The end-point distribution of self-avoiding walks on a crystal lattice

1971 J. Phys. A: Gen. Phys. 4 456

(http://iopscience.iop.org/0022-3689/4/4/007)

J. Phys. A: Gen. Phys., 1971, Vol. 4. Printed in Great Britain

The end-point distribution of self-avoiding walks on a crystal lattice

J. L. MARTIN and M. G. WATTS

Department of Physics, King's College, University of London, London, England

MS. received 30*th November* 1970

Abstract. A recursive relation connecting the numbers of self-avoiding walks on an arbitrary network with the numbers of certain related topologies is given, and is used to obtain exact end-point distributions of self-avoiding walks on a variety of Bravais lattices in two and three dimensions. The growth (proportional to n^{γ}) of the mean-square end-to-end distance of such walks of n steps is re-examined. We estimate (i) for the triangular lattice: $\gamma = 1{\cdot}488 \pm 0{\cdot}002$; (ii) for the face-centred cubic lattice: $\gamma = 1{\cdot}20 \pm 0{\cdot}01$. The first estimate is lower than $\gamma = 1{\cdot}5$ conjectured by some workers in this field; however, there are differences in assumptions of how the mean-square end-to-end distance ought to depend on n.

1. Introduction

The behaviour of finite self-avoiding walks on a regular lattice has attracted attention for several years, originally because it was thought that such walks may provide a reasonable model for a chain polymer in dilute solution. Wall and Hiller (1954), Hermans (1957), Casassa (1960) or Stockmeyer (1960) may be consulted for general reviews. Alternatively Fisher and Hiley (1961) give a brief summary up to 1960 and list other references. More recent papers include Domb (1963, 1969) and Domb *et al.* (1965). More recently the intrinsic combinatorial interest of the problem has attracted attention (see, for example, Hammersley 1957, Domb 1964, 1970, Domb and Hioe 1969a,b, Hiley and Sykes 1961, and Martin *et al.* 1967). In particular, on account of the fact that a walk is not permitted to visit the same lattice site more than once, the steps of the walk are correlated in an unobvious way. One of the more frequently discussed correlations is the distribution of the end-points of the n-step walks whose first point is at the origin: most of the work in this field has been numerical, and almost all the 'analytic results' are conjectures.

In this paper (§ 2) we derive recursive relations between the numbers of the self-avoiding walks of interest and the numbers of more complicated but less numerous auxiliary topologies. The prototype for such relations was given by Sykes (1961) and in a slightly different form by Fisher and Sykes (1959); the relations themselves have been known to us for some years. However, it is only recently that a rapid enough computer program has become available to provide the counts for the auxiliary topologies. Even so, the recurrence itself may contain thousands of terms, and the problem of managing it efficiently on a computer is not trivial. For example, it would not be easy to apply the method to a *random* network. However, the symmetries of the Bravais lattices result in a considerable simplification, and we have been able to obtain substantial new numerical results for the end-point distribution of n-step self-avoiding walks. Since they are extensive, the distributions themselves are not given here; they are to be found in Watts (1971). The new second moments

of the distributions are listed in table 1 for the usual Bravais lattices, and the results for the triangular lattice contribute to a fresh examination of the mean-square end-to-end distance of the walks on the close-packed lattices.

Table 1. New values for $c_n\rho_n$ for a variety of Bravais lattices

	n	$c_n\rho_n$†	c_n†
Square	19	2 23743 23436	3351 16620
	20	6 47029 14336	8976 97164
Triangular	11	4847 03142	176 68938
	12	23478 61440	753 55206
	13	1 12365 80322	3207 34686
	14	5 32257 80532	13627 91250
Simple cubic	14	11 27437 96632	44689 11678
	15	58 00522 60230	2 11751 46054
Body-centred cubic	9	4788 31624	356 52680
	10	36053 80256	2362 91096
	11	2 67798 55704	15680 49560
	12	19 67227 17504	1 03686 69992

† c_n = number of n-step self-avoiding walks, and ρ_n = mean square end-to-end distance of n-step self-avoiding walks. The already known values for c_n are also reproduced.

