Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Exact enumeration of self-avoiding walks

R D Schram[1,2], G T Barkema[1] and R H Bisseling[2]

[1] Institute for Theoretical Physics, Utrecht University, PO Box 80195, 3508 TD Utrecht, The Netherlands
[2] Mathematical Institute, Utrecht University, PO Box 80010, 3508 TA Utrecht, The Netherlands
E-mail: raouldschram@gmail.com, g.t.barkema@uu.nl and R.H.Bisseling@uu.nl

Received 12 April 2011
Accepted 9 June 2011
Published 27 June 2011

Online at stacks.iop.org/JSTAT/2011/P06019
doi:10.1088/1742-5468/2011/06/P06019

Abstract. A prototypical problem on which techniques for exact enumeration are tested and compared is the enumeration of self-avoiding walks. Here, we show an advance in the methodology of enumeration, making the process thousands or millions of times faster. This allowed us to enumerate self-avoiding walks on the simple cubic lattice up to a length of 36 steps.

Keywords: loop models and polymers, critical exponents and amplitudes (theory), exact results

J. Stat. Mech. (2011) P06019

 1742-5468/11/P06019+8$33.00

Contents

1. Introduction

According to renormalization group theory, the scaling properties of critical systems are insensitive to microscopic details and are governed by a small set of universal exponents. Polymers can be considered as critical systems in the limit where their length N (the number of chained monomers) grows [1]. For instance, the free energy F_N of an isolated polymer in a swollen phase behaves asymptotically as $\exp(-F_N) \equiv Z_N \approx A\mu^N N^\theta$. Here, the connectivity constant μ and the amplitude A are non-universal (model-dependent) quantities. The exponent θ, however, characterizing the leading correction to the scaling behavior, is believed to be universal; it is related to the entropic exponent $\gamma = \theta + 1 \approx 1.157$. The average squared distance between the end points of such polymers scales as $N^{2\nu}$, where $\nu \approx 0.588$ in three dimensions is also a universal critical exponent.

Universal exponents such as θ and ν can be measured most accurately in computer simulations of the most rudimentary models in the universality class of swollen polymers, which arguably is that of self-avoiding walks (SAWs) on a lattice. Estimates of these exponents can be obtained by counting the number Z_N of SAWs of all lengths up to $N_{\max}$, and calculating the sum P_N of their squared end-to-end extensions, which scales as $P_N \sim Z_N N^{2\nu}$. The exponents can then be obtained from

$$\theta = \frac{N^2-4}{4}\left[\log\frac{Z_N^2}{Z_{N+2}Z_{N-2}}\right] \tag{1}$$

and

$$\nu = \frac{N-1}{4}\left[\log\frac{P_{N+1}}{Z_{N+1}} - \log\frac{P_{N-1}}{Z_{N-1}}\right], \tag{2}$$

respectively, in the limit of increasing N. In equation (1), the values of N are taken a distance two apart, so that the formula involves either only even N or only odd N; this is more accurate than mixing even and odd values. Similar considerations lead to equation (2). The accuracy of the estimates improves significantly with increasing $N_{\max}$, but unfortunately at the expense of an exponentially growing number of walks. In two dimensions, various algorithmic improvements have allowed for the enumeration of all SAWs up to $N_{\max} = 71$ steps [2], but these methods cannot be used effectively in three

doi:10.1088/1742-5468/2011/06/P06019

dimensions, which is the most relevant dimensionality for practical purposes. Hence, to date, the enumeration of three-dimensional SAWs stops at $N_{\max} = 30$ steps [3].

Counting SAWs has a long history, see e.g. [4]. In a paper by Orr [5] from 1947, Z_N was given for all N up to $N_{\max} = 6$; these values were calculated by hand. In 1959, Fisher and Sykes [6] enumerated all SAWs in three dimensions (3D) up to $N_{\max} = 9$. More recently, in 1987 Guttmann [7] enumerated longer SAWs up to $N_{\max} = 20$ and extended this by one step in 1989 [8]. In 1992, MacDonald *et al* [9] reached $N_{\max} = 23$ and in 2000 MacDonald *et al* [10] reached $N_{\max} = 26$. In 2007, Clisby *et al* [3] reached $N_{\max} = 30$, which is currently the best result.

