National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Studies of Tethered Network Models

by

Edward Levinson

B. Sc. Worcester Polytechnic Institute, 1976

M. Sc. Simon Fraser University, 1984

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the Department

of

Physics

SIMON FRASER UNIVERSITY

October 1992

ISBN 0-315-91095-X

PARTIAL COPYRIGHT LICENSE

Title of Thesis/Project/Extended Essay

Studies of Tethered Network Models

Author:

(signature)

Edward Levinson

(name)

27 October 1992

(date)

APPROVAL

Name: Edward Levinson

Degree: Doctor of Philosophy

Title of thesis: Studies of Tethered Network Models

Examining Committee: Prof. R. F. Frindt
Chair

Prof. Michael Wortis
Senior Supervisor

Prof. Michael Plischke

Prof. K. S. Viswanathan

Prof. Joseph Rudnick
External Examiner
Department of Physics
Univ. of California at Los Angeles

Date Approved:

Prof. David Boal
September 25, 1992

Abstract

This thesis investigates three models that are particular examples of tethered networks. Starting from the viewpoint that a random walk is a collection of linearly connected vertices (internal dimension $D = 1$), the tethered network model is introduced as a natural generalization of the random walk to allow vertices to have a higher degree of internal connectivity ($D > 1$).

The first model studied is a closed two dimensional random walk with a pressure difference p between the inside and outside. Since all terms in the Hamiltonian are quadratic in the position-vector field $\mathbf{r}$, the partition function and its derivatives can be evaluated exactly. The most notable feature of this model is an instability, which occurs at $|p| = p_c$. For $|p| < p_c$, the system has a finite algebraic area and an anisotropic shape; for $|p| \geq p_c$, the algebraic area diverges and the shape is circular. The asphericity is also calculated. A form of bending rigidity, also quadratic in $\mathbf{r}$, is introduced into the model; however, the resulting macroscopic properties are quite different from those one would ordinarily expect. This difference can be traced to the absence of a fixed monomer size in the model.

Studies of the other two models are generally focused on trying to determine how the size of a network embedded in d dimensions scales with the network's maximum linear size L. Three possibilities are described: (1) flat, for which the scaling in two directions is linear in L, and the scaling in the other $d - 2$ directions is with some smaller power of L ; (2) rough, which is similar to flat except that the scaling in the two large directions is not linear

in L; (3) crumpled, for which scaling is uniform in all d directions.

The second model studied is a self-avoiding tethered network with an internal connectivity of a two-dimensional hexagonal lattice, known as a self-avoiding tethered membrane. Monte Carlo simulations of these membranes embedded in $d = 4, 5$ are described. Results for the $d = 4$ simulations suggest that the network is rough. Results of the $d = 5$ simulations are somewhat ambiguous, and are consistent with interpretation either as a rough or crumpled phase.

The third study consists of Monte Carlo simulations of tethered networks possessing the connectivity of a $b = 2$ Sierpiński gasket ($D \approx 1.365$), both with and without self avoidance in a range of embedding dimensions $3 \leq d \leq 14$. The simulation results indicate that $b = 2$ Sierpiński gaskets are crumpled. Measurements of the network size scaling exponents and the upper critical dimension (above which self avoidance is irrelevant) are in good agreement with theoretical predictions.

I dedicate this thesis to my patron saint, Antonio Salieri, *and to my bretheren in the underworld:* The Danaïdes, Tantalus, and Sisyphus.

There is no fate that cannot be surmounted by scorn.

Albert Camus

The Myth of Sisyphus

Acknowledgements

In any undertaking as tedious, long, and dubious as this one, it is hard to know whether or not my collaborators would wish to have their association with this project publicized. I suppose I could ask them, but honesty compels me to name all the guilty parties. Let it be a lesson to them for the future.

Although I am a classicist at heart, I will take the time-honored tradition of thanking the supervisor first and the spouse last, and stand it on its head. My wife, Cathy, has endured her time as a thesis widow with more patience, humor and good grace than anyone could reasonably ask for. I promise never to do this to her again.

The work in this thesis is mine, but the ideas, for the most part, belong to others. David Boal suggested that I study tethered membranes in four and five dimensions (Chapter 4). Hans-Werner Diehl suggested the asphericity calculation for the model described in Chapter 3, and it was Michael Wortis' questions that led to a more general analysis of this model. The idea for the Sierpiński gasket simulations (Chapter 5) was mine, but without Mike Plischke's insistence that the idea was potentially interesting, I would not have pursued it.

I would also like to thank Gary Grest, Yakov Kantor, Damin Liu, Joe Rudnick and Udo Seifert for various discussions on the work presented here.

In retrospect, it seems to me that I have learned most of what I know about physics from a small number of people who were willing to patiently explain things to me in simple terms and who never seemed to tire of my questions. To Jim Glosli, Mike Plischke, Zoltan

Rácz, Madan Rao, Gary Scott, Mark Singer, and Richard Woloshyn, I am grateful for all the lessons.

Finally, I wish to thank my supervisor. Michael Wortis has not been directly involved in any of this research, but he has nonetheless been its most thoughtful and careful critic. His careful questioning and willingness to spend the time required to understand the calculations and ideas have resulted in a much better finished product. It has not been easy to do research that is up to his standards, but it has certainly been educational.

Contents

List of Tables

List of Figures

List of Abbreviations

LSF	Leibler, Singh, and Fisher (Ref. [2])
MC	Monte Carlo
PRW	pressurized random walk with bending rigidity
PRW_0	pressurized random walk without bending rigidity
PSAW	pressurized self-avoiding walk
PSG	phantom Sierpiński gasket
PTM	phantom tethered membrane
PTM_κ	phantom tethered membrane with bending rigidity
PTN	phantom tethered network
QCD	quantum chromodynamics
RBC	red blood cell
RG	renormalization group
RS	random surface
RW	random walk
SA	self avoiding, self avoidance
SAPM	self-avoiding plaquette model
SASG	self-avoiding Sierpiński gasket
SATM	self-avoiding tethered membrane
SATN	self-avoiding tethered network

SAW	self-avoiding walk
TM	tethered membrane
TN	tethered network

Chapter 1

Introduction

1.1 The random walk paradigm

Within science there occasionally arises a paradigm whose applicability spans an especially wide range of disciplines. One such paradigm is the random walk (RW).

A simple definition of a RW is a set of d-dimensional vectors, $\mathbf{r}(s)$ that are random functions of a discrete arc-length $s = n\ell$, where n is a natural number. The vectors share a common origin and respect a connectivity constraint,

$$|\mathbf{r}(s+\ell) - \mathbf{r}(s)| = \ell, \tag{1.1}$$

where ℓ is a constant called the *monomer size*. The nature of the random distribution is determined by the particular model. In this thesis, it is convenient to generalize this definition somewhat and define a RW as an embedding of a discrete one-dimensional space $\mathcal{S}$ ($s \in \mathcal{S}$) in a d-dimensional space $\mathbb{R}^d$ ($\mathbf{r} \in \mathbb{R}^d$), with the following two restrictions on the embedding: 1) it must be random; and 2) it must be order-preserving, i. e., adjacent points in $\mathcal{S}$ must map to nearby points in $\mathbb{R}^d$. The RW described by Eq. (1.1) will subsequently be referred to as a rigid-rod RW in order to distinguish it from the general RW.

Heuristically, the restriction to order-preserving mappings provides the RW with its

linear character. Without it, the system would simply be a gas of monomers. Eq. (1.1) is an example of a mechanism that ensures an order-preserving mapping, but it is not the only one. In fact, we will have occasions in Sec. 1.3 and Chap. 2 modify Eq. (1.1) to

$$|\mathbf{r}(s+\ell) - \mathbf{r}(s)| \leq \ell. \tag{1.2}$$

This condition is sometimes referred to as a tethering constraint and the resulting model will be referred to as a tethered RW. From a physicist's viewpoint, tethered RW's do not differ from rigid-rod RW's in a significant way. The reason is that physicists are usually interested in the continuum limit of these models, and it turns out that they are all equivalent.

RW's have proved to be useful in such diverse fields as probability theory, ecology, molecular biology, polymer chemistry, and several branches of physics, including quantum field theory, statistical mechanics and biophysics. One of the great strengths of the random walk idea is that it has given rise to an entire class of models (hence the term paradigm) that have often had great success in describing scientific phenomena. This is especially true in physics, where the models of interest include: the ordinary random walk, which has been used to describe diffusion processes, spin glasses, and the space-time trajectories of quantum fields; and the self-avoiding walk (SAW)[1], which has been used to describe polymers and the propagation of fermions [3].[2]

Although much of the interest in RW's is due to their usefulness in modelling real systems, it is important to realize that the RW paradigm has developed a life of its own. Many RW models are studied because they are regarded as being intrinsically interesting, although these more theoretical studies sometimes turn out to have important applications for real systems. The best-known example is the study of the SAW embedded in four dimensional space, which has proved crucial to calculating the scaling exponents of a SAW in two or three dimensions [8].

[1] Crudely speaking, a SAW is a RW that does not intersect itself. The notion of self-avoidance will be made more precise in Chap. 2.

[2] This list is far from complete. For a survey of types of RW's and their uses, see Refs. [4, 5, 6, 7].

1.2 The random surface paradigm

In the last fifteen years there has emerged within physics a natural generalization of the RW paradigm, which is thought by some [9, 10] to have the same potential range of applicability and importance as the RW. This new paradigm is the random surface (RS).[3] To date the RS idea has found applicability in biophysics [11], the study of interfaces and crystal growth [12], polymerized membranes [11, 13], and quantum field theory, where it has been used in the context of relativistic string theory and the confining phase of non-Abelian gauge theories such as quantum chromodynamics [14][4].

Because the RS paradigm is a generalization of the RW paradigm, much of RS study is closely related to RW study. For example, the techniques of analysis and the choice of physical observables used to characterize RS's are identical to those for RW's. Also, just as with RW models, it is useful to study particular examples of RS models that do not have obvious experimental implications but serve, rather, to elucidate the general structure of the theory [12]. This last point is even more important for RS's than it is for RW's, since the theoretical tools for analyzing RS's are not as well developed as they are for RW's. Consequently, researchers must oftentimes content themselves with studying models whose main distinction is tractability rather than applicability to real systems. This is also true for RS simulations, which are often more difficult to code and always require more computing time than comparable RW simulations.

While there are many similarities between the RW and RS paradigms, there are, of course, differences. The most important one is that the two-dimensional internal space of a RS model has a non-trivial structure of its own, and this can have a qualitative effect on the RS behavior. Thus, unlike RW's, where the order-preserving restriction on the mapping

[3] A random surface is defined in much the same way as a RW, the difference being that $\mathcal{S} \rightarrow \mathcal{S}^2$ so that s is now a two-dimensional vector $\mathbf{s}$. The requirement that the mapping be order-preserving is the same as for the RW, but its implementation for an RS is generally more complicated. See Sec. 1.3 below and Sec. 2.1 in Chap. 2.

[4] For an overview of the RS idea from a condensed-matter viewpoint see Refs. [13, 15]. For the quantum-field-theory view see Ref. [9].

leads to a unique continuum limit, there is a need to classify RS's on the basis of their $\mathcal{S}^2$ structure.

1.3 A survey of random surfaces

Physical membranes are made up of molecules, bound together into a $D = 2$ structure by the combined effects of intermolecular and membrane-solvent forces. These effects determine the type of internal structure of the membrane, just as they do for flat surfaces. Before taking up the problem of classifying RS internal-space structure, it is helpful to review the three broad classes of flat surfaces:

fluid: In a $D = 2$ fluid, each molecule has on the average six nearest neighbors; however, sites of five-fold and seven-fold coordination, etc., can and do occur. Such cites are called disclinations, and, in the fluid phase, they occur in a random manner. This random distribution of disclinations destroys all translational and orientational order in the fluid phase and the individual monomers (molecules) flow freely. Thus, the identity of the "intrinsic" nearest neighbors (monomers adjacent to a given monomer in $\mathcal{S}^2$) varies in time and from place to place along the surface. The free flow of monomers also means that fluid surfaces have zero local in-plane shear modulus.

