Modern Birkhäuser Classics

Neal Madras
Gordon Slade

The Self-Avoiding Walk

Birkhäuser

Modern Birkhäuser Classics

Many of the original research and survey monographs, as well as textbooks, in pure and applied mathematics published by Birkhäuser in recent decades have been groundbreaking and have come to be regarded as foundational to the subject. Through the MBC Series, a select number of these modern classics, entirely uncorrected, are being re-released in paperback (and as eBooks) to ensure that these treasures remain accessible to new generations of students, scholars, and researchers.

The Self-Avoiding Walk

Neal Madras
Gordon Slade

Reprint of the 1996 Edition

Birkhäuser

Neal Madras
Department of Mathematics and Statistics
York University
Toronto, Ontario
Canada

Gordon Slade
Department of Mathematics
The University of British Columbia
Vancouver, British Columbia
Canada

Originally published in the series *Probability and Its Applications*

ISBN 978-1-4614-6024-4 ISBN 978-1-4614-6025-1 (eBook)
DOI 10.1007/978-1-4614-6025-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012952314

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.birkhauser-science.com)

Neal Madras
Gordon Slade

The Self-Avoiding Walk

Birkhäuser
Boston · Basel · Berlin

Neal Madras
Department of Mathematics
and Statistics
York University
Downsview, Ontario
Canada M3J 1P3

Gordon Slade
Department of Mathematics
and Statistics
McMaster University
Hamilton, Ontario
Canada L8S 4K1

Printed on acid-free paper

Birkhäuser

Hardcover edition, printed in 1993 by Birkhäuser Boston

ISBN 0-8176-3891-1
ISBN 3-7643-3891-1

Camera-ready text prepared in LATEX by the authors.
Printed and bound by Quinn-Woodbine, Woodbine, NJ.
Printed in the U.S.A.
9 8 7 6 5 4 3 2 1

To Joyce and Joanne

Contents

Preface

A self-avoiding walk is a path on a lattice that does not visit the same site more than once. In spite of this simple definition, many of the most basic questions about this model are difficult to resolve in a mathematically rigorous fashion. In particular, we do not know much about how far an n-step self-avoiding walk typically travels from its starting point, or even how many such walks there are. These and other important questions about the self-avoiding walk remain unsolved in the rigorous mathematical sense, although the physics and chemistry communities have reached consensus on the answers by a variety of nonrigorous methods, including computer simulations. But there has been progress among mathematicians as well, much of it in the last decade, and the primary goal of this book is to give an account of the current state of the art as far as rigorous results are concerned.

A second goal of this book is to discuss some of the applications of the self-avoiding walk in physics and chemistry, and to describe some of the nonrigorous methods used in those fields. The model originated in chemistry several decades ago as a model for long-chain polymer molecules. Since then it has become an important model in statistical physics, as it exhibits critical behaviour analogous to that occurring in the Ising model and related systems such as percolation. It is also of interest in probability theory as a basic example which does not respond well to standard probabilistic methods. Methods originating in mathematical physics and combinatorics have been more successful. Computer simulations have played an important role in formulating conjectures, and interesting computational and algorithmic issues have arisen in the process.

We have attempted to make this book as self-contained as possible. It should be accessible to graduate students in mathematics and to graduate students in physics and chemistry who are mathematically inclined.

Chapter 1 gives a general introduction to the basic questions and conjectures about the self-avoiding walk. The important notion of subadditivity is introduced in Section 1.2. Its relevance was pointed out by Hammersley and Morton (1954), and its interplay with concatenation is a recurring theme throughout the book.

Chapter 2 is devoted to a discussion of some nonrigorous and applied topics, namely scaling theory, the relation to polymers and the Flory argument, and the identification of the self-avoiding walk as a "zero-component" ferromagnet.

In 1962, Hammersley and Welsh proved an upper bound on the number of n-step self-avoiding walks which remains the best available bound in two dimensions. Their proof is given in Section 3.1. Shortly afterward, Kesten (1964) improved the Hammersley–Welsh bound in three or more dimensions. Kesten's bound is proven in Section 3.3; it remains the best bound in dimensions three and four. Bounds on the number of self-avoiding polygons are proven in Section 3.2. Subadditivity is the driving force in these arguments, implemented at times with considerable sophistication.

Chapter 4 is concerned with the decay of the subcritical two-point function. In particular, existence of a mass (or inverse correlation length) is proven, as well as existence of Ornstein–Zernike decay near a coordinate axis. This chapter also makes use of subadditivity in a fundamental way, using bridges and irreducible bridges. It has close connections with probabilistic renewal theory.