Here, we present the new length-doubling method which allowed us to reach $N_{\max} = 36$ using 50 000 h of computing time, a result that would have taken roughly fifty million hours with traditional methods, or alternatively we would have to wait another 20 years by Moore's law (which states that the number of transistors on a computer chip doubles every two years) before we could undertake the computation.

2. Length-doubling method

In the length-doubling method, we determine for each non-empty subset S of lattice sites the number $Z_N(S)$ of SAWs with length N and originating in the origin that visit the complete subset. Let $|S|$ denote the number of sites in S. The number Z_{2N} of SAWs of length $2N$ can then be obtained by the length-doubling formula

$$Z_{2N} = Z_N^2 + \sum_{S \neq \emptyset} (-1)^{|S|} Z_N^2(S). \tag{3}$$

This equation can be understood as follows. Let $N \geq 1$ be fixed. Let A_i be the set of pairs (v, w) of SAWs of length N that both pass through lattice point i. Here, a walk v starts in 0 and then passes through $v_1, \ldots, v_N$. Since the distance reached from the origin is at most N, there exist only finitely many non-empty sets A_i. Then, the total number of SAWs of length $2N$ equals

$$Z_{2N} = Z_N^2 - \left| \bigcup_i A_i \right|, \tag{4}$$

because every pair (v, w) of the Z_N^2 possible pairs can be used to construct a SAW of length $2N$, except if v and w intersect in a lattice point i. The resulting walk

$$(v, w) \equiv (v_{N-1} - v_N, \ldots, v_1 - v_N, -v_N, w_1 - v_N, \ldots, w_N - v_N) \tag{5}$$

of length $2N$ is obtained by connecting the two walks at 0 and translating the result over a distance $-v_N$. The new starting point 0 is then the translated old end point of v and the new end point is the translated old end point of w. Note that from a SAW of length $2N$ we can also create a non-intersecting pair (v, w) by using equation (5), so that indeed we have a bijection between such pairs and SAWs of length $2N$.

J. Stat. Mech. (2011) P06019

The inclusion–exclusion principle from combinatorics, see for instance [11, chapter 10], states that

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \cdots \\ + (-1)^{n+1}|A_1 \cap A_2 \cdots \cap A_n|, \qquad (6)$$

for the union of n sets A_i. We can apply this principle, noting that for a non-empty set $S = \{i_1, \ldots, i_r\}$ the intersection $A_{i_1} \cap \cdots \cap A_{i_r}$ has $Z_N^2(S)$ elements, where $Z_N(S)$ is defined as the number of SAWs of length N that pass through all the sites of S. The sign of the term corresponding to the set S in the expansion (6) is $(-1)^{r+1}$, where $r = |S|$. Substituting this in equation (6) and combining with equation (4) yields the length-doubling formula equation (3). The length-doubling method is illustrated in figure 1.

3. Application of the length-doubling formula

The usefulness of this formula lies in the fact that the numbers $Z_N(S)$ can be obtained relatively efficiently:

- Each SAW of length N is generated.
- For each SAW, each of the 2^N subsets S of the lattice sites is generated and the counter for each specific subset is incremented. Multiple counters for the same subset S must be avoided; this can be achieved by sorting the sites within each subset in an unambiguous way.
- As the last step, the squares of these counters are summed, with a positive and negative sign for subsets with an even and odd number of sites, respectively, as in equation (3).

With Z_N walks of length N, each visiting 2^N subsets of sites, the computational complexity is $\mathcal{O}(2^N Z_N) \sim (2\mu)^N$ times some polynomial in N which depends on implementation details. This compares favorably to generating all $Z_{2N} \sim \mu^{2N}$ walks of length $2N$, provided $2\mu < \mu^2$. This is clearly the case on the simple cubic lattice where $\mu \approx 4.684$.