When a $D = 2$ fluid is cooled, the effect of the intermolecular forces becomes larger, the energy cost of these free disclinations becomes unsupportable, and the system undergoes a transition to one of the two more-ordered phases with one or more non-zero shear moduli.

hexatic: In the so-called hexatic phase, the disclinations bind into, for example, 5–7 pairs, which are equivalent to dislocations. These dislocations have an energy that scales logarithmically with the system size. Their entropy also scales logarithmically with the system size, so their creation becomes favorable above some finite temperature

T_M, called the melting temperature. The disappearance of free disclinations leads to quasi-long-range (i. e., power-law) orientational order that gives rise to an orientational rigidity. On the other hand, the presence of free dislocations destroys translational order, so there is no conventional shear rigidity.

crystalline: The binding of dislocations into nearby plus-minus pairs (i. e., closely coupled 5-7–5-7 complexes) induces long-range orientational order and a translational order that is quasi-long-ranged because of the low intrinsic dimensionality ($D = 2$). This phase is the $D = 2$ crystalline phase and it has both orientational and translational rigidity.

We now describe how this classification scheme is modified for random surfaces. Fluid and hexatic RS's have the same characteristics as their flat surface counterparts, but crystalline RS's do not [16]. The presence of a dislocation in a crystalline RS will cause the RS to buckle out of plane. This buckling process reduces the internal stresses in the RS to such a degree that the energy cost of a dislocation is now independent of the system size. On the other hand, the dislocation entropy still diverges logarithmically with system size. Thus, any sufficiently large crystalline RS will contain free dislocations, which will destroy the $D = 2$ crystalline quasi-long-range order, and leave the system in the hexatic phase. Strictly speaking, then, crystalline RS's do not exist.

The foregoing argument is certainly relevant when discussing, for example, lipid bilayers below T_M, where the lipid molecules are held together by weak van der Waals forces and the formation of dislocations requires very little energy. In contrast, graphite oxide sheets studied in recent experiments [17, 18] are sufficiently small ($\sim 1\mu$m) and the covalent bonds sufficiently strong that it is unlikely that any dislocations are formed. Since this system lacks both free disclinations and free dislocations, it is reasonable that many of its properties can be modelled by a random surface with a uniform fixed connectivity. The connectivity constraint may be modelled by Eqs. (1.1) or (1.2). In more physical terms, we can think

of fixed-connectivity RS's as having a strong attractive potential between intrinsic nearest neighbors. The monomers are allowed to move relative to one another a short distance in the embedding space, but the "bonds" between them cannot be broken. This generates a non-zero shear modulus.

The RS versions of constraint Eqs. (1.1) and (1.2) give rise to what are known as "crystalline"[5] and tethered RS's, respectively. Although these have some obvious differences, the common features of finite shear modulus and fixed intrinsic nearest neighbors found in both models are thought to be sufficient to ensure that they belong to the same universality class [19]. "Crystalline" RS's are better models of real systems (e. g. graphite oxide), but tethered RS's are easier to work with for various technical reasons (to be discussed throughout the thesis). In this thesis, all of the fixed-connectivity RS's will be of the tethered type.

As is usually the case, the nomenclature in the literature is not as consistent as one might like. The term RS is used almost exclusively in QCD and string theory discussions. In the condensed matter literature, the term RS is replaced by the word membrane; hence, tethered membrane (TM), fluid membrane, etc. The condensed-matter terminology will be used for the remainder of this thesis.

1.4 Overview of the thesis

The similarities between tethered RW's and TM's naturally lead to a unified description of the two, known as the tethered network (TN) [19].[6] A TN is a random mapping of $\mathcal{S}^D$ ($D = 1$ for a tethered RW; $D = 2$ for a TM) in $\mathbb{R}^d$ that respects an appropriately generalized version of Eq. (1.2). The TN idea can be further generalized by allowing D to take on any

[5]The term "crystalline" does not have the same meaning here as it does on page 5. It is used here only in the sense that the spacing between monomers is fixed. It does not imply that the monomers are allowed to diffuse through the lattice or that the monomers create lattice vibrations as they would in a real crystal.

[6]Of course, a similar unified description could be made for rigid-rod RW's and crystalline membranes, but we will not need it here.

real value. This may not seem like a physically interesting generalization, but it has been useful in the analysis of continuum TN models [20] and will be of great importance in this thesis. Chapter 2 contains a detailed description of the geometry, structure, statistical mechanics, and relevant physical observables of TN's. Models obtained from the continuum limit of a TN are also discussed in Chap. 2.

Each of Chaps. 3, 4, and 5 describes calculations and simulations done for a particular example of a TN. These are:

1. 2d closed pressurized RW. For some time now there has been considerable effort invested in trying to understand the shape conformations and fluctuations of closed pressurized fluid membranes, in general, and of human red blood cells in particular [21, 22]. This has turned out to be a formidable problem in terms of both analysis and simulation. These difficulties have inspired some [2, 23, 24, 25, 26, 27, 28, 29] to investigate a reduced-dimension model consisting of a closed pressurized fluid chain embedded in $d = 2$. This model possesses two important simplifications over the fluid membrane model. The first is that a fluid chain is identical to a tethered RW (for equilibrium properties) [2], so the difficult problems in obtaining thermal averages for fluid membranes disappear. The other simplification is that embedding the tethered RW in $d = 2$ greatly simplifies the geometrical analysis needed. In Chap. 3 I analyze the relationships among a set of continuum models of closed pressurized RW's, and I calculate some shape ratios (see Sec. 2.2.4 of Chap. 2 for an explanation of this term) for a particularly simple (Gaussian) model. The details of the shape calculations are relegated to the Appendix.

2. Self-avoiding tethered membrane. The study of self-avoiding TM's (SATM) has been

filled with controversy [11, 30], mostly regarding the existence of a phase transition from a low-temperature "flat" phase, in which a SATM occupies only a two-dimensional slice of $\mathbb{R}^d$, to a high-temperature "crumpled" phase, in which its embedding in $\mathbb{R}^d$ is more uniform (These terms will be made precise in Chap. 2). Most of this controversy has centered on the existence of a crumpled phase in $d = 3$. The first simulations [31, 32] indicated that such a phase did exist, but later work [33, 34] came to the opposite conclusion. Chapter 4 describes Monte Carlo (MC) simulations done on SATM's embedded in $d = 4$ and $d = 5$ that look for a crumpled phase. The increased size of the phase space should make the crumpled phase easier to discover, if it exists. No crumpled phase is found, but indications of a possible third "rough" phase that is intermediate to the other two phases are discovered. There is also discussion of controversies regarding recent experiments [17, 18], the existence of the rough phase [1], and the appropriateness of using the SATM model to describe experimentally realizable systems [35].

3. Self-avoiding Sierpiński gasket. The controversy and lack of understanding of SATM's is in stark contrast to the great depth of knowledge we have about SAW's. It is desirable, then, to understand how the change from $D = 1$ to $D = 2$ leads to so many problems. One possible way to explore this issue is to study self-avoiding TN's (SATN) for which $1 < D < 2$. The generalization of D to real values is more than a theoretical nicety. Surprisingly, SATN's with non-integer D can be realized (at least on a computer) in the form of a self-avoiding Sierpiński gasket (SASG). Chapter 5 describes MC simulations of SASG's embedded in $\mathbb{R}^d$ with $3 \leq d \leq 14$ and compares the results for the scaling behavior of the radius of gyration with those obtained from mean-field and field-theory calculations. All three methods are found to be reasonably consistent. Shape parameters are also determined from the MC data.

Figure 1.1 illustrates the connectivity of the three models that will be investigated in

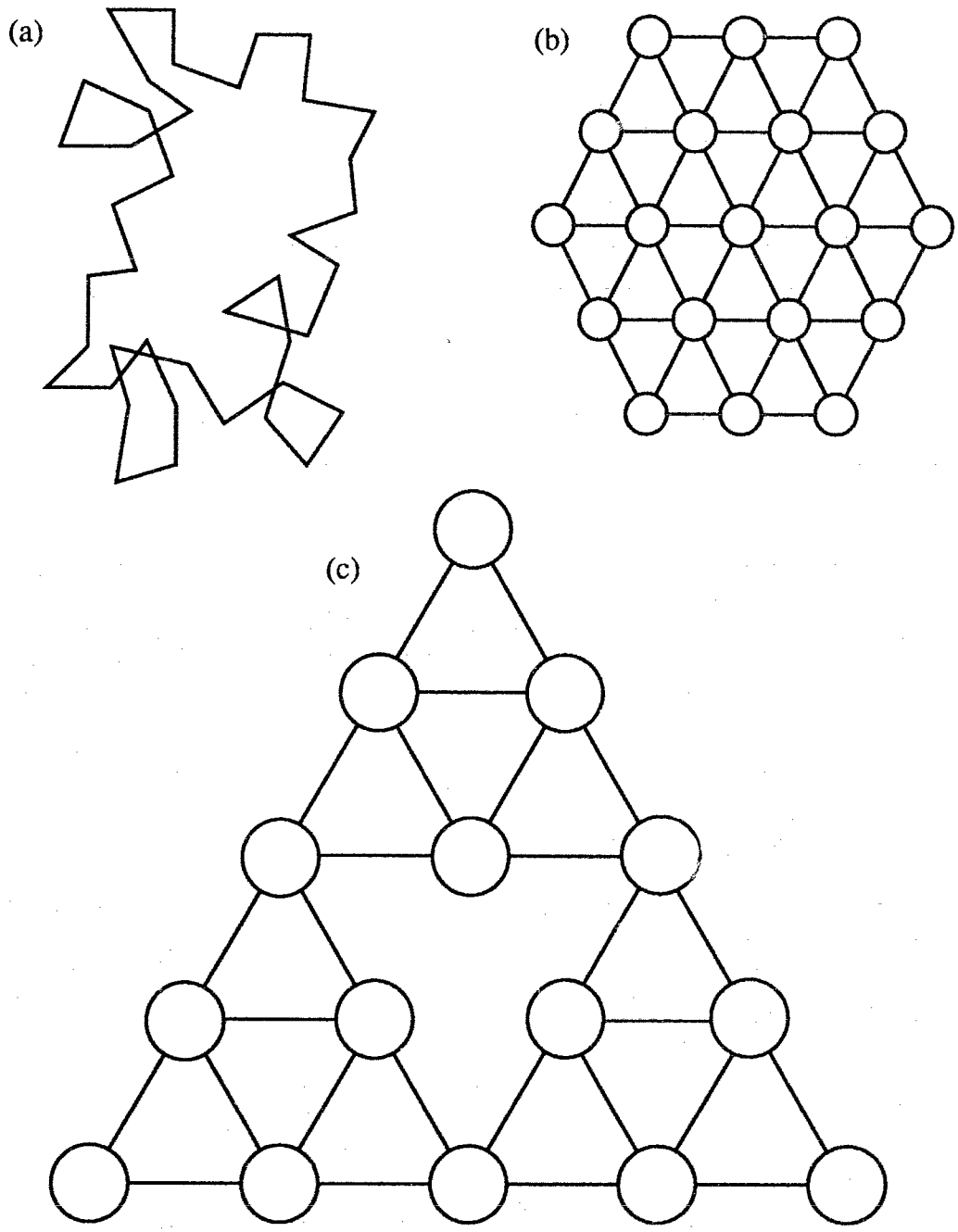


Figure 1.1: The TN models studied in this thesis. The lines represent the tethers and the open circles represent d-dimensional hard spheres which are used to enforce a self-avoidance constraint (see Chap. 2). (a) A closed RW without self-avoidance. The analysis in Chap. 3 deals mainly with the continuum limit of this model. (b) A self-avoiding tethered membrane. (c) A self-avoiding Sierpiński gasket.

these chapters.

The last chapter, Chap. 6, summarizes the conclusions obtained from this thesis.

If the reader has become alarmed at the proliferation of abbreviations in this chapter, s/he may find some comfort on page xvi, which contains a list of abbreviations used in this thesis.

Chapter 2

The Tethered Network Model

2.1 Structure of tethered networks

2.1.1 Geometry

A nice feature of TN's is that they can be described in purely geometrical terms: A TN consists of a collection of N vertices (monomers), each connected to some set of "intrinsic nearest neighbors" by "tethers" (strings) of length ℓ, usually in a regular manner. The tethers are a geometrical means of enforcing the constraint Eq. (1.2) (or its appropriate generalization). In the absence of other constraints and energetic considerations, each vertex is free to move anywhere in the embedding space provided that none of its tethers is stretched beyond its maximum length. The precise nature of the connections determines the topology of the network; the measure of this topology is the network's intrinsic or topological dimension D. For example, the tethered RW ($D = 1$) consists of vertices connected in a linear way; for a TM ($D = 2$), the tethers of one vertex are connected to all its intrinsic nearest neighbors so as to form a homogeneous two-dimensional lattice. The generalization to larger integer values of D is straightforward.[1] The resulting network has a characteristic linear size $L \sim N^{1/D}$ (measured in units of ℓ). The vertices can be labelled either by a discrete

[1] This concept will be extended to include any real-valued D in section 2.4.2.