Chapters 5 and 6 are concerned with the self-avoiding walk above four dimensions and the recent proof by Hara and Slade (1992a,1992b) of mean-field behaviour in five or more dimensions. The main tool in the proof is the lace expansion of Brydges and Spencer (1985), which is the subject of Chapter 5. Section 5.5 indicates how the lace expansion can also be applied to lattice trees, lattice animals and percolation, and attempts to describe the expansions for the various models in a manner which emphasizes their similarities. The results in high dimensions are summarized in Section 6.1, before proving convergence of the lace expansion in Section 6.2. The convergence proof uses a number of estimates for ordinary random walk; these are given in Appendix A. In order to keep the convergence proof as simple as possible, we do not prove mean-field behaviour for the nearest-neighbour model in five or more dimensions, but rather consider the nearest-neighbour model in sufficiently high dimensions and the "spread-out" model above four dimensions; this allows us to present the simplest proof of convergence of the lace expansion to appear in print to date. Sections 6.3 to 6.8 show how convergence of the lace expansion leads to existence of critical exponents and other results stated in Section 6.1.

Chapter 7 is devoted to a proof of the pattern theorem of Kesten (1963) and a discussion of some of its consequences. These consequences are primarily in the form of ratio limit theorems for the number of n-step self-avoiding walks and related quantities.

Chapter 8 contains a short potpourri of additional results: upper bounds on the critical exponent α_{sing}, comments on self-avoiding walks in restricted geometries, construction of the infinite bridge, and some comments on the occurrence of knots in self-avoiding polygons.

Chapter 9 gives an extensive survey of various Monte Carlo algorithms, both static and dynamic, that have been used to simulate self-avoiding walks. Special attention is paid to the rigorous analysis of ergodicity properties and autocorrelation times of the algorithms.

Finally in Chapter 10 a brief discussion is given of four related topics: the Edwards model and weakly self-avoiding walk, the loop-erased self-avoiding random walk, intersection properties of simple random walks, and the "myopic" or "true" self-avoiding walk.

With the exception of Chapter 10 and the Appendices, most references to the literature are postponed to Notes which follow each chapter.

Enumerations of self-avoiding walks are proceeding at a rapid pace as computer technology advances. Appendix C gives tables of the number of self-avoiding walks, the square displacement, and the number of polygons, on the hypercubic lattice in dimensions two through six.

An overview of the key concepts, results, and methods of the book can be obtained from a reading of all of Chapter 1, together with Sections 3.1, 3.2, 4.1, 4.2, 5.1, 5.2, 5.4, 6.1, 6.2, 7.1, and 9.1. Any section not on the above list is rarely referenced outside its own chapter. In fact, the chapters of this book to a large extent can be read in any order, with the following exceptions: Chapter 1 should be read first; Chapter 3 should precede Chapter 4; Chapter 5 should precede Chapter 6; and Chapters 3, 4, and 7 should precede Chapter 8.

In view of our emphasis on rigorous results, we have omitted any description of such important ideas as the exact solution in two dimensions arising from connections with the Coulomb gas due to Nienhuis, and with conformal invariance in work of Duplantier and others. Nor have we attempted to describe the work on renormalization in the physics and polymer literature.

We have benefitted from the help of various people throughout the course of writing this book. Takashi Hara, Alan Sokal and Stu Whittington each read various portions of the manuscript and made numerous suggestions for improvements. Harry Kesten provided extensive notes for Section 3.3. Greg Lawler clarified several issues for us in Sections 10.2 and 10.3. Tony Guttmann kept us up-to-date with the latest world records in self-avoiding walk enumerations, and guided us through the related literature. We extend our thanks to all these, and to the many others who have offered advice in correspondence or conversation. NM is indebted to Alan Sokal and Stu Whittington for teaching him much about this field while collaborating on various enjoyable and fascinating projects. GS expresses special thanks to David Brydges for inspiring his initial interest in the lace expansion and suggesting that it might converge in five dimensions, and to Takashi Hara for the pleasure of four years of collaboration and for permission to include unpublished joint work. We gratefully acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada. Finally we offer our thanks and deep appreciation to Joyce Kruskal and Joanne Nakonechny for their encouragement, support, and tolerance.

Toronto and Hamilton
July 7, 1992

Chapter 1

Introduction

1.1 The basic questions

Imagine that you are standing at an intersection in the centre of a large city whose streets are laid out in a square grid. You choose a street at random and begin walking away from your starting point, and at each intersection you reach you choose to continue straight ahead or to turn left or right. There is only one rule: you must not return to any intersection already visited in your journey. In other words, your path should be self-avoiding. It is possible that you will lead yourself into a trap, reaching an intersection whose neighbours have all been visited already, but barring this disaster you continue walking until you have walked some large number N of blocks. There are two basic questions:

- How many possible paths could you have followed?
- Assuming that any one path is just as likely as any other, how far will you be on the average from your starting point?

These questions are straightforward enough, but the answers are only known for small values of N. It is widely accepted that a search for general exact formulas is an enormously difficult problem which lies beyond the reach of current methods. A less difficult question would be to ask for the asymptotic behaviour of the answers as N becomes very large, but this too is very hard. Physicists and chemists who are interested in this and related problems have applied a variety of methods and have produced many intriguing results, but a great deal of work is still needed to settle these issues in a mathematically rigorous way. In this book we will state some of the

N. Madras and G. Slade, *The Self-Avoiding Walk*,
Modern Birkhäuser Classics, DOI 10.1007/978-1-4614-6025-1_1,

results of nonrigorous work in the field, and describe the rigorous work in some detail.