We have seen work by other authors (Hioe 1970, Chay 1971) concerned with recurrence relations similar to those of this paper. Their 'counting theorems' are couched in the language of generating functions, and though the basic concepts are similar in outline, the details are not appropriate to the analysis carried out here.

2. The recurrence relations

The problem is the following. We are given a network of *points* linked by *bonds*, the structure of which may be specified by the adjacency matrix η defined by

$$\eta(AB) = \eta(BA) = \begin{cases} 1, \text{ if points } A \text{ and } B \text{ are linked by a bond;} \\ 0, \text{ otherwise.} \end{cases}$$

An *n-step self-avoiding walk* on the network is a continuous sequence of n bonds visiting exactly $n+1$ distinct points, including the starting and ending points. The aim is to develop practicable methods of enumerating the n-step self-avoiding walks on the network.

One of the most rewarding techniques is based on noticing that the addition of a new bond to the end of a self-avoiding walk is normally more likely to result in a new self-avoiding walk than anything else; it therefore pays to count the *failures*, and to obtain the successes by using an appropriate recurrence relation. In this technique, it is necessary to classify the walks according to their behaviour near their ends; we therefore introduce the notation,

$$c_n(AB \ldots D, X \ldots Z)$$

to denote the number of n-step self-avoiding walks which begin by visiting the sequence of points $AB \ldots D$, and end by visiting the sequence $X \ldots Z$. (We shall say that these walks go *from* $AB \ldots D$ *to* $X \ldots Z$.)

The first recurrence relation is obtained by observing that, if $\eta(XY) = 1$, every n-step walk from AB to XY may be obtained by adding the bond XY to an $(n\text{-}1)$ step walk from AB to X. We therefore consider what happens when we add the bond XY to each of the $c_{n-1}(AB, X)$ walks from AB to X; as well as every one of the n-step walks from AB to XY we are likely to generate failures of three distinct types:

$$\begin{aligned} c_{n-1}(AB, X) = {} & c_n(AB, XY) \\ & + c_{n-1}(AB, YX) && (XY \text{ is an immediate reversal of the previous step}) \\ & + \delta_{AY} u_n(AB, XA) && (\text{if } Y \text{ is the same point as } A, \text{ we obtain a } \textit{loop} \text{ from } AB \text{ to } XA\text{: the number of these is } u_n(AB, XA)) \\ & + t_n(AB, XY) && (\text{if } Y \text{ is any other previous point of the walk, including } B, \text{ we obtain a } \textit{tadpole} \text{ from } AB \text{ to } XY\text{; number} = t_n(AB, XY)). \end{aligned} \tag{1}$$

Thus information about u_n and t_n will yield c_n, provided c_{n-1} is known.

A second relation may be obtained in a similar way. If $\eta(AB) = 1$, every tadpole with n bonds from AB to XY may be obtained either by adding the bond AB to a tadpole with $(n-1)$ bonds from B to XY or, if B and Y coincide, by adding the bond AB to a loop with $(n-1)$ bonds from B to XB. Thus

$$\begin{aligned} t_{n-1}(B, XY) + \delta_{BY} u_{n-1}(B, XB) = {} & t_n(AB, XY) \\ & \left.\begin{aligned} & + t_{n-1}(BA, XY) \\ & + \delta_{BY} u_{n-1}(BA, XB) \end{aligned}\right\} (\text{if } AB \text{ is an immediate reversal}) \\ & + \delta_{BY}\delta_{AX} u_{n-1}(AB, A) \quad (\text{if } B = Y, \text{ and } BA \text{ falls along } YX) \\ & + e_n(AB, XY) \quad (\text{all other failures}). \end{aligned} \tag{2}$$

The types of configuration which may appear in e_n are illustrated in figure 1. It is important that they are far less numerous than the walks with the same number of bonds.