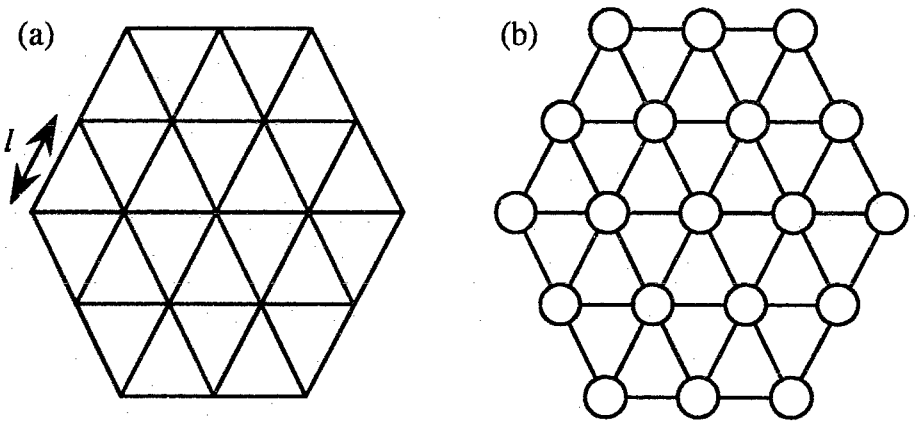


Figure 2.1: Two versions of the $D = 2$ TN model: (a) The PTN. The lines represent the tethers, which cannot be stretched beyond some maximum distance, ℓ. (b) The SATN. The open circles represent hard d-dimensional spheres. In addition to the tethering constraint, the hard spheres are not allowed to overlap.

index α ($1 \leq \alpha \leq N$), or by a D-dimensional lattice vector $\mathbf{s} = (s_1, \ldots, s_D)$, $0 \leq s_i \leq \mathcal{L}_0$ ($\mathcal{L}_0 \equiv L\ell$).

The network is then embedded in a d-dimensional Euclidean space $\mathbb{R}^d$. Operationally, this means that every vertex is assigned a d-dimensional vector $\mathbf{r}(\mathbf{s})$ or $\mathbf{r}^{(\alpha)}$ that respects the constraints of the particular TN model.

As illustrated in Fig. 2.1, there are two broad classifications of TN models:

phantom TN (PTN): For a PTN, all vertices are free to move through the embedding space subject only to the constraint that *connected* intrinsic nearest-neighbor vertices be within a distance ℓ of each other, i. e.,

$$\left|\mathbf{r}^{(\alpha)} - \mathbf{r}^{(\beta)}\right| \leq \ell; \tag{2.1}$$

α and β label connected nearest neighbors.

self-avoiding TN (SATN): The embedding of a TN in $\mathbb{R}^d$ defines a D-dimensional volume, which is determined by triangulating the nearest-neighbor vertices. Embeddings

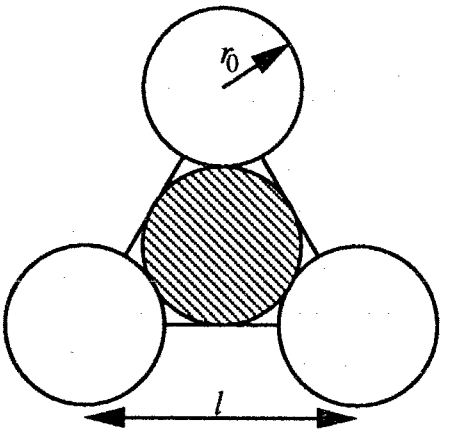


Figure 2.2: How r_0 is determined for a triangularly connected TN. The three open circles each have a radius r_0 and are tethered to one another. The shaded circle, also of radius r_0, is not tethered to the others. Simple geometry shows that the circles just touch when $r_0 = \ell/2\sqrt{3}$. Note that r_0 is a function of ℓ.

of SATN's in $\mathbb{R}^d$ that result in self-intersection of this volume are not allowed. This constraint can be enforced by regarding every vertex as a hard d-dimensional sphere of radius r_0, and requiring

$$\left|\mathbf{r}^{(\alpha)} - \mathbf{r}^{(\beta)}\right| \geq 2r_0; \tag{2.2}$$

α and β label *any* two vertices. The size of r_0 is chosen sufficiently large so that a sphere cannot pass through the open spaces in the TN, even when the tethers are fully stretched. Figure 2.2 illustrates this idea. Note that this constraint is non-local in the sense that vertices separated by a large distance in the internal space will be restricted by it.

SATN's are usually far more realistic models of polymerized networks than PTN's, but the non-local nature of the self-avoidance condition makes them much more difficult to study.

2.1.2 Energetics

Although, highly non-trivial, the TN model described above is not yet very realistic. A more realistic model would include a Hamiltonian, $H_d^{(L)}$, accounting for the "deformation energy" of the TN, e. g., the energy cost of bending or stretching it [36]. While it is certainly desirable to include these effects, they are mostly ignored in this thesis. The reason is that the study of $D > 1$ models containing only constraint terms has proved to be sufficiently complicated and controversial, so that there seems little advantage to studying more complicated models now. Energetics are included in the $D = 1$ model studied in Chap. 3, where the Hamiltonian contains a (pressure)×(area) energy term and a bending-energy term.

For simulation purposes, Eqs. (2.1) and (2.2) are the most convenient ways of expressing the tethering and SA constraints, respectively, but they are not very useful for analytic work. Instead, the constraints are better expressed as terms in a Hamiltonian. The first task is to characterize mathematically the effects of the tethers. A potential that reproduces the effects of the constraint Eq. (2.1) exactly is[2]

$$\beta H_t^{(L)} = \sum_{\langle \mathbf{s},\mathbf{s}' \rangle} u_t(s,s'), \tag{2.3}$$

where

$$u_t(s,s') = \begin{cases} 0 & \text{if } |\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')| < \ell \\ \infty & \text{otherwise} \end{cases}, \tag{2.4}$$

and the notation $\sum_{\langle \mathbf{s},\mathbf{s}' \rangle}$ indicates a sum over all pairs of tethered vertices. This is still not a very convenient analytical form. We therefore look for another, more computationally useful, potential that contains the same essential physics.

The standard approach [19] is to drop the "hard" constraint Eq. (2.3) and to model each tether as a rigid-rod RW of n steps, each of length η. The probability distribution for the

[2]In anticipation of our ultimate interest in using statistical mechanics to study TN's, the definitions for Hamiltonian terms will include a factor of $\beta = 1/k_B T$.

end-to-end distance, x, of a rigid-rod RW is [4]

$$P_n(x) \sim \exp\left(\frac{-dx^2}{2n\eta^2}\right), \tag{2.5}$$

provided that n is large enough. In the language of statistical mechanics, $dx^2/n\eta^2$ is an effective potential for the model. If the end-to-end distance for the rigid-rod RW is now identified with the distance between a pair of tethered vertices and $\ell = \sqrt{n}\eta/\sqrt{d}$, then the effective potential becomes

$$\beta H_0^{(L)} = \sum_{\langle \mathbf{s},\mathbf{s}' \rangle} \frac{1}{2\ell^2} |\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')|^2. \tag{2.6}$$

The end result of this procedure is to replace the tethers with springs of zero equilibrium length. Because P_n represents a Gaussian distribution, the model defined by $H_0^{(L)}$ is often referred to as the Gaussian model.

$H_t^{(L)}$ and $H_0^{(L)}$ define different models, not different representations of the same model, but the differences between them are not as large as one might think. Monte Carlo simulations [37, 38] and numerical Migdal-Kadanoff renormalization group studies [32] indicate that $H_t^{(L)}$ and $H_0^{(L)}$ belong to the same "universality class" under most circumstances.[3], i. e., they have the same generic long-distance behavior. It turns out for technical reasons that $H_0^{(L)}$ is much more convenient to work with than $H_t^{(L)}$.

The SA constraint can similarly be represented as a potential

$$\beta H_{\mathrm{SA}}^{(L)} = \sum_{\{\mathbf{s},\mathbf{s}'\}} u_{\mathrm{SA}}(s, s'), \tag{2.7}$$

where

$$u_{\mathrm{SA}}(s, s') = \begin{cases} \infty & \text{if } |\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')| \leq 2r_0 \\ 0 & \text{otherwise} \end{cases}, \tag{2.8}$$

[3] An exception occurs for configurations where the typical $|\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')| \gg \ell$. Such highly stretched configurations are forbidden by $H_t^{(L)}$, but are allowed by $H_0^{(L)}$, although usually they are very improbable. Thus, under normal circumstances, this is an unimportant difference between them. However, the application of external forces to the system can make these configurations typical, and it is here that there will be important differences between the two potentials. Such a situation occurs in Chap. 3.

and the notation $\sum_{\{s,s'\}}$ indicates a sum over all pairs of vertices. Unfortunately, an approximation procedure similar to the one used to go from Eq. (2.3) to Eq. (2.6) is not known for the SA constraint. The reason is that it seems to be very difficult to find such a procedure that maintains the non-local character of the interaction. Although no one has succeeded in finding a useful approximation to Eq. (2.7) [39], a truly SA potential is known for continuum models (see Eq. (2.45) below).

2.2 Physical observables

We are interested in describing the characteristic size and shape of a TN as a whole. To do this requires a two-step procedure: The first step, which does not involve any statistical mechanics, is to construct macroscopic physical observables from the microscopic degrees of freedom (the **r**'s). The second step is to perform the usual statistical mechanics average of the macroscopic variables over the ensemble of configurations.

2.2.1 Invariants of the inertia tensor

In searching for a useful measure of the macroscopic properties of an individual TN, we are guided by considerations of simplicity and symmetry. Simple observables are the ones that are easiest to interpret physically. Symmetry considerations arise from the observation that, provided that the ensemble is isotropic, its averages implicitly include an average over the angular degrees of freedom for the entire TN; only rotational invariants will survive this averaging process. The moment of inertia tensor,[4]

$$Q_{ij}^{(L)} = \frac{1}{2}\sum_{\alpha,\beta=1}^{N}\left[r^{(\alpha)}-r^{(\beta)}\right]_i\left[r^{(\alpha)}-r^{(\beta)}\right]_j$$

[4]This definition has been chosen for computational convenience, but it is slightly different from the one found in mechanics texts. Assuming that the center of mass is located at the origin, the two are related by $Q_{ij}^{(L)} = I\delta_{ij} - I_{ij}$, where I_{ij} is the usual moment of inertia tensor [40], and I is the usual moment of inertia. Note that since Eq. (2.9) implies that all the vertices have unit mass, the engineering dimension of $Q_{ij}^{(L)}$ is length^2 instead of the usual mass $\times$ length^2.

$$= \frac{1}{2}\sum_{\mathbf{s},\mathbf{s}'} \left[r(\mathbf{s}) - r(\mathbf{s}')\right]_i \left[r(\mathbf{s}) - r(\mathbf{s}')\right]_j \tag{2.9}$$

$(i, j = 1, \ldots, d)$, is a simple macroscopic quantity that contains all the information needed to construct a simple class rotational invariants. $Q^{(L)}$ is real symmetric; thus a rotation can always be found to diagonalize $Q^{(L)}$ so that $Q_{ij}^{(L)} \xrightarrow{\text{rotation}} \tilde{Q}_{ij}^{(L)} = \lambda_i^{(L)}\delta_{ij}$. Therefore, only d of the d^2 elements of $Q^{(L)}$ are independent.