At first glance one might expect that the easiest way to answer the above questions, at least approximately, would be to use a computer. Much numerical work has been done in this direction, and in Chapter 9 some of it will be discussed. Here too, however, the situation is not so easy: exact enumeration of all possible routes has been done to date only for $N \leq 34$, with further enumerations made difficult because of the exponential growth in the number of paths as N increases. Larger values of N can be studied by extrapolation of the exact enumeration data, or by Monte Carlo simulations.

There is no need to restrict the walk to a two-dimensional grid, and it is easy to generalize the above questions to general dimension d. It is also possible to generalize the problem by changing from a rectangular to a triangular or other type of grid. There is at least one case where the above questions can be easily answered, and this is the case of a one-dimensional walk. A self-avoiding walker in one dimension has no alternative but to continue travelling in the direction initially chosen, so there are exactly two paths for every value of N and the distance travelled is exactly N blocks. That was easy, but not very interesting. Higher dimensions provide a vastly richer structure.

In general, a self-avoiding walk takes place on a graph. A graph (more precisely, an undirected graph) is a collection of points, together with a collection of pairs of points known as *edges*. The basic example that will concern us most is the *d-dimensional hypercubic lattice* $\mathbf{Z}^d$. The points of this graph are the points of the d-dimensional Euclidean space $\mathbf{R}^d$ whose components are all integers, and the edges are given by the set of all unit line segments joining neighbouring points. The points will be referred to as *sites*, and the unit line segments as *nearest-neighbour bonds*. Sites will typically be denoted by letters such as u, v, x, y, and their components by subscripts: $x = (x_1, x_2, \ldots, x_d)$. The usual Euclidean dot product on $\mathbf{Z}^d$ will be written $x \cdot y = \sum_{i=1}^d x_i y_i$, and the Euclidean norm will be written $|x| = \sqrt{x \cdot x}$. We will also use the notation $||x||_p = (\sum_{i=1}^d x_i^p)^{1/p}$, and $||x||_\infty = \max\{|x_i| \colon i = 1, \ldots, d\}$.

An N-step self-avoiding walk ω on $\mathbf{Z}^d$, beginning at the site x, is defined as a sequence of sites $(\omega(0), \omega(1), \ldots, \omega(N))$ with $\omega(0) = x$, satisfying $|\omega(j+1) - \omega(j)| = 1$, and $\omega(i) \neq \omega(j)$ for all $i \neq j$. We write $|\omega| = N$ to denote the length of ω, and we denote the components of $\omega(j)$ by $\omega_i(j)$ ($i = 1, \ldots, d$). Let c_N denote the number of N-step self-avoiding walks beginning at the origin. By convention, $c_0 = 1$. Then the first of our basic questions above is asking for the value of c_N. More modestly, we could ask

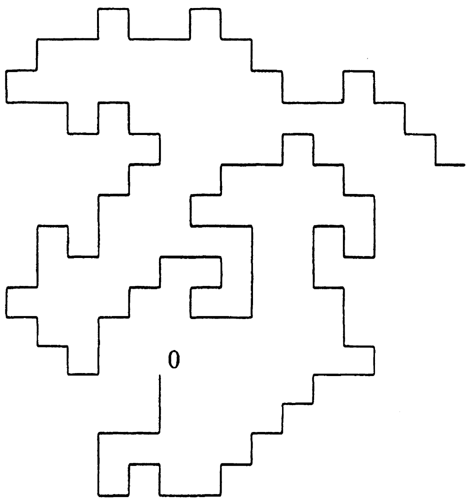

Figure 1.1: A two-dimensional self-avoiding walk with 115 steps.

for the asymptotic form of c_N as $N \to \infty$. It is easy to find the exact values of c_N (as a function of d) for very small values of N, for example $c_1 = 2d$, $c_2 = 2d(2d-1)$, $c_3 = 2d(2d-1)^2$, and $c_4 = 2d(2d-1)^3 - 2d(2d-2)$ (for c_4 the second term subtracts the contribution of squares to the first term). However, the combinatorics quickly become difficult as N increases and then soon become intractable. Tables in Appendix C give enumerations of c_N for dimensions two through six.

The simplest bounds on the behaviour of c_N are obtained as follows. An upper bound on c_N is given by the number of walks which have no immediate reversals, or in other words which never visit the same site at times i and $i+2$. Avoiding immediate reversals allows $2d$ choices for the initial step, and $2d-1$ choices for the $N-1$ remaining steps, for a total of $2d(2d-1)^{N-1}$. For a lower bound we simply count the number of walks in which each step is in one of the d positive coordinate directions. Such walks are necessarily self-avoiding. Thus we have

$$d^N \leq c_N \leq 2d(2d-1)^{N-1}. \tag{1.1.1}$$

To discuss the average distance from the origin after N steps, we need to introduce a probability measure on N-step self-avoiding walks. The measure that we shall use throughout this book is the uniform measure, which assigns equal weight c_N^{-1} to each N-step self-avoiding walk. It is worth noting that although we originally introduced the self-avoiding walk in terms of a walker moving in time, the uniform measure is a measure on paths of length N and does not define a stochastic process evolving in time (for example, a walk may be trapped and impossible to extend without introducing a self-intersection).