The relation which forms the basis of the numerical work of this paper is obtained by eliminating $t_n(AB, XY)$ from (1) and (2); it is

$$\begin{aligned} & c_n(AB, XY) - \{c_{n-1}(B, XY) + c_{n-1}(AB, X) - c_{n-1}(AB, YX) - c_{n-1}(BA, XY)\} \\ & \quad + \{c_{n-2}(B, X) - c_{n-2}(B, YX) - c_{n-2}(BA, X) + c_{n-2}(BA, YX)\} \\ & \quad = e_n(AB, XY) - \delta_{AY} u_n(AB, XA) + \delta_{AX}\delta_{BY} u_{n-1}(AB, A) \end{aligned} \tag{3}$$

which is true for $n \geqslant 3$, $\eta(AB) = 1 = \eta(XY)$. In applying this recurrence, it is sufficient to know appropriate values of $e_n(AB, XY)$, since the terms in u_n contribute only when at least two of the points A, B, X and Y coincide, that is, only when $c_n(AB, XY)$ is known to be zero anyway.

The relation (3) is the first in a hierarchy of such relations. The form of the left side takes full account of the prohibition of immediate reversals, while the right side represents the fact that there are other sequences of bonds which violate the self-avoidance requirement. Apart from complexity, there is no reason not to account for

other prohibitions (such as triangular loops) in the recurrence on the left rather than in the correction on the right; in this way the hierarchy is obtained. However, the relations soon become too unwieldy to be of practical use.

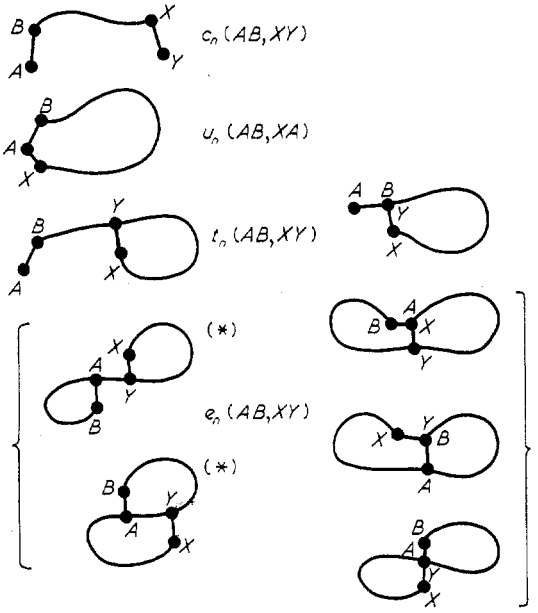

Figure 1. The types of configurations contributing to the values of c, u, t and e. Those which may be expected to appear when two of A, B, X and Y are equal are shown on the right. Only those marked (*) need to be known in applying the recurrence.

The derivation of (3) is valid for any network whatever. For some networks it is possible to 'condense' the relation in a way described in the appendix. In particular, if every point of the network has the same number $\sigma+1$ of bonds incident on it, we may sum (3) with A, B, X and Y ranging over all points of the network to obtain

$$c_n - 2\sigma c_{n-1} + \sigma^2 c_{n-2} = e_n - u_n + u_{n-1}$$

a relation first obtained and used by Sykes (1961) for the enumeration of self-avoiding walks on crystal lattices.

3. The computer program

It is not feasible to apply the recurrence relation to any useful problem without the aid of a computer, since it is necessary to handle many thousands of terms. Even so, there are problems of computer storage and time; special techniques are required, and what starts as a basically simple computation develops complex features which will not be discussed here. Therefore, though it would have been possible to write a program to apply the recurrence to a general network, we have thought it best to confine attention to the problem of the distribution of end-points of n-step self-avoiding walks on a selection of Bravais lattices. The extent of the new results may be inferred from table 1.