A practical problem which is encountered already at relatively low N, is the memory requirement for storing the counters for all subsets. An efficient data structure to store these is based on a tree structure. The occurrence of a subset $\{a, b, c, d, e\}$, in which a, b, c, d, and e are site numbers ordered such that $a < b < c < d < e$, is stored in the path $a \to b \to c \to d \to e$, where a is directly connected to the root of the tree and e is a leaf.

We added two further refinements to the method sketched above. First, we exploit symmetry. Two subsets S_1 and S_2 which are related by symmetry will end up with the same counter. One can therefore safely keep track of the counter belonging to only one subset S out of each group of symmetry-related subsets. This reduces the memory requirement by a factor close to 48 (slightly less because of subsets with an inherent symmetry); in practice, the computational effort goes down by a similar factor.

The second refinement is tree splitting. Rather than computing the full tree, we split the tree into non-overlapping subtrees, using for instance as a criterion the value of the

J. Stat. Mech. (2011) P06019

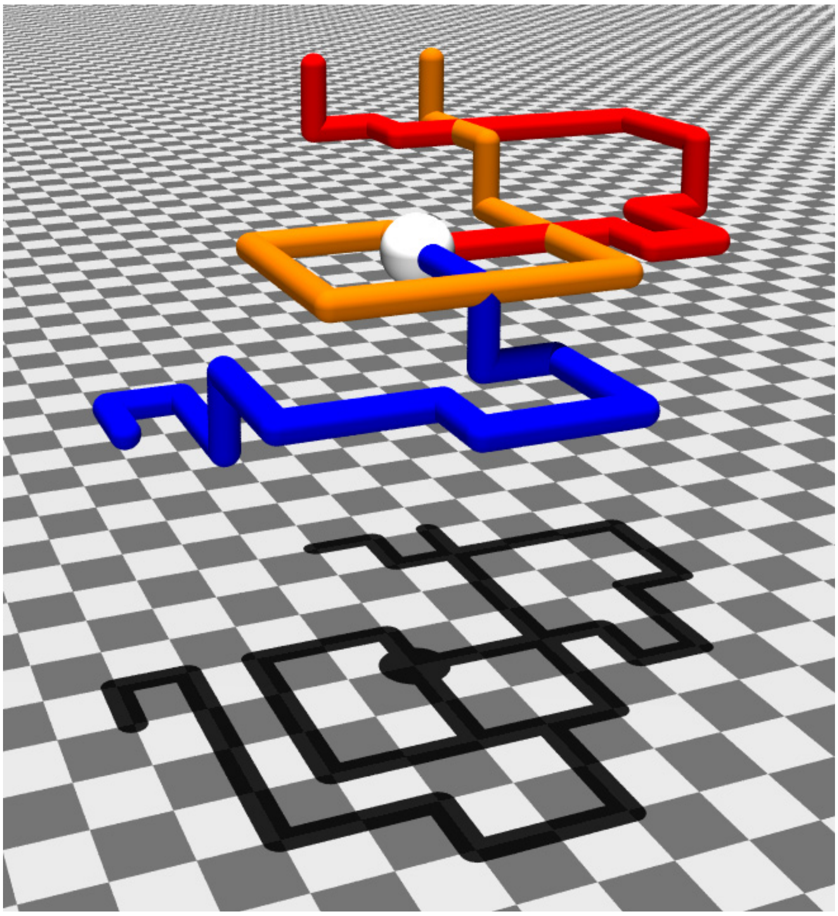

Figure 1. Illustration of the length-doubling algorithm, using a small subset of three walks of length $N = 18$. Ignoring intersections, there are $Z = 3$ candidates for SAWs of length 36: the blue-red, blue-orange and red-orange combinations. Ignoring double counting, Z should be reduced by three because of the intersections $a = (2, 3, 1)$, $b = (2, 0, 0)$ and $c = (0, -2, 0)$. Correcting for double counting because of the pair of sites $S = \{a, b\}$, the number of self-avoiding combinations is thus $3-3+1 = 1$. Indeed, only the red-blue combination is self-avoiding. Using a computer, we applied this approach to combinations of all walks of length $N = 18$.

J. Stat. Mech. (2011) P06019

site with the highest number. Another criterion is the subset size $|S|$. This splits up the summation in equation (3) into independent sums, which can be computed in parallel.