Denoting expectation with respect to the uniform measure by angular brackets, the average distance (squared) from the origin after N steps is then given by the *mean-square displacement*

$$\langle |\omega(N)|^2 \rangle = \frac{1}{c_N} \sum_{\omega:|\omega|=N} |\omega(N)|^2. \tag{1.1.2}$$

The sum over ω is the sum over all N-step self-avoiding walks beginning at the origin. Like c_N, the mean-square displacement can also be calculated by hand for very small values of N, but the combinatorics quickly become intractable as N increases. Enumerations are tabulated in Appendix C.

It is instructive to compare the behaviour of the self-avoiding walk with that of the simple random walk. An N-step simple random walk on $\mathbf{Z}^d$, starting at the origin, is a sequence $\omega = (\omega(0), \omega(1), \ldots, \omega(N))$ of sites with $\omega(0) = 0$ and $|\omega(j+1) - \omega(j)| = 1$, with the uniform measure on the set of all such walks. Without the self-avoidance constraint the situation is rather easy. Indeed, since each site has $2d$ nearest neighbours, the number of N-step simple random walks is exactly $(2d)^N$. To analyse the mean-square displacement, we represent the simple random walk in the following way. Let $\{X^{(i)}\}$ be independent and identically distributed random variables with $X^{(i)}$ uniformly distributed over the $2d$ (positive and negative) unit vectors. Then the position after N steps can be represented as the sum $S_N = X^{(1)} + X^{(2)} + \ldots + X^{(N)}$. Expanding $|S_N|^2$, the mean-square displacement is given by

$$\langle |S_N|^2 \rangle = \sum_{i,j=1}^{N} \langle X^{(i)} \cdot X^{(j)} \rangle. \tag{1.1.3}$$

For $i \neq j$, $\langle X^{(i)} . X^{(j)} \rangle = 0$, using independence and the fact that $\langle X^{(i)} \rangle = 0$. Since $\langle X^{(i)} . X^{(i)} \rangle = 1$, it follows that the mean-square displacement is equal to N. Similarly, if we consider a random walk in $\mathbf{Z}^d$ in which steps lie in a symmetric finite set $\Omega \subset \mathbf{Z}^d$ of cardinality $|\Omega|$, with each possible step equally likely, then the number of N-step walks is $|\Omega|^N$ and the mean-square displacement is $N\sigma^2$, where σ^2 is the mean-square displacement of a single step.

For the self-avoiding walk it is believed that there is exponential growth of c_N with power law corrections, unlike the pure exponential growth of the simple random walk. It is also believed that the mean-square displacement will not always be linear in the number of steps, in contrast to the diffusive behaviour of the simple random walk. These beliefs are in harmony with known properties of other models of statistical mechanics, and are supported by numerical and nonrigorous calculations. The conjectured behaviour of c_N and $\langle|\omega(N)|^2\rangle$ is thus

$$c_N \sim A\mu^N N^{\gamma-1} \tag{1.1.4}$$

and

$$\langle|\omega(N)|^2\rangle \sim DN^{2\nu}, \tag{1.1.5}$$

where A, D, μ, γ and ν are dimension-dependent positive constants. We shall refer to μ as the *connective constant*, and γ and ν are examples of *critical exponents*. In four dimensions the above two relations should be modified by logarithmic factors; see (1.1.13) and (1.1.14) below. Here $f(N) \sim g(N)$ means that f is asymptotic to g as $N \to \infty$:

$$\lim_{N\to\infty} \frac{f(N)}{g(N)} = 1.$$

For ordinary random walk (1.1.4) and (1.1.5) hold with $\gamma = 1$ and $\nu = 1/2$, both for the nearest-neighbour and more general walks.

In the next section the existence of the limit

$$\mu = \lim_{N\to\infty} c_N^{1/N} \tag{1.1.6}$$

will be proven, which is the first step in justifying (1.1.4). The simple bounds of (1.1.1) then immediately imply that

$$d \leq \mu \leq 2d - 1. \tag{1.1.7}$$

The exact value of μ is not known for the hypercubic lattice in any dimension $d \geq 2$, although for the honeycomb lattice in two dimensions there is nonrigorous evidence that $\mu = \sqrt{2+\sqrt{2}}$. Improvements to (1.1.7) will be discussed in the next section. For high dimensions it is known that as $d \to \infty$

$$\mu = 2d - 1 - \frac{1}{2d} - \frac{3}{(2d)^2} + O\left(\frac{1}{(2d)^3}\right); \tag{1.1.8}$$

references are given in the Notes. In fact Fisher and Sykes (1959) established the coefficients in the $1/d$ expansion up to and including order d^{-4}, although there is no rigorous control of their error term. Intuitively (1.1.8) says that in high dimensions the principal effect of the self-avoidance constraint is to rule out immediate reversals.