The evaluation of the starting values (c_n) and right sides (e_n) of the recurrence is done by a fast and efficient counting program which was developed for a very general type of topological enumeration: to count the number of distinct ways in which one given network may be embedded in another. This program is written in IBM 360 assembler language and is organized to guarantee that the counting speed is high; in the face-centred cubic lattice, for example, successful embeddings are counted at rates exceeding 10^6/minute. The program designed to execute the recurrence is written in FORTRAN, and includes the counting program as a subroutine.

Usually, the best way to use the recurrence is to choose a starting value of n for which the c_n are small enough to be evaluated directly, and to continue until the e_n become prohibitively large. However, when the number of walks to be counted is small (as it is for end-points near the edge of the distribution) it is more economical to count the walks directly, and to make no use of the recurrence.

4. Analysis of the results

Suppose that $c_n(A)$ is the number of n-step self-avoiding walks from the origin to the point A on a regular lattice, and that $r(A)$ is the cartesian distance of A from the origin. An important quantity is the *mean-square end-to-end distance* of the n-step self-avoiding walks:

$$\rho_n = \frac{\Sigma_A r^2(A) c_n(A)}{\Sigma_A c_n(A)}. \tag{4}$$

There is good numerical evidence (see, for example, Domb 1963) that the dominant behaviour of ρ_n for large n is given by

$$\rho_n \sim \text{const} \times n^{\gamma} \ (n \text{ large})$$

where γ depends only on the dimensionality of the lattice. Such a behaviour implies

$$\frac{\rho_{n+1}}{\rho_n} = 1 + \frac{\gamma}{n} + o\left(\frac{1}{n}\right).$$

Here we shall make a much stronger assumption that ρ_{n+1}/ρ_n may be expanded as a power series in n^{-1}; this is to say that ρ_n may be regarded as a coefficient in a generating function whose singularities are 'simple' (in a sense which we shall not discuss here). Thus we shall write

$$\frac{\rho_{n+1}}{\rho_n} = 1 + \frac{\gamma}{n}\left(1 + \frac{\alpha}{n}\right) + o\left(\frac{1}{n^2}\right). \tag{5}$$

Under these circumstances, estimates for γ are provided by the sequence

$$\gamma_n = \frac{(\rho_{n+1} - \rho_n)(\rho_n - \rho_{n-1})}{\rho_n^2 - \rho_{n+1}\rho_{n-1}} \tag{6}$$

$$= \gamma + o\left(\frac{1}{n}\right).$$

(Even without our strong assumption, the sequence of γ_n will still converge to γ, but not so rapidly.)

Table 2. Mean-square end-to-end length (ρ_n) of n-step walks on the triangular lattice

n	ρ_n	γ_n
7	14·2038857	
8	17·2421898	1·48667
9	20·4664849	1·48792
10	23·8664230	1·48597
11	27·4325000	1·48746
12	31·1572559	1·48743
13	35·0338794	1·48782
14	39·0564443	

Estimates for γ are provided by equation (6). (The end-point distributions for $n = 11$ to 14 are new.)

Table 3. Mean-square end-to-end length (ρ_n) of n-step walks on the face-centred cubic lattice, with estimates for γ

n	ρ_n	γ_n
5	6·3972150	
6	7·9498258	1·20574
7	9·5559603	1·20407
8	11·2091892	1·20264
9	12·9045503	1·20131
10	14·6380756	

The distributions up to $n = 10$ have previously been obtained by McKenzie (1967).

Values of ρ_n and γ_n are given in tables 2 and 3 for the triangular and face-centred cubic lattices. Based on these results, our estimates of γ are

$$\gamma(\text{triangular}): 1{\cdot}488 \pm 0{\cdot}002$$

$$\gamma(\text{fcc}) \quad\quad : 1{\cdot}20 \pm 0{\cdot}01.$$

The sequences for the loose-packed lattices exhibit the usual oscillation, but are not inconsistent with the conjecture that the true values of γ depend only on the dimensionality of the lattice. The full analysis for these lattices has not been carried out.