The $\lambda_i^{(L)}$'s are referred to as the principal moments of inertia and are just the eigenvalues of $Q^{(L)}$. Although they are independent of rotations, their ordering along the diagonal of $\tilde{Q}^{(L)}$ is not unique, since rotations can always be found whose only effect is to scramble their order. This non-uniqueness in labelling can be removed by imposing the convention

$$\lambda_i^{(L)} \leq \lambda_j^{(L)}, \text{ when } i < j, \tag{2.10}$$

which will be used in all future discussions. In simulation studies, where one has access to the individual configurations, imposing this convention is a trivial numerical problem, and, as a result, the calculation of the ensemble averages of the $\lambda_i^{(L)}$'s is straightforward. On the other hand, it is impossible in practice to diagonalize $Q_{ij}^{(L)}$ analytically. Analytic studies therefore require rotationally invariant macroscopic quantities obtained directly from $Q_{ij}^{(L)}$, whose rotational invariance does not depend on convention (2.10).

Combinations of eigenvalues that are symmetrized with respect to the i-index are just such quantities. A general form for these invariants is

$$M_j^{(L)} = \sum_{\substack{i_j = j \\ i_1 < i_2 < \cdots < i_j}}^{d} \prod_{\alpha=1}^{j} \lambda_{i_\alpha}^{(L)}, \tag{2.11}$$

where $j = 1, \ldots, d$. Three well-known special cases are:

1. the radius of gyration $(j = 1)$:

$$M_1^{(L)} = \sum_{i=1}^{d} \lambda_i^{(L)} = \mathrm{Tr}\, Q^{(L)} \equiv \left(R_g^{(L)}\right)^2; \tag{2.12}$$

2. the determinant ($j = d$):

$$M_d^{(L)} = \prod_{i=1}^{d} \lambda_i^{(L)} = \det Q^{(L)}; \tag{2.13}$$

3. the sum of the minors ($j = 2$):

$$\begin{aligned} M_2^{(L)} &= \sum_{i<j} \lambda_i^{(L)} \lambda_j^{(L)} \\ &= \frac{d-1}{2d} \left(\mathrm{Tr}\, Q^{(L)} \right)^2 - \frac{1}{2} \mathrm{Tr} \left(\hat{Q}^{(L)2} \right), \end{aligned} \tag{2.14}$$

where

$$\hat{Q}_{ij}^{(L)} = Q_{ij}^{(L)} - \frac{1}{d} \mathrm{Tr}\, Q^{(L)} \delta_{ij}. \tag{2.15}$$

2.2.2 Statistical mechanics

The preceding discussion has been entirely in terms of a single configuration; but, if we are to do statistical mechanics, an average over the ensemble of configurations must also be performed.

The partition function, $Z^{(L)}$, for a TN model is just a sum over all configurations with the usual $e^{-\beta E}$ thermal weight function. Constraint conditions that suppress certain "impossible" configurations may be enforced directly or via terms in the Hamiltonian that assign infinite energy to the unwanted configurations, as was done in Sec. 2.1.2. Thus, for PTN's, which are subject to constraint Eq. (2.1),

$$Z^{(L)} = \int D[\mathbf{r}] \prod_{\langle \mathbf{s},\mathbf{s}' \rangle} \theta(\ell - |\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')|)\, e^{-\beta H_d^{(L)}}, \tag{2.16}$$

where

$$D[\mathbf{r}] \equiv \prod_{\mathbf{s}} d\mathbf{r}(\mathbf{s}), \tag{2.17}$$

and the Hamiltonian $H_d^{(L)}$ contains only the deformation energy. Alternatively, we could write

$$Z^{(L)} = \int D[\mathbf{r}]\, e^{-\beta \left(H_d^{(L)} + H_t^{(L)} \right)}, \tag{2.18}$$

where the explicit constraint has been replaced by tethering Hamiltonian, H_t from Eq. (2.3). Similarly, for SA models, the SA constraint can be included in $Z^{(L)}$ directly by including the extra factor $\prod_{\{\mathbf{s},\mathbf{s}'\}}\theta(|\mathbf{r}(\mathbf{s})-\mathbf{r}(\mathbf{s}')|-2r_0)$, or indirectly by including the Hamiltonian $H_{\rm SA}^{(L)}$ from Eq. (2.7) in the exponent. In practice, including the constraints (or some approximation to the constraints) in the Hamiltonian is always easier than including them in the integration measure. Thus, we will have occasion (especially in Sec. 2.4) to refer to the tethering term or the SA term in a Hamiltonian.

Another useful quantity is the pair correlation function,

$$G_{ij}^{(L)}(\mathbf{s},\mathbf{s}') = \langle r_i(\mathbf{s}) r_j(\mathbf{s}')\rangle = \frac{1}{Z^{(L)}}\int D[\mathbf{r}]\, r_i(\mathbf{s}) r_j(\mathbf{s}') e^{-\beta H^{(L)}}, \tag{2.19}$$

where $\langle\cdots\rangle$ denotes an ensemble average. Note that in this equation all relevant constraints have been absorbed into $H^{(L)}$. Eq. (2.19) can be used to construct physical observables. For example, combining Eqs. (2.9) and (2.19) gives

$$\left\langle Q_{ij}^{(L)}\right\rangle = \frac{1}{2}\sum_{\mathbf{s},\mathbf{s}'}\left[G_{ij}^{(L)}(\mathbf{s},\mathbf{s}) + G_{ij}^{(L)}(\mathbf{s}',\mathbf{s}') - 2G_{ij}^{(L)}(\mathbf{s},\mathbf{s}')\right]. \tag{2.20}$$

Such formulas are useful for the shape calculations described in Chap. 3.

2.2.3 Scaling laws

Scaling laws are one of the most useful ways of characterizing the long-distance behavior of a TN. Properly speaking, scaling exponents only exist in the thermodynamic limit ($L\to\infty$ with ℓ fixed); but, L-dependent exponents can be introduced in order to make contact with simulations, which are necessarily done on finite-sized systems.

The best-known result from the RW literature is that the average end-to-end distance of a RW is proportional to $\sqrt{L}$. It is easy to show [41] that this same scaling relationship holds for $\left\langle R_g^{(L)\,2}\right\rangle$. By analogy with RW's, the scaling exponent, $\nu^{(L)}$, for a TN is defined through the formula,

$$\left\langle R_g^{(L)\,2}\right\rangle \sim L^{2\nu^{(L)}}. \tag{2.21}$$

Similarly defined is $\nu_i^{(L)}$ through

$$\left\langle \lambda_i^{(L)} \right\rangle \sim L^{2\nu_i^{(L)}}. \tag{2.22}$$

The true scaling exponents ν and ν_i are obtained by taking the thermodynamic limit of the L-dependent ones. The thermodynamic limit and Eq. (2.10) imply

$$\nu_i \leq \nu_j \text{ when } i < j. \tag{2.23}$$

Eqs. (2.12), (2.21), and (2.22) further imply that

$$\nu = \nu_d. \tag{2.24}$$

The range of allowable values for ν is restricted by simple geometric arguments. Eqs. (2.21) and (2.12) give an upper bound $\nu \leq 1$, and it is obvious that $\nu \geq 0$. The lower bound on ν for a SATN can be further restricted by considering that a SATN of intrinsic volume L^D, when "close packed" (i. e., isotropically embedded in $\mathbb{R}^d$ in a way that minimizes its volume in $\mathbb{R}^d$), will have a typical linear size $L^{D/d}$. These considerations combine to give

$$\begin{aligned} &0 \leq \nu \leq 1 \quad \text{for a PTN;} \\ &\frac{D}{d} \leq \nu \leq 1 \quad \text{for a SATN.} \end{aligned} \tag{2.25}$$

There are other scaling exponents such as γ, defined, once again in analogy with RW's, through the equation

$$Z^{(L)} \sim L^{\gamma - 1}. \tag{2.26}$$

The calculation of γ for tethered membranes ($D = 2$ TN's) has become a controversial subject in the literature [20] because its value for a TM appears to depend upon the shape of the TM boundary, and is therefore not universal. This subject will not be discussed further here.

A basic question in the study of TN's is how the $\lambda_i^{(L)}$'s scale relative to one another as a function of L. There are different types of scaling behavior, which are thought to represent

different thermodynamic phases. Much of the subsequent work in this thesis is devoted to determining the phases of specific TN models. There are three known phases[5]:

flat, for which $\nu = \nu_d = \nu_{d-D+1} = 1$ and $\nu_i < 1$ for $i = 1, \ldots, d-D$. As $L \to \infty$ the TN becomes increasingly concentrated in a D-dimensional slice of $\mathbb{R}^d$. The term flat comes from the $D = 2$ case, which is the one most commonly discussed.

crumpled, for which $\nu = \nu_i < 1$ for all values of i. In this phase all the $\lambda_i^{(L)}$'s have the same scaling behavior, so the shape of the TN does not change as $L \to \infty$.

rough, for which $\nu = \nu_d = \nu_{d-D+1} < 1$ and $\nu_i < \nu$ for $i = 1, \ldots, d-D$. The $L \to \infty$ behavior is similar to that for the flat phase, but, for a given L, a flat TN will always be more concentrated in a D-dimensional slice of $\mathbb{R}^d$ than a rough TN.

The above list of phases is the most general one possible that is consistent with the following two assumptions about the thermodynamic limit:

1. $\nu = \nu_k$ for all $d - D + 1 \leq k \leq d$

2. $\nu_i = \nu_j$ for all $i, j \leq d - D$

How valid are these assumptions? It is tempting to argue that the intrinsic D-dimensional structure of a TN will either leave the SO(d) symmetry of $\mathbb{R}^d$ intact and result in a crumpled phase, or, at most, break the SO(d) symmetry into SO($d-D$) $\times$ SO(D) and result in either a flat or rough phase. This is, in fact, an argument about the symmetry of the $Q_{ij}^{(L)}$ *eigenvalues*, i. e., that $\lambda_d = \cdots = \lambda_{d-D+1}$ and $\lambda_i = \lambda_j$ for all $i, j \leq d - D$. It is well-known [43, 44, 45] that the eigenvalues of TN's do not have this symmetry (they are all different). From an "experimental" point of view, there is currently no evidence from simulations to indicate that the first assumption is false; however Chap. 4 describes simulations that suggest that the second assumption may be false (see Sec. 4.4.3).

[5]The following descriptions of the flat and rough phases only make sense for integer D. It is not clear how to properly generalize these descriptions when D is not an integer. The usual practice [42, 43] is to replace D with the next largest integer.

2.2.4 Shapes

Analyses of critical phenomena are often focused on critical exponents such as ν and γ. Much of the interest in these exponents lies in their universality—several models that differ in their detailed characterization may, nonetheless, have the same critical exponents. In addition to being described by universal critical exponents, TN's also seem to possess universal *amplitude ratios* [46, 47], which characterize the equilibrium shape of the surface in terms of the $\lambda_i^{(L)}$'s. The two most commonly used ratios are the *shape parameter*,

$$S_d^{(L)} = \left\langle \frac{\lambda_1^{(L)}}{\lambda_d^{(L)}} \right\rangle, \tag{2.27}$$

and the *asphericity*,

$$\begin{aligned} A_d{}^{(L)} &= \left\langle \frac{d\mathrm{Tr}\left(\hat{Q}^{(L)2}\right)}{\left(\mathrm{Tr}\, Q^{(L)}\right)^2} \right\rangle \\ &= \left\langle \frac{\sum_{i,j=1}^{d} \left(\lambda_i^{(L)} - \lambda_j^{(L)}\right)^2}{2\,(d-1)\left(\sum_{i=1}^{d} \lambda_i^{(L)}\right)^2} \right\rangle. \end{aligned} \tag{2.28}$$

Note that $A_d^{(L)}$ is composed entirely of $M_j^{(L)}$'s, but $S_d^{(L)}$ is not. Thus, while both $A_d^{(L)}$ and $S_d^{(L)}$ can be measured from simulation data, only $A_d^{(L)}$ can be calculated analytically in practice. Like the scaling exponents $\nu_i^{(L)}$, $A_d^{(L)}$ and $S_d^{(L)}$ possess well-defined thermodynamic limits, denoted by A_d and S_d, respectively.

These two shape measures are sensitive to different characteristics of a TN. To see how they differ, consider three different scenarios:

1. $\nu_i < \nu_j$ for $i < j \leq d$. From Eq. (2.27) it follows that $S_d \sim L^{\nu_1 - \nu_d} \to 0$. The behavior of A_d is more complicated. If $\nu_{d-1} < \nu_d$, so that there is a unique largest ν_j $(j = d)$, the TN will have a "cigar shape" and inspection of Eq. (2.28) shows that $A_d \to 1$. If $j < d$ and $\nu_j = \nu_{j+1} = \ldots = \nu_d$, the TN will be effectively embedded in a $(d - j + 1)$-dimensional space, and A_d will measure the shape in this reduced space.