Concerning γ, we will show in Section 1.2 that $c_N \geq \mu^N$ and hence $\gamma \geq 1$ in all dimensions. There is still no proof, however, that γ is finite in two, three or four dimensions, where the best bounds are

$$c_N \leq \begin{cases} \mu^N \exp[KN^{1/2}] & d = 2 \\ \mu^N \exp[KN^{2/(2+d)} \log N] & d = 3,4 \end{cases} \tag{1.1.9}$$

for a positive constant K; these bounds will be discussed in Sections 3.1 and 3.3. In Chapter 6 we will describe a proof that (1.1.4) holds with $\gamma = 1$ for $d \geq 5$. In addition to characterizing the asymptotic behaviour of c_N, the exponent γ provides a measure of the probability that two N-step self-avoiding walks starting at the same point do not intersect. In fact, this probability is equal to c_{2N}/c_N^2, and assuming (1.1.4) we have

$$\frac{c_{2N}}{c_N^2} \sim \frac{2^{\gamma-1}}{A} N^{1-\gamma}. \tag{1.1.10}$$

If $\gamma > 1$ then this probability goes to zero as $N \to \infty$, while if $\gamma = 1$ it remains positive. For the simple random walk the analogous probability is known to remain positive as $N \to \infty$ for $d > 4$, and roughly speaking to go to zero like $(\log N)^{-1/2}$ for $d = 4$ and as an inverse power of N for $d = 2, 3$. A survey of the simple random walk results is given in Section 10.3.

Intuitively it is to be expected that the repulsive interaction of the self-avoiding walk will tend to drive the endpoint of the walk away from the origin faster than for simple random walk, or in other words that $\nu \geq 1/2$. However it is still an open question to prove that this "obvious" inequality $\langle |\omega(N)|^2 \rangle \geq CN$ holds in all dimensions. On the other hand, bounding $|\omega(N)|^2$ above by N^2 in (1.1.2) gives the upper bound $\langle |\omega(N)|^2 \rangle \leq N^2$, or $\nu \leq 1$. This bound is optimal in one dimension, but seems far from optimal in two or more dimensions. No upper bound of the form $CN^{2-\epsilon}$ $(C, \epsilon > 0)$, or in other words $\nu < 1$, has been proven for dimensions two, three or four, however. For $d \geq 5$ it has been proved that $\nu = 1/2$; this proof will be described in Chapter 6. It will also be shown that for high dimensions the diffusion constant D is strictly greater than the simple random walk value of 1. Thus in high dimensions the self-avoiding walk does move away from the origin more quickly than the simple random walk, but only at the level of the diffusion constant and not at the level of the exponent ν. The tendency of the self-avoiding walk to move away from the origin more quickly than the simple random walk should become less pronounced as the dimension increases, and hence it is to be expected that ν is a nonincreasing function of the dimension.

The critical exponents γ and ν are believed to be dimension dependent, but independent of the type of allowed steps (as long as there are only finitely many possible steps and the allowed steps are symmetric) or even of the type of lattice—the exponents are believed, for example, to be the same for the square and triangular lattices. This lack of dependence on the detailed definition of the model is known as *universality*, and models with the same exponents are said to be in the same *universality class*. The *connective constant* μ appearing in (1.1.4) represents the effective coordination number of the lattice and is not universal—it depends on the details of the allowed steps and the underlying lattice, as well as the dimension d.

It seems clear that in high dimensions the self-avoiding walk should be closer to the simple random walk than in low dimensions, since a simple random walk is less likely to intersect itself in high dimensions. Four dimensions plays a special role: for simple random walk the expected time of the first return to the origin, conditioned on the event that this return occurs, is finite for $d > 4$; this suggests that above four dimensions self-avoidance is a short-range effect rather than a long-range one, and hence that it will not affect the critical exponents. In addition, as mentioned above, the probability that two independent simple random walks of length N do not intersect remains bounded away from zero as $N \to \infty$ for $d > 4$, but not for $d \leq 4$.

The conjectured values of γ and ν are as follows:

$$\gamma = \begin{cases} \frac{43}{32} & d = 2 \\ 1.162... & d = 3 \\ 1 \text{ with logarithmic corrections} & d = 4 \\ 1 & d \geq 5 \end{cases} \qquad (1.1.11)$$

$$\nu = \begin{cases} \frac{3}{4} & d = 2 \\ 0.59... & d = 3 \\ \frac{1}{2} \text{ with logarithmic corrections} & d = 4 \\ \frac{1}{2} & d \geq 5 \end{cases} \qquad (1.1.12)$$

Currently the only rigorous results which prove power law behaviour and confirm the conjectured values of γ and ν are for $d \geq 5$. These are discussed in detail in Chapter 6. The conjectured logarithmic corrections to γ and ν in four dimensions, predicted by the renormalization group, are given by:

$$c_N \sim A\mu^N[\log N]^{1/4}, \quad d = 4 \qquad (1.1.13)$$

$$\langle|\omega(N)|^2\rangle \sim DN[\log N]^{1/4}, \quad d = 4. \qquad (1.1.14)$$

Equations (1.1.11) to (1.1.14) are typical of what is found for other statistical mechanical models, such as the Ising model or percolation. A common feature is the existence of a certain dimension, the so-called *upper critical* *dimension*, at which there are logarithmic corrections to critical exponents and above which all critical exponents are dimension independent and are given by the corresponding critical exponents for a simpler model, known as the *mean-field*[1] model. For the self-avoiding walk the mean-field model is the simple random walk and the simple random walk critical exponents are sometimes referred to as the mean-field exponents.