The estimate for the triangular lattice is distinctly lower than $\gamma = 1{\cdot}5$—proposed by Domb (1963) who, however, assumes a different dependence of ρ_n on n. This illustrates a difficulty which sometimes arises, that any extrapolation must be preceded by an *assumption* about the general behaviour of the extrapolate. Any such assumption may be supported by physical insight or by the numerical consistency of the extrapolation itself. In the case of the triangular lattice, we find numerical consistency which encourages the assumption of (5).

Acknowledgments

The assembler language program for counting embeddings was developed in 1967 when one of us (J.L.M.) was a guest at the Department of Physics, University of Alberta. The other author (M.G.W., who prepared the FORTRAN program which executes the recurrence) is grateful for the award of an SRC Research Studentship.

Appendix 1

In certain circumstances, the recurrence relations may be condensed in a useful way. Suppose it is possible to parcel the points of the network in to sets α, β, γ, ... such that for any two of the sets (possibly the same set),

$\eta_{\alpha\beta}$ = the number of bonds from a fixed point A in α to the points of β is *independent of the choice of* A in α.

The *trivial* parcelling is always possible: each point of the network is placed in a parcel by itself. In this case η becomes $\eta(AB)$, the adjacency matrix of the network. Not every network admits other parcellings; see, for example, figure 2.

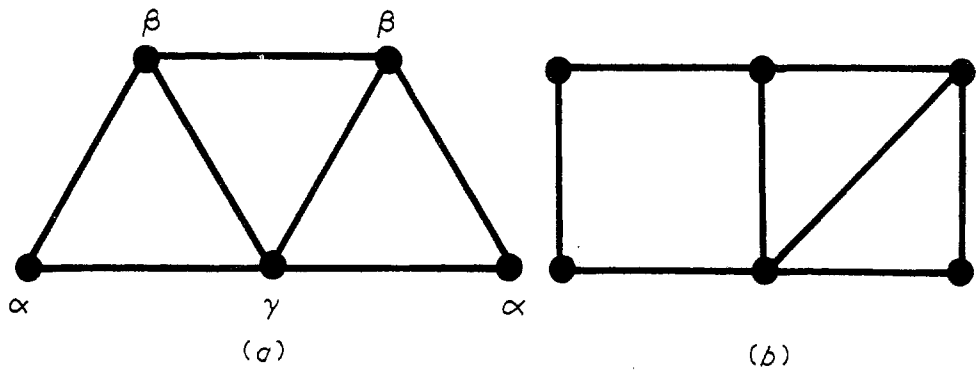

Figure 2. (*a*) a network admitting a parcelling into 3 sets, α, β and γ; (*b*) a network admitting only the trivial parcelling.

We define

$c_n(\alpha\beta, \xi\eta)$ = the number of n-step self-avoiding walks whose first two points are in α and β respectively, and whose last two points are in ξ and η respectively

with analogous definitions for u_n and e_n. By summing (3) over all A, B, X, Y in α, β, ξ, η respectively, with $\eta(AB) = 1 = \eta(XY)$, it is straightforward to show that

$$\begin{aligned} c_n(\alpha\beta, \xi\eta) - (\eta_{\beta\alpha} c_{n-1}(\beta, \xi\eta) + \eta_{\xi\eta} c_{n-1}(\alpha\beta, \xi) - c_{n-1}(\alpha\beta, \eta\xi) - c_{n-1}(\beta\alpha, \xi\eta)) \\ + (\eta_{\beta\alpha}\eta_{\xi\eta} c_{n-2}(\beta, \xi) - \eta_{\beta\alpha} c_{n-2}(\beta, \eta\xi) - \eta_{\xi\eta} c_{n-2}(\beta\alpha, \xi) + c_{n-2}(\beta\alpha, \eta\xi)) \\ = e_n(\alpha\beta, \xi\eta) - \delta_{\alpha\eta} u_n(\alpha\beta, \xi\alpha) + \delta_{\alpha\xi}\delta_{\beta\xi} u_{n-1}(\alpha\beta, \alpha). \end{aligned}$$

This recurrence is very similar to (3), and reduces to it for the trivial parcelling; for other parcellings it may be compact enough to be executed by hand.