2. All ν_i's equal, but the $\lambda_i^{(L)}$'s unequal. In this case, both A_d and S_d will be non-zero. Because A_d contains information about all the $\lambda_i^{(L)}$'s, it provides a more detailed characterization of the shape than S_d, which contains information about only two of them.

3. All $\lambda_i^{(L)}$'s equal. Inspection of Eqs. (2.27) and (2.28) shows that $S_d = 1$ and $A_d = 0$.

Some workers prefer to use variants of these amplitude ratios, namely,

$$s_d^{(L)} = \frac{\left\langle \lambda_1^{(L)} \right\rangle}{\left\langle \lambda_d^{(L)} \right\rangle}, \tag{2.29}$$

and

$$a_d^{(L)} = \frac{\left\langle \sum_{i,j=1}^{d} \left(\lambda_i^{(L)} - \lambda_j^{(L)} \right)^2 \right\rangle}{2\,(d-1) \left\langle \left(\sum_{i=1}^{d} \lambda_i^{(L)} \right)^2 \right\rangle}. \tag{2.30}$$

These amplitude ratios also have well-defined thermodynamic limits, denoted by a_d and s_d, respectively. Comparing Eqs. (2.27) and (2.28) with Eqs. (2.29) and (2.30), respectively, one can see that the only difference between them is that the former two are ensemble averages of ratios for individual configurations, whereas the latter two are ratios of ensemble averages. The distinction between S_d and s_d is not important here. (The results for S_d in the above-mentioned scenarios hold equally well for s_d.) The only point of interest is whether or not they are non-zero (see section 2.2.5 below). On the other hand, as a measure of the shape of a typical configuration, A_d is preferable to a_d precisely because it is an ensemble average of a quantity calculated for each configuration.

2.2.5 Criteria for crumpling

Two central issues in the study of TN's are the existence of a crumpled phase and of a flat-crumpled phase transition. These questions have been especially important and controversial in the study of TM's [30].

phase	ν	S_d
flat	1	0
rough	< 1	0
crumpled	< 1	> 0

Table 2.1: Criteria for distinguishing between the three known phases of a TN.

If there were only the flat and crumpled phases (as many workers have assumed), then knowing ν would suffice to determine which one was observed. However, since $\nu < 1$ cannot distinguish between the rough and crumpled phases, some other additional test is needed. In principle, all the ν_i's could be used, but the same difficulty arises here as arose for the analytic calculation of the eigenvalues in Sec. 2.2.1. The ν_i's are also difficult to measure accurately from simulation data. The shape parameter S_d (or s_d), however, does provide a useful means of distinguishing between the flat and rough phases ($S_d = 0$) on the one hand, and the crumpled phase ($S_d \neq 0$) on the other. Thus, combined measurements of both ν and S_d suffice to distinguish between the three known phases of a TN. These conclusions are summarized in Table 2.1.

There has been a great deal of effort in the literature devoted to resolving analytically the question of which phases of TN's are actually realized [11]. The importance of S_d in determining the phase and the difficulty in calculating it make it unlikely that phase existence questions can be resolved analytically. Note that it is sometimes possible to compute the asphericity; however, this is no help since it does not provide any indication of the phase, i. e., all values of the asphericity are possible in each of the phases.

In subsequent chapters I will discuss three things about each TN: the scaling behavior of the $\lambda_i^{(L)}$'s, various ratios of the $\lambda_i^{(L)}$'s that characterize the shape of the TN, and the phase of the TN.

2.3 The continuum limit

The description of TN's given so far has been for discrete ($\ell > 0$) versions of the models. Discrete versions are appropriate and necessary for simulations, but statistical field theory usually requires continuum models for reasons of computational tractability. In passing to the continuum limit, one would like the continuum model to be closely related to the discrete one—derived from it by a straightforward application of the $\ell \to 0$ limit. There are, however, some subtleties involved in taking this limit, which I now discuss. For simplicity the argument is given in terms of a random walk, but it is applicable to any TN when suitably generalized. In addition to being more tractable than discrete models, continuum models are interesting in their own right as examples of random fractals.

Before the continuum limit is taken, there are three length scales that characterize the RW: the step length ℓ, the intrinsic length $\mathcal{L}_0$ and a macroscopic distance such as R_{g0}. It is important to realize that all three quantities have the same engineering dimensions. Thus, if R_{g0} depends on both ℓ and $\mathcal{L}_0$, then these quantities must be related in the following way [48]:

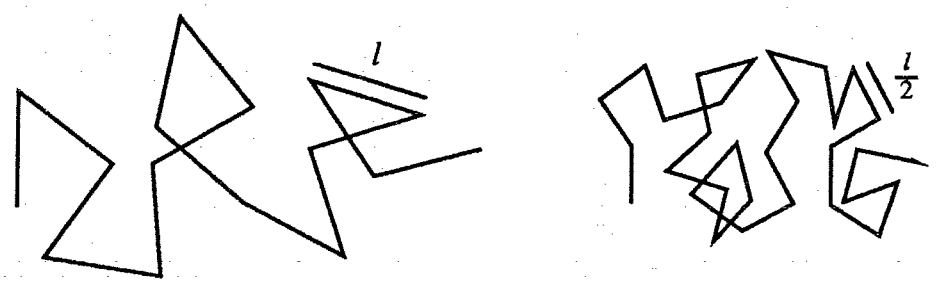
$$R_{g0} \sim \sqrt{\ell \mathcal{L}_0} \sim \ell \sqrt{L}. \tag{2.31}$$

We can begin to understand the subtlety of the continuum limit by asking the following question: What happens to the RW when ℓ decreases, say, $\ell \to \ell' = \ell/2$? The simple answer is that it depends on what happens to $\mathcal{L}_0$ and R_{g0}, i. e., demanding $\ell \to 0$ does not uniquely specify a continuum limit — we must also specify what happens to $\mathcal{L}_0$ and R_{g0} in the process. Although it is possible to consider a wide variety of scenarios, I will discuss only two:

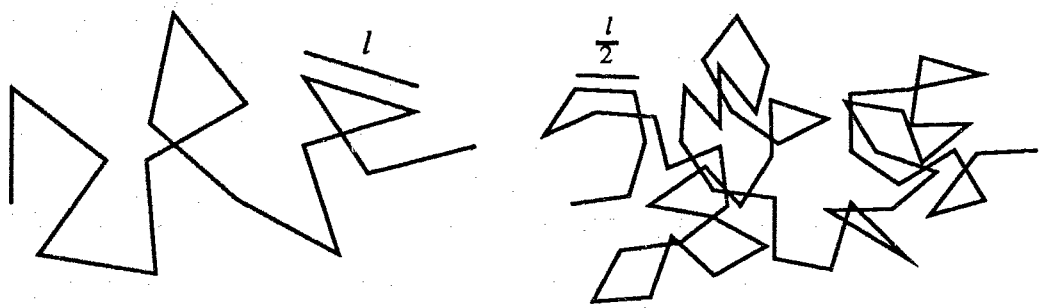
1. $\ell \to \ell/2$ and $\mathcal{L}_0$ fixed. From Eq. (2.31),

$$\left.\begin{array}{ccc} \ell & \to & \ell/2 \\ \mathcal{L}_0 & & \text{fixed} \end{array}\right\} \Longrightarrow \left\{\begin{array}{ccc} L & \to & 2L \\ R_{g0} & \to & R_{g0}/\sqrt{2} \end{array}\right. . \tag{2.32}$$

This limit is illustrated in Fig. 2.3a.



(a)



(b)

Figure 2.3: Two possible limiting procedures for the discrete RW (a) $\ell \to \ell/2$ and $\mathcal{L}_0$ fixed. This limit is not very physical because the physical observable R_{g0} shrinks as a result of applying the limiting procedure. (b) $\ell \to \ell/2$ and R_{g0} fixed. This limit has the desirable property that the physical observable R_{g0} is not altered by the limiting procedure. However, the intrinsic length of the RW does increase.

2. $\ell \to \ell/2$ and R_{g0} fixed. From Eq. (2.31),

$$\left.\begin{array}{rcl} \ell & \to & \ell/2 \\ R_{g0} & & \text{fixed} \end{array}\right\} \Longrightarrow \left\{\begin{array}{rcl} \mathcal{L}_0 & \to & 2\mathcal{L}_0 \\ L & \to & 4L \end{array}\right. . \tag{2.33}$$

This limit is illustrated in Fig. 2.3b.

The physics, of course, should not depend on the limiting procedure, which is just a computational convenience. Clearly, then, the physically interesting limit is the one described in scenario 2, since it fixes the physical observable R_{g0}. As the RW is divided in to smaller and smaller steps, the number of steps (L) and the intrinsic length $\mathcal{L}_0$ must grow in order to keep R_{g0} fixed. In the limit $\ell \to 0$, $\mathcal{L}_0 \to \infty (\sim 1/\ell)$ and $L \to \infty (\sim 1/\ell^2)$. This is the central point: the limiting procedure leaves neither ℓ nor $\mathcal{L}_0$ finite; only the product

$$\begin{aligned} \mathcal{L} &\equiv \ell^{1/\nu} L \\ &= \ell \mathcal{L}_0 \end{aligned} \tag{2.34}$$

remains finite. Thus, the limiting procedure replaces two finite parameters, each having engineering dimensions of length, with a single finite parameter having engineering dimensions of [length]^2.

From a more mathematical point of view, the continuum limit generates a random fractal[6]. As with all fractals, the infinite intrinsic length scale (in this case $\mathcal{L}_0$) is a signal that the fractal covers some finite portion of a larger space. The size of this space is measured by the Hausdorff or fractal dimension d_f. For a TN, the number of points scales like ${\mathcal{L}_0}^D$ and the size of the TN, R_g, scales like ${\mathcal{L}_0}^\nu$ (assuming ℓ fixed). Hence the the number of points available to cover the plane scales like $R_g^{D/\nu}$ and $d_f = D/\nu$. For a RW, $D = 1$ and $\nu_0 = \frac{1}{2}$, so $d_{f0} = 2$. Thus, a RW generates a so-called "Brownian area", which covers some finite fraction of the plane when the RW is projected onto an $\mathbb{R}^2$ slice of $\mathbb{R}^d$. This Brownian area is measured by $\mathcal{L}$.

[6] A random fractal is an object whose scaling behavior is independent of the length scale at which it is probed (hence the term fractal). In the case of the RW, $R_{g0} \sim \mathcal{L}^{1/2}$, independent of the size of $\mathcal{L}$. However, unlike a regular fractal, the precise shape varies randomly from one part of the object to another (hence the term random).

2.4 Continuum models of TN's

Having discussed the meaning of the continuum limit, we now apply this knowledge to the Hamiltonians described in Sec. 2.1.2.

2.4.1 The Gaussian model

It has already been mentioned in Sec. 2.1.2 that for practical calculations $\beta H_t^{(L)}$ (Eq. (2.3)) must be replaced by $\beta H_0^{(L)}$ (Eq. (2.6)). Another reason for making this replacement is that $\beta H_t^{(L)}$ does not have a well-defined continuum limit, but $\beta H_0^{(L)}$ does. Starting from ($D = 1$)

$$\lim_{\ell \to 0} \frac{|\mathbf{r}(s) - \mathbf{r}(s')|}{\ell^2} = \frac{\partial \mathbf{r}(s)}{\partial s}, \tag{2.35}$$

it follows that

$$\lim_{\ell \to 0} \beta H_0^{(L)} = \beta H_0 = \frac{1}{2} \int_0^{\mathcal{L}} ds \left(\frac{\partial \mathbf{r}}{\partial s} \right)^2. \tag{2.36}$$

(Continuum versions of all physical quantities will be distinguished from their discrete counterparts by dropping the (L) superscript.) Note that ℓ^2 appears in the denominator of Eq. (2.35), not ℓ. This is a direct consequence of the fractal nature of the continuum limit. From another point of view, the requirement that βH_0 be dimensionless can only be fulfilled if $\ell^2 \to ds$.