The rational values for two dimensions given in (1.1.11) and (1.1.12) come from a nonrigorous exact solution of the $O(N)$ spin model which includes the self-avoiding walk as the special case $N = 0$ (see Section 2.3). This remarkable work exploits a connection between the $O(N)$ model and the Coulomb gas and uses the renormalization group. From a different approach, nonrigorous conformal invariance arguments reproduce the same rational values. There is no analogous exact solution in three dimensions, and the $d = 3$ values given in (1.1.11) and (1.1.12) are from numerical results and field-theoretic calculations using the ϵ-expansion. References for these topics are given in the Notes.

An early conjecture for the values of ν was made by Flory, and will be discussed in Section 2.2. The Flory exponents are given by $\nu_{Flory} = 3/(2+d)$ for $d \leq 4$ and $\nu_{Flory} = 1/2$ for $d > 4$. This agrees with Equation (1.1.12) for $d = 2$ and $d \geq 4$ (apart from the logarithmic correction when $d = 4$), and comes very close for $d = 3$. The exact Flory value $\nu_{Flory} = 3/5$ in three dimensions has been ruled out by numerical work, however.

1.2 The connective constant

If (1.1.4) correctly represents the behaviour of c_N for large N, then the limit

$$\mu = \lim_{N\to\infty} c_N^{1/N} \tag{1.2.1}$$

must exist. One purpose of this section is to prove the existence of this limit as a simple consequence of a subadditive property of $\log c_N$. It then follows immediately from (1.1.1) that

$$d \leq \mu \leq 2d - 1. \tag{1.2.2}$$

The proof involves the notion of concatenation of two self-avoiding walks.

[1]This terminology has its origin in the Ising model. For the Ising model the upper critical dimension is also four, and above four dimensions critical exponents are given by the exactly solvable model in which a spin interacts with the *average* of all the other spins. References are given in the Notes.

Definition 1.2.1 *The* concatenation $\omega^{(1)} \circ \omega^{(2)}$ *of an M-step self-avoiding walk $\omega^{(2)}$ to an N-step self-avoiding walk $\omega^{(1)}$ is the $(N+M)$-step walk ω, which in general need not be self-avoiding, given by*

$$\begin{array}{lcll} \omega(k) & = & \omega^{(1)}(k), & k = 0, \ldots, N \\ \omega(k) & = & \omega^{(1)}(N) + \omega^{(2)}(k-N) - \omega^{(2)}(0), & k = N+1, \ldots, N+M. \end{array}$$

The product $c_N c_M$ is equal to the cardinality of the set of $(N+M)$-step simple random walks which are self-avoiding for the initial N steps and the final M steps, but which may not be completely self-avoiding. This can be seen by concatenations of M-step walks to N-step walks, and implies that

$$c_{N+M} \leq c_N c_M. \tag{1.2.3}$$

In fact equality holds in (1.2.3) only if N or M is zero, since otherwise there will be at least one M-step walk whose concatenation with a given N-step walk fails to be self-avoiding. Taking logarithms in (1.2.3) shows that the sequence $\{\log c_n\}$ is *subadditive:*

$$\log c_{N+M} \leq \log c_N + \log c_M. \tag{1.2.4}$$

The existence of the limit (1.2.1) is a consequence of (1.2.4) and the following standard result; this was first observed by Hammersley and Morton (1954).

Lemma 1.2.2 *Let $\{a_n\}_{n\geq 1}$ be a sequence of real numbers which is subadditive, i.e., $a_{n+m} \leq a_n + a_m$. Then the limit $\lim_{n\to\infty} n^{-1}a_n$ exists in $[-\infty, \infty)$ and is equal to*

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n}. \tag{1.2.5}$$

Proof. It suffices to show that

$$\limsup_{n\to\infty} \frac{a_n}{n} \leq \frac{a_k}{k} \tag{1.2.6}$$

for every k, since taking the $\liminf_{k\to\infty}$ in (1.2.6) gives existence of the limit, and then (1.2.5) can be seen by taking the $\inf_{k\geq 1}$ in (1.2.6).

