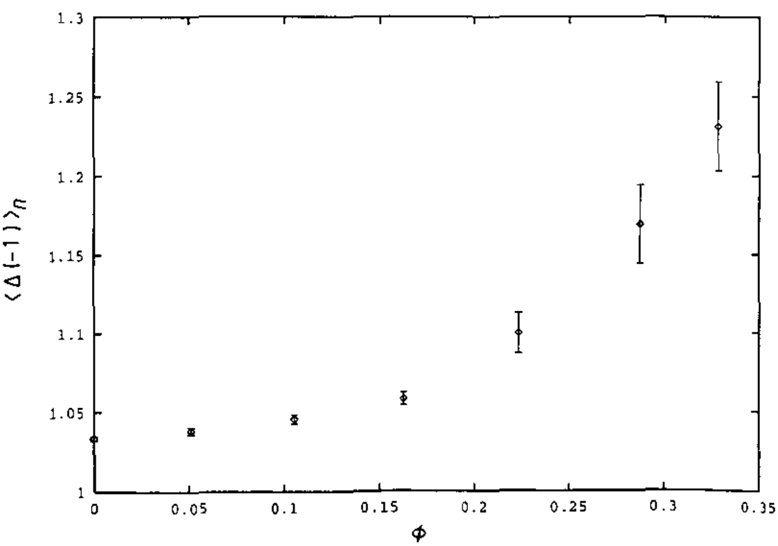

Figure 5. The dependence of entanglement complexity on solvent quality for walks with 1000 edges.

4. Discussion

The primary focus of this paper has been the investigation of two methods to describe and quantify the entanglement complexity of self-avoiding walks. Since walks are homeomorphic to 1-balls they are unknotted in the strict topological sense but we have shown rigorously that one can arrive at reasonable measures of the entanglement complexity by converting a walk to a polygon and using measures of the entanglement complexity of the polygon to characterize that of the walk. Using these measures we have shown that almost all sufficiently long walks are self-entangled and we have examined the n-dependence of these measures by a Monte Carlo calculation. In addition, by incorporating a pseudo-potential to reflect the effect of solvent quality on the conformational properties of a linear polymer, we have shown that the entanglement complexity increases rapidly as the solvent becomes worse.

Acknowledgments

This work was financially supported, in part by NSERC of Canada and in part by the NSF of the USA.

References

de Gennes P G 1984 *Macromol.* **17** 703

Delbruck M 1962 *Mathematical Problems in the Biological Sciences (Proc. Symp. Appl. Math. 14)* (Providence, RI: American Mathematical Society) p 55

Diao Y 1990 On knotting of randomly embedded polygons in R^3 *PhD Thesis* Florida State University

Diao Y, Pippenger N and Sumners D W 1992 in preparation

Edwards S F 1967 *Proc. Phys. Soc.* **91** 513

Freedman M H 1988 *J. Fluid Mech.* **194** 549

Freedman M H and He Z X 1991 *Ann. Math.* **134** 189

Frisch H L and Klempner D 1970 *Adv. Macromolecular Chem.* **2** 149

Frisch H L and Wasserman E 1961 *J. Am. Chem. Soc.* **83** 3789

Hammersley J M 1961 *Proc. Camb. Phil. Soc.* **57** 516

Hammersley J M and Welsh D J A 1962 *Q. J. Math. Oxford* **13** 108

Janse van Rensburg E J and Whittington S G 1990 *J. Phys. A: Math. Gen.* **23** 3573

Kendall W S 1979 *J. London Math. Soc.* (2) **19** 378

Kesten H 1963 *J. Math. Phys.* **4** 960

Koniaris K and Muthukumar M 1991 *J. Chem. Phys.* **95** 2873

Lal M 1969 *Mol. Phys.* **17** 57

Madras N and Sokal A D 1987 *J. Stat. Phys.* **47** 573

Michels J P J and Wiegel F W 1984 *Proc. R. Soc.* A **403** 269

Pippenger N 1989 *Discrete Applied Math.* **25** 273

Soteros C E, Sumners D W and Whittington S G 1992 *Math. Proc. Camb. Phil. Soc.* **111** 75

Sumners D W and Whittington S G 1988 *J. Phys. A: Math. Gen.* **21** 1689

—— 1990 *J. Phys. A: Math. Gen.* **23** 1471

Vologodskii A V, Lukashin A V, Frank-Kamenetskii M D and Anshelevich V V 1974 *Sov. Phys.–JETP* **39** 1059