The extension to an arbitrary value of D is accomplished by a natural generalization[7] of Eq. (2.36),

$$\beta H_0 = \frac{1}{2} \int_0^{\mathcal{L}} d^D s \sum_{i=1}^{D} \left(\frac{\partial \mathbf{r}}{\partial s_i} \right)^2. \tag{2.37}$$

The resulting partition function,

$$Z_0 = \int D[\mathbf{r}] \exp\left(-\beta H_0\right), \tag{2.38}$$

can readily be evaluated using standard functional integration techniques [49].

[7] Strictly speaking, $g(D) = \int d^D s\, f(s)$ is not defined for non-integer values of D. The correct approach is to regard D as an integer-valued parameter of g and then to attempt an analytic continuation of $g(D)$ to real values of D. The notations $\int d^D s$ and $\sum_1^D$ are merely a shorthand for this procedure.

A comparison of the results from analysis of the βH_0 model with the results of MC studies of the $\beta H_0^{(L)}$ model indicate that they are in the same universality class [32].

Taking the continuum limit also generates a scale-invariant theory. We can see this by first rewriting H_0 in terms of the dimensionless parameters $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$, defined through

$$\begin{aligned} \mathbf{s} &= \boldsymbol{\sigma}\mathcal{L} \quad \text{with } 0 \leq \boldsymbol{\sigma} \leq 1 \\ \text{and} \quad \mathbf{r} &= \boldsymbol{\rho}\mathcal{L}^{\nu_0}, \end{aligned} \tag{2.39}$$

which gives

$$H_0 = \frac{1}{2}\mathcal{L}^{2\nu_0+D-2}\int_0^1 d^D\sigma \left(\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}}\right)^2. \tag{2.40}$$

The integral is now a pure number, so the only way H_0 can remain dimensionless is if it is scale invariant (i. e., independent of $\mathcal{L}$). Requiring scale invariance gives

$$\nu_0 = \frac{2-D}{2}. \tag{2.41}$$

From this scale invariance we can also obtain the scaling behavior of $\langle R_{g0}\rangle$. From Eqs. (2.9) and (2.12) we find

$$\left\langle R_{g0}^2 \right\rangle \sim \sum_{\mathbf{s},\mathbf{s}'} \left\langle \left[\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')\right]^2 \right\rangle \sim \mathcal{L}^{2\nu_0} \langle \cdots \rangle , \tag{2.42}$$

where $\langle \cdots \rangle$ is dimensionless. Thus, $\langle R_{g0}\rangle \sim \mathcal{L}^{\nu_0}$.

The fractal dimension for a TN is

$$d_{f0} = \frac{D}{\nu_0} = \frac{2D}{2-D}. \tag{2.43}$$

For $D = 1$ we recover the result that a RW effectively fills the plane. On the other, a TM ($D = 2$) fills any space, regardless of its dimensionality.

2.4.2 The Edwards model

Having dealt with the tethering constraint, we now turn to the issue of self-avoidance. The continuum limit of a SATN is thought to be represented by the Edwards model [20, 50],

$$H_{\mathrm{E}} = H_0 + H_{\mathrm{SA}}, \tag{2.44}$$

where

$$\begin{aligned}\beta H_{\rm SA} &= \frac{v_2}{2}\int_{\substack{0\\ \mathbf{s}\neq\mathbf{s}'}}^{\mathcal{L}} d^D s d^D s\, \delta^d[\mathbf{r}(\mathbf{s})-\mathbf{r}(\mathbf{s}')] \\ &= \frac{v_2}{2}\mathcal{L}^{2D-d\nu_0}\int_{\substack{0\\ \boldsymbol{\sigma}\neq\boldsymbol{\sigma}'}}^{1} d^D\sigma d^D\sigma'\,\delta^d[\boldsymbol{\rho}(\boldsymbol{\sigma})-\boldsymbol{\rho}(\boldsymbol{\sigma}')]. \end{aligned} \tag{2.45}$$

v_2 measures the two-body interaction strength and has dimensions $[v_2] = \mathcal{L}^{\nu_0 d-2D}$. $\beta H_{\rm SA}$ does not contribute to $H_{\rm E}$ unless the configuration self-intersects, i. e., $\mathbf{r}(\mathbf{s}) = \mathbf{r}(\mathbf{s}')$ for at least one pair of distinct $\mathbf{s}$-values. Therefore, it has no effect on SA configurations, but it does generate a finite positive value for self-intersecting configurations. The restriction $\mathbf{s} \neq \mathbf{s}'$ prevents self-interaction of individual points. The full Hamiltonian is

$$\beta H_{\rm E} = \frac{1}{2}\int_0^1 d^D\sigma \left(\frac{\partial\boldsymbol{\rho}}{\partial\boldsymbol{\sigma}}\right)^2 + \frac{z}{2}\int_{\substack{0\\ \boldsymbol{\sigma}\neq\boldsymbol{\sigma}'}}^{1} d^D\sigma d^D\sigma'\,\delta^d[\boldsymbol{\rho}(\boldsymbol{\sigma})-\boldsymbol{\rho}(\boldsymbol{\sigma}')], \tag{2.46}$$

where

$$z \equiv v_2\mathcal{L}^{2D-(2-D)d/2}. \tag{2.47}$$

Once again, the scaling behavior is completely determined by the exponent of $\mathcal{L}$. The model remains scale invariant along the curve

$$\begin{aligned} d(D) &= \frac{4D}{2-D} \\ &\equiv d_{uc}, \end{aligned} \tag{2.48}$$

so that, in the thermodynamic limit,

$$\mathcal{L}\to\infty \Rightarrow z \to \begin{cases} 0 & \text{for } d > d_{uc} \quad \text{SA irrelevant} \\ v_2 & \text{for } d = d_{uc} \quad \text{SA marginal} \\ \infty & \text{for } d < d_{uc} \quad \text{SA relevant} \end{cases}. \tag{2.49}$$

Thus, when $d > d_{uc}$, the SA configurations are given the same weight as the self-intersecting configurations, and SA is irrelevant. When $d = d_{uc}$, the self-intersecting configurations are

suppressed, but not very strongly. In this situation $\beta H_{\rm SA}$ gives logarithmic corrections to R_g, but does not effect the scaling exponents [20]. When $d < d_{uc}$, any positive contribution arising from integrating over a self-intersecting configuration is multiplied by $z \to \infty$. Thus, in the thermodynamic limit, self-intersecting configurations are *completely excluded* from the partition function [51]. The Edwards model is, therefore, truly self-avoiding.

It is also possible to consider n-body interaction terms such as

$$H_{\rm SA}^{(n)} = \frac{v_n}{2} \int_0^{\mathcal{L}} d^D s_1 \cdots d^D s_n \, \delta^d[\mathbf{r}(\mathbf{s_1}) - \mathbf{r}(\mathbf{s_2})] \ldots \delta^d[\mathbf{r}(\mathbf{s_{n-1}}) - \mathbf{r}(\mathbf{s_n})]. \tag{2.50}$$

These terms also have a critical line,

$$\begin{aligned} d_{ucn} &= \left(\frac{n}{n-1}\right)\frac{D}{\nu_0} \\ &= \left(\frac{n}{n-1}\right)\frac{2D}{2-D}. \end{aligned} \tag{2.51}$$

However, the $n > 2$ terms are never relevant in the thermodynamic limit if the $n = 2$ term is present. We can see this from the scaling coefficient,

$$z_n = v_n \mathcal{L}^{-n(d\nu - D) + d\nu}, \tag{2.52}$$

for each term $H_{\rm SA}^{(n)}$. Naive scaling analysis shows that that, near the Gaussian fixed point, the large-n terms can be the most important ones. This will occur when $D > 2d/(2+d)$. However, renormalization will drive the system away from the Gaussian fixed point and towards another fixed point with a larger value of ν. As ν increases, the small-n terms become increasingly important. For $\nu > D/d$ (the "close-packed" lower bound on ν for a SATN; see Eq. (2.25)), the $n = 2$ term dominates all others. Hence, the renormalized interaction will always be dominated by the $n = 2$ term.

This effect has been studied in some detail, for the case of polymers, using so-called *k-tolerant walks* [52]. In a k-tolerant walk, the walker is allowed to visit a site as many as k times, but not more. In the Edwards-model language, a k-tolerant walk has the k-body term as its lowest order bare interaction term. Renormalization group analyses [53] and

more formal treatments [54] show (although not rigorously) that the k-tolerant walk is in the same universality class as the SAW. Thus, the two-body interaction in the renormalized Edwards model always dominates the interaction terms even when it is initially not present. We will return to a discussion of the higher-order interactions in Sec. 4.5.2.

In the preceding approach, the scale invariance of H_0 was enforced and $H_{\rm SA}$ was allowed to become relevant, marginal (scale invariant), or irrelevant as a function of d. When $H_{\rm SA}$ is irrelevant, it is obvious that $\nu = \nu_0$. This is also true (but not obvious) when $H_{\rm SA}$ is marginal. The situation is more complicated when $H_{\rm SA}$ is relevant, since it is not easy to characterize the forbidden configurations. We can get an estimate of ν by replacing the requirement that $H_{\rm E}$ be scale invariant with the demand that both terms in Eq. (2.46) have the same scaling behavior. Equating the exponents of the $\mathcal{L}$-factors in Eqs. (2.40) and (2.45) leads to [20]

$$\nu_F = \frac{D+2}{d+2}. \tag{2.53}$$

This approach has no rigorous foundation and is, in fact, equivalent to a mean-field analysis of the theory. This mean-field argument (originally due to Flory [4, 55] in the context of polymers) consists of writing the free energy as a sum of two competing terms,

$$F \sim R_g^2 + v_2 \frac{N^2}{R_g^d}. \tag{2.54}$$

The first term is the elastic energy required to stretch a TN to a size R_g, and it corresponds to H_0 in the Edwards model. The second term is the SA-interaction energy, which scales like the square of the density of monomers for a two-body interaction. It is here that the mean-field approximation is introduced, since the free-energy density, which is expected to depend on the density-density correlations $\langle c^2 \rangle$, is replaced by $\langle c \rangle^2 \sim N^2/R_g^{2d}$. This term must then be multiplied by a volume factor $\sim R_g^d$ to obtain an energy. Recalling that $N \sim L^D \sim R_g^{D/\nu}$ and minimizing F with respect to R_g leads to Eq. (2.53). Despite the approximate nature of the argument, ν_F gives a surprisingly good estimate of ν for polymers and Sierpiński gaskets (see Chap. 5).

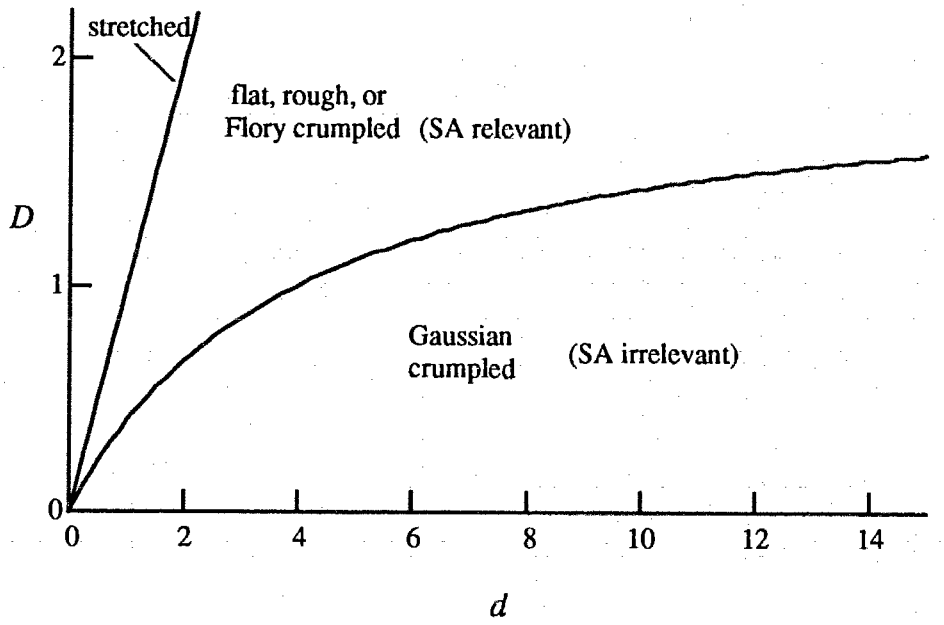


Figure 2.4: The D versus d phase diagram. Along the line labelled "stretched" $D = d$ and the TN fills the embedding space, so $\nu = 1$. The other curve, which separates the SA-relevant and SA-irrelevant regimes is a plot of Eq. (2.48). Several points in the (D, d)-plane are studied in Chaps. 4 and 5.