In the special case of a network whose points all have the same degree, $\sigma+1$, all points may be placed in the *same* parcel α. If we write $c_n(\alpha\alpha, \alpha\alpha) = c_n$ simply, and $\eta_{\alpha\alpha} = \sigma+1$, we obtain

$$c_n - 2\sigma c_{n-1} + \sigma^2 c_{n-2} = e_n - u_n + u_{n-1}$$

first found by Sykes (1961).

Appendix 2

Complete tables of end-point distributions are too extensive to publish here; however, they have been obtained by the following authors who produce tables unless otherwise stated.

SQ	$N =$ 1 to 8	Orr (1947).
	1 to 12	Martin (1962): no tables.
	1 to 16	O'Flaherty (1961): no tables.
	17 & 18	Domb *et al.* (1965): no table for 17.
	1 to 20	Watts (1971).
T	$N =$ 1 to 9	Martin (1962): no tables.
	10	Hioe (1967): no table.
	1 to 14	Watts (1971).
SC	$N =$ 1 to 10	Martin (1962).
	11 to 13	Domb *et al.* (1965): no tables for 11 & 12.
	1 to 15	Watts (1971).
BCC	$N =$ 1 to 8	Martin (1962).
	1 to 12	Watts (1971).
FCC	$N =$ 1 to 7	Martin (1962).
	1 to 10	McKenzie (1967).

References

Casassa, E. F., 1960, *Ann. Rev. Phys. Chem.*, **11,** 477–500.
Chay, T. R., 1971, *J. chem. Phys.*, **54,** 1852–65.
Domb, C., 1963, *J. chem. Phys.*, **38,** 2957–63.
—— 1964, *J. R. Statist. Soc., Series* B, **26,** 367–97.
—— 1969, *Adv. chem. Phys.*, **15,** 229–59.
—— 1970, *J. Phys.* C: *Solid St. Phys.*, **3,** 256–84.
Domb, C., Gillis, J., and Wilmers, G., 1965, *Proc. Phys. Soc.*, **85,** 625–44.
Domb, C., and Hioe, F. T., 1969a, *J. chem. Phys.*, **51,** 1915–9.
—— 1969b, *J. chem. Phys.*, **51,** 1920–8.
Fisher, M. E., and Hiley, B. J., 1961, *J. chem. Phys.*, **34,** 1253–67.
Fisher, M. E., and Sykes, M. F., 1959, *Phys. Rev.*, **114,** 45–58.
Hammersley, J. M., 1957, *Proc. Camb. Phil. Soc.*, **53,** 642–5.
Hermans, J. J., 1957, *Ann. Rev. Phys. Chem.*, **11,** 477–500.
Hiley, B. J., and Sykes, M. F., *J. chem. Phys.*, **34,** 1531–7.
Hioe, F. T., 1967, *Ph.D. Thesis*, University of London.
—— 1970, *J. Phys. Chem.*, **74,** 4401–10.
McKenzie, D. S., 1967, *Ph.D. Thesis*, University of London.
Martin, J. L., 1962, *Proc. Camb. Phil. Soc.*, **58,** 92–101.
Martin, J. L., Sykes, M. F., and Hioe, F. T., 1967, *J. chem. Phys.*, **46,** 3478–81.
O'Flaherty, M. P., 1961, *Tech. Rpt. No.* 11, *Applied Mathematics and Statistics Laboratories* (California: Stanford University), 3 July.
Orr, W. J. C., 1947, *Trans. Farad. Soc.*, **43,** 12–27.
Stockmayer, W. H., 1960, *Makromol Chem.*, **35,** 54–79.
Sykes, M. F., 1961, *J. math. Phys.*, **2,** 52–62.
Wall, F. T., and Hiller, L. A., 1954, *Ann. Rev. Phys. Chem.*, **5,** 267–90.
Watts, M. G., 1971, *Ph.D. Thesis*, University of London.