To prove (1.2.6), we fix k and let

$$A_k = \max_{1\leq r\leq k} a_r. \tag{1.2.7}$$

Given a positive integer n we let j denote the largest integer which is strictly less than n/k. Then $n = jk + r$ for some integer r with $1 \leq r \leq k$. Using subadditivity, we have

$$a_n \leq j a_k + a_r \leq \frac{n}{k} a_k + A_k. \tag{1.2.8}$$

Dividing by n and taking the $\limsup_{n\to\infty}$ then gives (1.2.6).

Equation (1.2.5) shows that $\lim_{n\to\infty} n^{-1}a_n < \infty$. In general, the possibility that the limit equals $-\infty$ cannot be excluded, as is illustrated by the example of $a_n = -n^2$. For many applications, however, this is ruled out by an *a priori* bound such as $a_n \geq 0$. □

Together with (1.2.4), Lemma 1.2.2 implies the existence of the limit $\log\mu \equiv \lim_{N\to\infty} N^{-1}\log c_N$, and hence gives (1.2.1). In fact (1.2.5) shows more:

$$\log\mu = \inf_{N\geq 1} N^{-1}\log c_N, \tag{1.2.9}$$

and hence

$$\mu^N \leq c_N, \quad N \geq 1. \tag{1.2.10}$$

This inequality can be summarized by the statement $\gamma \geq 1$, where γ is as introduced in (1.1.4), although strictly speaking we do not know that γ exists. Equation (1.2.10) also yields $\mu \leq c_N^{1/N}$. This gives a sequence of upper bounds for μ, but they converge to μ very slowly. A better bound is

$$\mu \leq \left(\frac{c_N}{c_1}\right)^{1/(N-1)}, \quad N \geq 2. \tag{1.2.11}$$

References for this and other improvements are given in the Notes.

Another sequence of upper bounds for μ can be obtained by considering walks which are self-avoiding only over a finite time scale or *memory* τ. We define $c_{N,\tau}$ to be the number of N-step walks ω beginning at the origin, for which $\omega(i) \neq \omega(j)$ whenever $0 < |i-j| \leq \tau$. Self-intersections occurring after an interval of more than τ steps are permitted. For example, $c_{N,2} = 2d(2d-1)^{N-1}$ for $N \geq 1$, since memory $\tau = 2$ simply rules out immediate reversals. For $\tau \geq N$, $c_{N,\tau} = c_N$. Memory $\tau = 0$ corresponds to the simple random walk.

The sequence $\{\log c_{N,\tau}\}_{N=1}^{\infty}$ is subadditive for every τ (for the same reason that $\{\log c_N\}_{N=1}^{\infty}$ is), and hence by Lemma 1.2.2 there is a μ_τ such that

$$\mu_\tau = \lim_{N\to\infty} c_{N,\tau}^{1/N} = \inf_{N\geq 1} c_{N,\tau}^{1/N}. \tag{1.2.12}$$

Since $c_{N,\tau} \geq c_N$, μ_τ provides an upper bound for μ. The next lemma shows that this sequence of upper bounds converges monotonically to μ.

Lemma 1.2.3 *$\mu_\tau \searrow \mu$ as $\tau \to \infty$.*

Proof. For $\sigma \leq \tau$, $c_{N,\sigma} \geq c_{N,\tau}$ and hence $\mu_\sigma \geq \mu_\tau$. By (1.2.12), $\mu_\tau \leq c_{N,\tau}^{1/N}$ for all N, τ. Taking $N = \tau$ gives

$$\mu \leq \mu_\tau \leq c_{\tau,\tau}^{1/\tau} = c_\tau^{1/\tau}. \tag{1.2.13}$$

Taking the limit $\tau \to \infty$ and using (1.2.1) gives the desired result. □

The connective constant for the walk with memory $\tau = 4$ was shown in Fisher and Sykes (1959) to be given by the largest root of the cubic equation

$$\theta^3 - 2(d-1)\theta^2 - 2(d-1)\theta - 1 = 0. \tag{1.2.14}$$

For $d = 2$ this gives $\mu_4(2) = 2.8312$, where we have made the dimension dependence explicit by writing $\mu_\tau(d)$.

A number of investigations into the self-avoiding walk have approached the problem via the limit of finite memory walks as the memory goes to infinity. This approach was used in particular by Brydges and Spencer (1985) in applying their lace expansion to study weakly self-avoiding walk for $d > 4$, and will be adopted in Section 6.8 to obtain an upper bound in high dimensions on $c_N(0,x)$, the number of N-step self-avoiding walks which begin at the origin and end at x.

A lower bound on μ can be obtained in terms of *bridges*.

Definition 1.2.4 *An N-step* bridge *is defined to be an N-step self-avoiding walk ω whose first components satisfy the inequality*

$$\omega_1(0) < \omega_1(i) \leq \omega_1(N)$$

for $1 \leq i \leq N$. The number of N-step bridges starting at the origin is denoted b_N. By convention, $b_0 = 1$.