It is implicit in the above discussion that D and d are integers, but the analytic continuation of the Edwards model theory to real-valued D and d is mathematically well-defined and potentially interesting. There are two reason for interest in this generalization. The first is that analysis of the Edwards model is difficult and can be done only using perturbation theory. In the most successful method, known as the ϵ-expansion, the theory is first solved at a point (D^*, d^*) along the critical line $d = d_{uc}$. Because self-avoidance is marginal along this line, the theory can be solved exactly. The theory is evaluated at the point of interest, (D_{phys}, d_{phys}), by regarding D_{phys} and d_{phys} as perturbations of D^* and d^* [20]. For this procedure to work, D and d must be regarded as continuous parameters.

From a second and more mathematical point of view, it is desirable to understand the structure of the theory over the entire (D, d)-plane (see Fig. 2.4), not only because

the structure may be intrinsically interesting, but also because it may provide insight into the theory defined at the physical points of the (D, d)-plane. This is especially true of membranes, for which $d_{uc} = \infty$. If the theory were restricted to integer-valued points in the (D, d)-plane, the physical point $(D = 2, d = 3)$ would be infinitely far away from where the theory could be solved. On the other hand, the physical point $(D = 2, d = 3)$ is only a finite distance away from the point $(D = 6/7, d = 3)$, which lies along the critical line $d = d_{uc}$. It is not yet clear whether $d_{uc} = \infty$ signals a regime with qualitatively new physics or, on the other hand, whether it simply creates computational difficulties for the perturbation theory. Exploring other regions of the (D, d)-plane may give some insight on this matter, and is, in fact, the motivation for the work in Chaps. 4 and 5.

Thou com'st in such a questionable shape,
That I will speak to thee.
William Shakespeare
Hamlet

Chapter 3

Two-Dimensional Closed Pressurized Random Walks

3.1 Introduction

In recent years there has been a surge of activity in the study of artificial and naturally occurring fluid vesicles [22]. Much of this activity has focused on understanding the various shape conformations of these systems. Although some vesicle shapes can be explained by studying the $T = 0$ energetics [56], there remain many questions about the role of thermal fluctuations [57]. To address these more difficult questions requires a statistical mechanical treatment of a model that incorporates the most important properties of fluid vesicles:

1. A vesicle has a fluid as opposed to tethered membrane. Alternatively, one can think of a fluid membrane as a TN whose connectivity is allowed to evolve over time. This is, in fact, how simulations of fluid membranes are done [58].

2. A vesicle membrane is self avoiding.

3. The energetics are controlled mainly by a surface bending rigidity [56], κ, that is sufficiently large so that the correlation length induced by it, l_κ, is comparable to the vesicle size, R. This turns out to be a crucial distinguishing feature of a vesicle[1]; if $l_\kappa \ll R$, the vesicle is simply a pressurized self-avoiding surface (or pressurized self-avoiding walk (PSAW) in the two-dimensional analogue), and many of the interesting shapes that resemble those seen in experiments on real vesicles disappear. Mindful of this distinction, the term vesicle will henceforth be used to describe only models for which $l_\kappa \gtrsim R$.

4. The surface area exceeds the minimum needed for the enclosed volume (the vesicle is deflated). There may also be, in general, a pressure difference, p, between the inside and the outside of the vesicle.

The non-local nature of the self-avoiding interaction and the complications that fluid surfaces create for ensemble averages make this model very difficult to analyze for $T > 0$. As a result, only mean-field analyses of phenomenological models have been done so far [56]. These difficulties have led some people to consider studying vesicle models via computer simulations, but it is only very recently [59, 60] that the computer power needed to do reasonable simulations of $d = 3$ fluid vesicles has become available.

3.1.1 Simulation models

In the time between these two developments, M. E. Fisher and collaborators [2, 23, 24, 25, 26, 27] have used Monte Carlo simulations to investigate a two-dimensional model that contains features 2–4 listed above (The distinction between fluid and tethered membranes is lost for equilibrium properties of one-dimensional surfaces embedded in two dimensions). Specifically, the model (referred to as the LSF model) consists of hard disks of diameter a connected by tethers of length $\frac{9}{5}a$ to form a ring. The bending energy term in the

[1] I thank Udo Seifert for pointing this out.

Hamiltonian is $(\kappa/a)\sum_{j=1}^{N}(1-\cos\theta_j)$, where θ_j is the angle determined by two adjacent links. There is also a (pressure)×(area) energy term, which accounts for the work done at constant pressure when the enclosed area of the vesicle changes.

In their original work [2], LSF observed scaling behavior for the enclosed area and the radius gyration of a PSAW and found different shape classes as a function of pressure. They also found dumbbell-like and ellipsoidal vesicles coexisting for certain combinations of negative pressure and bending rigidity. Subsequent work by the same group [24] demonstrated that the shape parameter, s_2, of a PSAW can be varied continuously by adjusting the pressure. Other work by this group includes studies of the scaling behavior of the enclosed area and radius of gyration of a PSAW in the high pressure regime [25] and studies on vesicles in the large bending rigidity limit [26]. Most recently [27] they have begun a detailed investigation of a negative pressure-rigidity phase diagram, and have found a wide and interesting variety of shape classes (e. g. , ellipsoidal, dumbbell-like, rod, and branched polymer). Susceptibility measurements from these simulations suggest that some of the shapes may be separated by phase boundaries. Most of this work is reviewed in Refs. [23, 27].

Boal [61] has studied a similar model, but with the bending rigidity replaced by an attractive potential. The resulting pressure-attraction phase diagram includes a branched-polymer phase and a dense phase, as well as the usual self-avoiding walk phase.

3.1.2 Continuum models

Computer simulations are certainly useful for studying vesicle shapes, but they cannot tell the whole story. Questions about the existence of phase transitions, such as those discussed in Ref. [27], are particularly difficult to answer from simulation data because of finite-size effects. Mean-field calculations are also unreliable in this respect. Thus, it is desirable to complement these computer studies with analytic work. Ideally, one would like to analyze a microscopic model with the same features as the LSF model. Such a model can be described

by a Hamiltonian with three parts:

$$H_{\rm LSF} = H_{\rm SA} + H_\kappa + H_p^{(abs)}; \tag{3.1}$$

$$\beta H_{\rm SA} = v_2' \int_0^{\mathcal{L}_0} ds_1 ds_2 \delta(\mathbf{r}(s_1) - \mathbf{r}(s_2)) \tag{3.2}$$

is identical to Eq. (2.45) except for the replacements $v_2 \to v_2'$ (these interaction strengths have different engineering dimensions) and $\mathcal{L} \to \mathcal{L}_0$ (The need for this distinction will be made clear at the end of this subsection.);

$$\beta H_\kappa = \int_0^{\mathcal{L}_0} ds \frac{\kappa}{2} \left(\frac{d^2\mathbf{r}}{ds^2} \right)^2 \tag{3.3}$$

is the natural continuum realization of the discretized bending energy $\hat{\mathbf{n}}(k) \cdot \hat{\mathbf{n}}(k+1)$ [48], where $\hat{\mathbf{n}}(k)$ is the unit normal of the kth link; and the last term,

$$\beta H_p^{(abs)} = -p \left| a[\mathbf{r}] \right|, \tag{3.4}$$

where

$$a[\mathbf{r}] = \frac{1}{2} \int_0^{\mathcal{L}_0} ds \epsilon^{ij} r_i \frac{dr_j}{ds}, \tag{3.5}$$

accounts for the energy due to the pressure difference p. Eq. (3.4) employs the convention that the inside pressure is greater than the outside pressure when $p > 0$. The functional $a[\mathbf{r}]$ (referred to as the algebraic area) is different from the true area, since $a[\mathbf{r}]$ is sensitive to the direction that the loop is traversed. Hence, the absolute value must be used.

The last ingredient needed to complete the model is the constraint,

$$\left(\frac{d\mathbf{r}}{ds} \right)^2 = 1, \tag{3.6}$$

which fixes the distance between adjacent links. This rigid-rod constraint does not correspond exactly to the LSF-model constraint (which uses flexible tethers), but, as discussed in Chap. 1, the long-distance behaviors and the continuum limits for these models are the same.

Eqs. (3.2)–(3.6) implicitly assume that the arc-length, s, is a continuous parameter. On the other hand, the evaluation of the partition function,

$$Z_{\rm LSF} = \int D[\mathbf{r}]\,\delta\left[\left(\frac{d\mathbf{r}}{ds}\right)^2 - 1\right]\exp\left(-\beta H_{\rm LSF}\right), \tag{3.7}$$

requires s to be discretized [48] into N links, each of length ℓ, so that

$$\mathcal{L}_0 \equiv N\ell \tag{3.8}$$

measures the perimeter of the vesicle. After the functional integral in Eq. (3.7) is performed, the continuum limit is obtained by the prescription

$$N \to \infty, \ell \to 0, \text{ such that } \mathcal{L}_0 \text{ is fixed.} \tag{3.9}$$

Eqs. (3.1)–(3.9) describe a continuum version of the LSF model, and should contain the same physics as the discrete version. Unfortunately, this model has several features that destroy its otherwise Gaussian character, thus rendering an exact calculation hopeless. I will now describe a sequence of models that are obtained by removing these non-Gaussian elements one at a time, eventually leading to a model that can be analyzed completely in closed form. The price that will be paid for this increased tractability will be a significant change in the physics that the resulting model describes.

The first task is to remove the two non-Gaussian elements of $H_{\rm LSF}$. One of these elements is $H_{\rm SA}$. While $H_{\rm SA}$ could be included in the analysis with the aid of the ϵ-expansion described in Sec. 2.4.2, the necessary starting point is still the exact solution of the Gaussian field theory. For the sake of simplicity, all subsequent discussion will ignore $H_{\rm SA}$. However, Sec. 3.3.1 does contain a brief discussion of Flory-theory self avoidance.

The second non-Gaussian element of $H_{\rm LSF}$ is the absolute value in Eq. (3.4), which is required to ensure that the area is positive. This problem can be handled simply by making the replacement

$$\beta H_p^{(abs)} \to \beta H_p = -pa[\mathbf{r}], \tag{3.10}$$

but this creates two problems of its own. The first occurs because closed random walks that cross themselves have both positive and negative contributions to $a[\mathbf{r}]$; this effect will become especially significant when $|p|$ is small. Of course, if $H_{\rm SA}$ were retained in the Hamiltonian, this problem would not arise. The second problem is that, for every vesicle with algebraic area A, there is an identical one with algebraic area $-A$. In the presence self avoidance, these two regimes are separated by an infinite energy barrier, but this does not solve the problem as far as equilibrium properties are concerned. This is a serious defect in the model and there does not seem to be any way to circumvent it without destroying the quadratic nature of H_p. I will return to this issue in Sec. 3.4.1.

The changes required to make the Hamiltonian quadratic in $\mathbf{r}$ are summarized by the equation

$$H_{\rm LSF} \to H_{\rm RR} = H_\kappa + H_p. \tag{3.11}$$

The model described by Eqs. (3.6)–(3.11) will be referred to as the rigid-rod model.

The final non-Gaussian element of the continuum LSF model is the constraint, Eq. (3.6). It can be removed by using the well-known procedure of replacing the constraint with an additional term in the Hamiltonian

$$H_{\rm RR} + \text{Eq. (3.6)} \to H_{\rm PRW} = H_0 + H_\kappa + H_p, \tag{3.12}$$

where

$$\beta H_0 = \frac{1}{4}\int_0^{\mathcal{L}} ds \left(\frac{d\mathbf{r}}{ds}\right)^2 \tag{3.13}$$

is the Gaussian Hamiltonian introduced in Sec. 2.4.1, with

$$\mathcal{L} \equiv N\ell^2. \tag{3.14}$$

The new partition function obtained by making the replacement

$$Z_{\rm LSF} \to Z_{\rm PRW} = \int D[\mathbf{r}] \exp(-\beta H_{\rm PRW}) \tag{3.15}$$

is a Gaussian functional integral.