With the length-doubling method, it is also possible to compute the squared end-to-end distance, summed over all SAW configurations. The squared end-to-end distance for walks of length N is defined by

$$P_N = \sum_w ||w_N||^2, \tag{7}$$

where the sum is taken over all the SAWs of length N and $||w_N||$ is the Euclidean distance of the end point w_N of walk w from the origin.

The length-doubling formula for the squared end-to-end distance then becomes

$$P_{2N} = 2Z_N P_N + 2\sum_{S\neq\emptyset}(-1)^{|S|}(Z_N(S)P_N(S) - ||E_N(S)||^2). \tag{8}$$

Here, $P_N(S)$ is the total squared end-to-end distance for all walks of length N that pass through the complete set S and the extension $E_N(S)$ is defined as the sum of w_N for all such walks w. This formula can be understood again by using the inclusion–exclusion principle, but now generalized to add (squared) distances for sets A_i instead of just counting numbers of elements. The first term of the right-hand side of equation (8) is obtained by computing

$$\begin{aligned}\sum_{(v,w)} ||w_N - v_N||^2 &= \sum_{(v,w)}(||w_N||^2 + ||v_N||^2 - 2v_N \cdot w_N)\\ &= Z_N \sum_w ||w_N||^2 + Z_N \sum_v ||v_N||^2 - 2\Big(\sum_v v_N\Big)\cdot\Big(\sum_w w_N\Big)\\ &= 2Z_N P_N,\end{aligned} \tag{9}$$

where the inner product vanishes because of the symmetry between v and $-v$. For walks passing through S a similar derivation holds, but now the inner product does not vanish and instead gives rise to the term $||E_N(S)||^2$. Computing P_{2N} by this formula requires additional counters for each subset S, namely for the total extension in the x-, y- and z-directions, as well as for the total squared extension $P_N(S)$.

4. Results

With length-doubling, we obtained Z_N up to $Z_{36} = 2\,941\,370\,856\,334\,701\,726\,560\,670$, with a squared end-to-end extension of $P_{36} = 230\,547\,785\,968\,352\,575\,619\,933\,376$. All values of Z_N and P_N for $N \leq 36$ are given in table 1.

The behavior of Z_N and P_N for large N is expected to follow

$$\begin{aligned} Z_N &\approx A\mu^N N^{\gamma-1}(1 + c_1 N^{-\Delta}),\\ P_N &\approx D\mu^N N^{\gamma+2\nu-1}(1 + c_2 N^{-\Delta}).\end{aligned} \tag{10}$$

Here, we left out finite-size corrections distinguishing even and odd lengths.

A preliminary analysis by Clisby, using the direct fitting method as described in [3] and utilizing the recent estimate $\Delta = 0.53(1)$ [12], yields $\mu = 4.684\,0401(50)$, $\gamma = 1.156\,98(34)$, $\nu = 0.587\,72(17)$, $A = 1.2150(22)$ and $D = 1.2177(38)$. The estimates for μ and γ are significantly improved by the availability of the longer series, whereas estimates for ν, A and D are comparable in accuracy to [3]; the central estimates are shifted with respect to [3] largely due to the use of a different central value for Δ. The estimate for γ agrees with the literature value $\gamma = 1.1573(2)$ as obtained by Hsu *et al* [13] using the pruned-enriched Rosenbluth method.

In the near future, we will apply our new approach for exact enumeration to other lattices such as face-centered-cubic and body-centered-cubic, adapt it to count self-avoiding polygons and generalize it to various other models in polymer physics, such as confined and branched polymers, and to various other models in statistical physics.

J. Stat. Mech. (2011) P06019

Table 1. Enumeration results on the number of three-dimensional SAWs Z_N and the sum of their squared end-to-end distances P_N.