The concatenation of two bridges will always yield another bridge, so

$$b_M b_N \leq b_{M+N}. \tag{1.2.15}$$

Hence $\{-\log b_n\}$ is subadditive and so by Lemma 1.2.2 the limit

$$\mu_{Bridge} \equiv \lim_{n\to\infty} b_n^{1/n} = \sup_{n\geq 1} b_n^{1/n} \tag{1.2.16}$$

exists. Clearly $b_n \leq c_n$. Therefore $\mu_{Bridge} \leq \mu$, and so by (1.2.16)

$$b_N^{1/N} \leq \mu_{Bridge} \leq \mu. \tag{1.2.17}$$

In Section 3.1 it will be shown that in fact $\mu_{Bridge} = \mu$. Although the lower bound (1.2.17) is very slowly convergent, a more sophisticated use of bridges leads to better lower bounds. References can be found in the Notes at the end of this chapter.

We conclude this section with a table showing the current best rigorous upper and lower bounds on μ, together with estimates of the precise value, for the hypercubic lattice in dimensions $d = 2, 3, 4, 5, 6$.

d	lower bound	estimate	upper bound
2	2.61987^a	$2.6381585 \pm 0.0000010^d$	2.69576^b
3	4.43733^c	4.6839066 ± 0.0002^e	4.756^b
4	6.71800^c	6.7720 ± 0.0005^f	6.832^b
5	8.82128^c	8.83861^g	8.881^b
6	10.871199^c	10.87879^g	10.903^b

Table 1.1: Current best rigorous upper and lower bounds on the hypercubic lattice connective constant μ, together with estimates of actual values.
a) Conway and Guttmann (to be published), b) Alm (1992), c) Hara and Slade (1992b), d) Guttmann and Enting (1988), e) Guttmann (1987), f) Guttmann (1978), g) Guttmann (1981).

1.3 Generating functions

A common tool for understanding the behaviour of a sequence is its generating function. The generating function of the sequence $\{c_N\}$ is defined by

$$\chi(z) = \sum_{N=0}^{\infty} c_N z^N = \sum_{\omega} z^{|\omega|}. \tag{1.3.1}$$

The sum over ω is the sum over all self-avoiding walks, of arbitrary length $|\omega|$, which begin at the origin. The parameter z is known as the *activity*. Physically the activity occurs in the study of a canonical ensemble of polymers of variable length, and in this context is nonnegative. From a mathematical point of view, however, it will sometimes be useful to consider χ to be an analytic function of complex z.

Given two sites x and y, let $c_N(x, y)$ be the number of N-step self-avoiding walks ω with $\omega(0) = x$ and $\omega(N) = y$. The *two-point function* is the generating function for the sequence $c_N(x, y)$, i.e.,

$$G_z(x, y) = \sum_{N=0}^{\infty} c_N(x, y) z^N = \sum_{\omega: x \to y} z^{|\omega|}. \tag{1.3.2}$$

On the right side, the sum over ω is the sum over all self-avoiding walks, of arbitrary length, which begin at x and end at y. This is clearly translation invariant, so $G_z(x, y) = G_z(0, y - x)$. The two-point function is the self-avoiding walk analogue of the simple random walk Green function with killing rate $1 - 2dz$:

$$C_z(x,y) = \sum_{N=0}^{\infty} p_N(x,y)(2dz)^N, \tag{1.3.3}$$

where $p_N(x,y)$ is the probability that an N-step simple random walk beginning at x ends at y.

The generating function for c_N can be written in terms of the two-point function as

$$\chi(z) = \sum_{x \in Z^d} G_z(0,x). \tag{1.3.4}$$

In analogy with spin systems (see Section 2.3) we will refer to the generating function $\chi(z)$ as the *susceptibility*. The power series defining the susceptibility has radius of convergence

$$z_c \equiv \left[\lim_{N\to\infty} c_N^{1/N}\right]^{-1} = \frac{1}{\mu}, \tag{1.3.5}$$

and hence defines an analytic function in the *complex* parameter z if $|z| < z_c$. Since $c_N(0,x) \leq c_N$, the two-point function has radius of convergence at least z_c. It will be shown in Section 3.2 that in fact the radius of convergence is equal to z_c, for all $x \neq 0$. We will refer to z_c as the *critical point*, since it plays a role analogous to the critical point in statistical mechanical systems such as the Ising model or percolation.

It follows from (1.2.10) that

$$\chi(z) \geq \sum_{N=0}^{\infty} (\mu z)^N = \frac{1}{1-\mu z} \tag{1.3.6}$$

and hence χ is "continuous" at the critical point, in the sense that $\chi(z) \to \infty$ as $z \nearrow z_c$. The manner of divergence of $\chi(z)$ at the critical point is related to the behaviour of the coefficients c_N for large N. To see this, we proceed as follows.

First we introduce the notation

$$f(x) \simeq g(x) \quad \text{as } x \to x_0 \tag{1.3.7}$$

to mean that there are positive constants C_1 and C_2 such that

$$C_1 g(x) \leq f(x) \leq C_2 g(x) \tag{1.3.8}$$

uniformly for x near its limiting value. Assuming that there is a γ such that

$$c_N \simeq \mu^N N^{\gamma-1} \quad \text{as } N \to \infty, \tag{1.3.9}$$