N	Z_N	P_N
1	6	6
2	30	72
3	150	582
4	726	4032
5	3534	25 566
6	16 926	153 528
7	81 390	886 926
8	387 966	4983 456
9	1853 886	27 401 502
10	8809 878	148 157 880
11	41 934 150	790 096 950
12	198 842 742	4166 321 184
13	943 974 510	21 760 624 254
14	4468 911 678	112 743 796 632
15	21 175 146 054	580 052 260 230
16	100 121 875 974	2966 294 589 312
17	473 730 252 102	15 087 996 161 382
18	2237 723 684 094	76 384 144 381 272
19	10 576 033 219 614	385 066 579 325 550
20	49 917 327 838 734	1933 885 653 380 544
21	235 710 090 502 158	9679 153 967 272 734
22	1111 781 983 442 406	48 295 148 145 655 224
23	5245 988 215 191 414	240 292 643 254 616 694
24	24 730 180 885 580 790	1192 504 522 283 625 600
25	116 618 841 700 433 358	5904 015 201 226 909 614
26	549 493 796 867 100 942	29 166 829 902 019 914 840
27	2589 874 864 863 200 574	143 797 743 705 453 990 030
28	12 198 184 788 179 866 902	707 626 784 073 985 438 752
29	57 466 913 094 951 837 030	3476 154 136 334 368 955 958
30	270 569 905 525 454 674 614	17 048 697 241 184 582 716 248
31	1274 191 064 726 416 905 966	83 487 969 681 726 067 169 454
32	5997 359 460 809 616 886 494	408 264 709 609 407 519 880 320
33	28 233 744 272 563 685 150 118	1993 794 711 631 386 183 977 574
34	132 853 629 626 823 234 210 582	9724 709 261 537 887 936 102 872
35	625 248 129 452 557 974 777 990	47 376 158 929 939 177 384 568 598
36	2941 370 856 334 701 726 560 670	230 547 785 968 352 575 619 933 376

J. Stat. Mech. (2011) P06019

Acknowledgments

Help with the series analysis by Nathan Clisby is gratefully acknowledged. Computations were carried out on the Huygens supercomputer at SARA in Amsterdam. We thank the Dutch National Computing Facilities Foundation (NCF) for providing us with computer resources under the project SH-174-10.

References

[1] de Gennes P-G, 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)

[2] Jensen I, *Enumeration of self-avoiding walks on the square lattice*, 2004 *J. Phys. A: Math. Gen.* **37** 5503
[3] Clisby N, Liang R and Slade G, *Self-avoiding walk enumeration via the lace expansion*, 2007 *J. Phys. A: Math. Theor.* **40** 10973
[4] Madras N and Slade G, 1993 *The Self-Avoiding Walk* (Boston, MA: Birkhäuser)
[5] Orr W J C, *Statistical treatment of polymer solutions at infinite dilution*, 1947 *Trans. Faraday Soc.* **43** 12
[6] Fisher M E and Sykes M F, *Excluded-volume problem and the Ising model of ferromagnetism*, 1959 *Phys. Rev.* **114** 45
[7] Guttmann A J, *On the critical behaviour of self-avoiding walks*, 1987 *J. Phys. A: Math. Gen.* **20** 1839
[8] Guttmann A J, *On the critical behaviour of self-avoiding walks: II*, 1989 *J. Phys. A: Math. Gen.* **22** 2807
[9] MacDonald D, Hunter D L, Kelly K and Jan N, *Self-avoiding walks in two to five dimensions: exact enumerations and series study*, 1992 *J. Phys. A: Math. Gen.* **25** 1429
[10] MacDonald D, Joseph S, Hunter D L, Moseley L L, Jan N and Guttmann A J, *Self-avoiding walks on the simple cubic lattice*, 2000 *J. Phys. A: Math. Gen.* **33** 5973
[11] van Lint J H and Wilson R M, 2001 *A Course in Combinatorics* 2nd edn (Cambridge: Cambridge University Press)
[12] Clisby N, *Accurate estimate of the critical exponent ν for self-avoiding walks via a fast implementation of the pivot algorithm*, 2010 *Phys. Rev. Lett.* **104** 055702
[13] Hsu H-P, Nadler W and Grassberger P, *Scaling of star polymers with 1–80 arms*, 2004 *Macromolecules* **37** 4658

J. Stat. Mech. (2011) P06019