Although Eq. (3.12) achieves the desired effect of replacing an unmanageable constraint with a Gaussian integral, it constitutes a major change to the model; rigid rods have been replaced by infinitely stretchable springs. The new model described by H_{PRW} will be referred to as the pressurized random walk (PRW) model. Despite their differences, the PRW model and the continuum LSF model should contain the same physics provided that they are compared in a regime where Eq. (3.6) does not strongly restrict the phase space of Z_{LSF}. However, we should not expect them to agree under conditions that produce highly stretched configurations in the rigid-rod model.

Just as with Z_{LSF}, Z_{PRW} is evaluated by discretizing the arc-length, s, into N links, each of length ℓ. However, the replacement of the constraint in Eq. (3.9) with H_0 requires that a different continuum limit be taken:

$$N \to \infty, \ell^2 \to 0, \text{such that } \mathcal{L} \text{ is fixed.} \tag{3.16}$$

(cf. Eq. (3.9)). The reason why H_0 requires this particular continuum limit has already been discussed in Sec. 2.4.1. Also, recall that ℓ no longer represents a microscopic length scale for the system in the usual sense. For example, the perimeter, $N\ell$, is divergent in this limit. In fact, the presence of an external field means that, even before the continuum limit is taken, specifying N and ℓ does not fix the perimeter (see the discussion below Eq. (3.24)).

The absence of a well-defined microscopic length scale has a profound effect on the model — particularly on the H_κ term. Consequently, all of the analyses in Sec. 3.3 deal only with the PRW_0 model, which is just the $\kappa = 0$ version of the PRW model. Sec. 3.4.2 discusses why the H_κ term in the PRW model does not correspond to an ordinary bending rigidity. This surprising effect of H_κ in the PRW model is particularly unfortunate because it is the existence of a large l_κ, more than anything else, that distinguishes the physics and especially the shapes of vesicles from those of random walks.

3.2 Asphericity calculations

Computer simulation studies of shapes have the luxury of being able to produce and to study individual configurations, and to compare them with real photomicrographs. However, in the absence of a quantitative measure of the shape, all such comparisons are subjective. Quantitative measures of polymer shapes have been proposed and investigated for some time now [44, 62]. In particular, both A_d and a_d (defined by Eqs. (2.28) and (2.30), respectively) have been measured from Monte Carlo simulations [63, 64]. Rudnick and coworkers [65, 66, 67] have described methods for calculating a_2, and have applied them to the PRW_0 model [28].

Although A_d is generally regarded as a better measure of the shape of a typical configuration, almost all calculations to date [65, 67, 68] have been of a_d, because calculation of A_d was considered difficult except in the $d \to \infty$ limit [66]. Recently, however, Diehl and Eisenriegler [69] have described a method for calculating the asphericity in any dimension, and have used it to calculate the shape of ordinary open and closed random walks embedded in a space of arbitrary dimension. The method uses a trick, which consists of writing the ensemble average of a ratio in the form

$$\left\langle \frac{a}{b^n} \right\rangle = \frac{1}{(n-1)!} \int_0^\infty dy y^{n-1} \left\langle a e^{-yb} \right\rangle. \tag{3.17}$$

Applying this trick to Eq. (2.28) gives ($n = 2$)

$$\begin{aligned} A_2 &= 2 \int_0^\infty dy y \left\langle \mathrm{Tr}\left(\hat{Q}^2\right) \exp\left(-y \mathrm{Tr} Q\right) \right\rangle \\ &= 2 \int_0^\infty dy y \left\langle \mathrm{Tr}\left(\hat{Q}^2\right) \right\rangle_{H(y)} \frac{Z(y)}{Z(0)}, \end{aligned} \tag{3.18}$$

where $\langle \cdots \rangle_{H(y)}$ denotes an ensemble average with respect to the Hamiltonian

$$\begin{aligned} \beta H(y) &= \beta H_{\mathrm{PRW}} + y \mathrm{Tr} Q \\ &= \int_0^{\mathcal{L}} ds \left[\frac{1}{4} \left(\frac{d\mathbf{r}}{ds} \right)^2 + \kappa \left(\frac{d^2\mathbf{r}}{ds^2} \right)^2 - \frac{p}{2} \epsilon^{ij} r_i \frac{dr_j}{ds} \right] \\ &\quad + \frac{y}{2\mathcal{L}^2} \int_0^{\mathcal{L}} ds_1 ds_2 \left[\mathbf{r}(s_1) - \mathbf{r}(s_2) \right]^2, \end{aligned} \tag{3.19}$$

and

$$Z(y) = \int D[\mathbf{r}]\, e^{-\beta H(y)}. \tag{3.20}$$

Inspection of Eq. (3.19) shows that $H(y)$ is quadratic in $\mathbf{r}$, so the asphericity can be calculated exactly.[2]

3.3 Results ($\kappa = 0$)

3.3.1 Mean-field results

Although virtually all interesting physical quantities can be extracted from the partition function in closed form, it is nonetheless instructive to examine first a simplified version of the theory whose configurations are restricted to circles. As we shall see, this restriction, which effectively generates a mean-field theory for the model, does not have an important effect on the physics discussed here. The energy of a circular configuration of radius R is

$$\beta H_{\mathrm{C}} = \frac{\pi^2 R^2}{\mathcal{L}} \left(1 - \frac{|\rho|\, R^{d-2}}{\rho_c} \right), \tag{3.21}$$

where

$$\rho \equiv p\mathcal{L}, \tag{3.22}$$

and $\rho_c = \pi$. Inspection of Eq. (3.21) shows that there is an instability for $d = 2$, with

$$R = \begin{cases} 0 & \text{for } \rho < \rho_c \\ \infty & \text{for } \rho > \rho_c \end{cases}, \tag{3.23}$$

which occurs because both terms in H_{C} scale as R^2. The trivial $\rho < \rho_c$ behavior predicted here reflects the complete absence of fluctuations in the above treatment. They can be introduced in a mean-field way by allowing R to fluctuate and then averaging over all values of R. This procedure gives

$$\langle R \rangle = \frac{1}{\pi} \sqrt{\frac{\mathcal{L}}{\rho_c - \rho}} \tag{3.24}$$

[2] I wish to thank H.-W. Diehl for pointing out the tractability of the A_d calculation for the PRW model, and for encouraging me to pursue the calculation.

The instability remains, but R now has the correct qualitative behavior for $0 < \rho < \rho_c$ (cf. Eqs. (3.29) and (3.32) below).

Eq. (3.24) clearly illustrates an important difference between the rigid-rod model and the PRW_0 model. In the rigid-rod model, R can never diverge, even at infinite pressure. Furthermore, the result in Eq. (3.24) is independent of whether or not the continuum limit is taken. Thus, even for N and ℓ finite, specifying $\mathcal{L}$ does not uniquely determine R since it also depends on p. Put another way, for a fixed value of N, the link length is not fixed by the model, but instead depends on p.

Mean-field methods can also be applied to other values of d. For $d = 1$, $p \to f$ (a force), and mean-field theory recovers the well-known result $R \sim f\mathcal{L}$. Once again, we see that fixing $\mathcal{L}$ does not set an upper bound on the overall size of the system. For $d = 3$ and $\rho = 0$, H_C has a minimum at $R = 0$, but for any $\rho > 0$ there are two minima separated by a finite barrier:[3] a local one at $R = 0$ and a global one at $R = \infty$. Thus, a collapsed system with no pressure difference between the inside and outside will spontaneously inflate some time after a pressure difference is introduced.

Self avoidance can be included by adding the Flory-theory term $v_2\mathcal{L}^2/R^2$ [4] to βH_C. In two dimensions the value of ρ_c is unchanged, but R is shifted away from zero, reflecting the system's reduced ability to collapse. In three dimensions the global minimum of βH_C is still at infinity, but the metastable R is shifted away from zero to a point which, once again, reflects the system's reduced ability to collapse. The critical value of ρ, above which the metastable region vanishes, depends on the coefficients of the terms in βH_C.

Finally, we can understand why the mean-field results are so accurate by Fourier analyzing the shape, and studying the mean-field behavior of higher frequency terms. The energy of an nth harmonic component with amplitude R_n is

$$\beta H_C^{(n)} = \frac{n^2\pi^2 R_n^2}{\mathcal{L}}\left(1 - \frac{|\rho|}{n\rho_c}\right), \tag{3.25}$$

[3]To make sense, the random walk must be generalized to a random surface, but within the context of mean-field theory, the R scaling behavior of H_C is the same for both.

and leads to

$$\frac{\langle R_n \rangle}{\langle R \rangle} = \sqrt{\frac{\rho_c - \rho}{n\,(n\rho_c - \rho)}}, \tag{3.26}$$

which vanishes as $\rho \to \rho_c$, if $n \geq 2$. Thus, the higher frequency terms become less important as $\rho \to \rho_c$, and, as a result, the inflated shape is a circle.

3.3.2 Exact results

Most of the results in this subsection have been published elsewhere [28, 39]. They are included here for the sake of completeness, and to allow ready comparison with the mean-field results. The first step in obtaining the exact results is the evaluation of $Z(y)$. A straightforward[4] but tedious calculation gives [39]

$$Z(y)|_{\kappa=0} = \frac{y}{2\pi\,(\cosh w - \cos\rho)}, \tag{3.27}$$

where

$$w = \sqrt{4y\mathcal{L} - \rho^2}. \tag{3.28}$$

Eq. (3.27) can be used to calculate the expectation value of the algebraic area [28],

$$\begin{aligned} \mathcal{A} \equiv \frac{\langle a[\mathbf{r}] \rangle}{\mathcal{L}} &= \frac{\partial}{\partial\rho} \log\left(\lim_{y\to 0} Z(y)\Big|_{\kappa=0} \right) \\ &= \frac{1}{\rho} - \cot\rho. \end{aligned} \tag{3.29}$$

The exact result contains the same instability as the mean-field result and has the same qualitative behavior for $\rho < \rho_c$ as Eq. (3.24). Note the symmetry $\mathcal{A}(\rho) = -\mathcal{A}(-\rho)$. This is due entirely to the algebraic nature of $a[\mathbf{r}]$. Since there is no additional information in the $\rho < 0$ sector of the theory, all subsequent discussions will implicitly assume $\rho \geq 0$, unless stated otherwise. Also note that, despite the differences between the PRW_0 model studied here and the LSF model, they have similar scaling behavior in some regions of parameter

[4] The only difficulty in evaluating $Z(y)$ arises from the non-local nature of the y-dependent term. This term can be converted to a Gaussian integral over a local field using a trick described in Ref. [39] and in Appendix A.

space. Eq. (3.29) is a special case of the scaling form for the area proposed in Ref. [2]. The only change required is $\nu_{SA} \to \nu_0$.

The fluctuations in the algebraic area are measured by

$$\begin{aligned} \chi \equiv \frac{1}{\mathcal{L}^2}\left[\left\langle a^2[\mathbf{r}]\right\rangle - \langle a[\mathbf{r}]\rangle^2\right] &= \frac{\partial^2}{\partial\rho^2}\log\left(\lim_{y\to 0} Z(y)\Big|_{\kappa=0}\right) \\ &= \frac{1}{\sin^2\rho} - \frac{1}{\rho^2}. \end{aligned} \tag{3.30}$$

Note that $\lim_{\rho\to 0}\chi = \frac{1}{3}$, which shows that a significant fraction of the configurations have a net algebraic area for $\rho = 0$, even though the antisymmetry of $\mathcal{A}$ with respect to ρ requires $\lim_{\rho\to 0}\mathcal{A} = 0$. At the critical point, ρ_c, the fluctuations diverge; but

$$\lim_{\rho\to\rho_c} \chi/\mathcal{A}^2 = 1 \tag{3.31}$$

shows that they are the same size as the average algebraic area.

The average radius of gyration can also be easily computed [28].

$$\begin{aligned} \mathcal{R}^2 \equiv \frac{\left\langle R_g^2\right\rangle}{\mathcal{L}} &= -\mathcal{L}\left.\frac{\partial \log Z(y)}{\partial y}\right|_{\kappa=y=0} \\ &= \frac{1}{\rho^2} - \frac{\cot\rho}{\rho}. \end{aligned} \tag{3.32}$$

Once again, this result is a special case of the scaling proposed in Ref. [2]. In the limit $\rho \to 0$, Eq. (3.32) reduces to the well-known result $\mathcal{R}^2 = \frac{1}{3}$.

3.3.3 Shapes

One of the simplest measures of the shape is $\mathcal{A}/\mathcal{R}^2 = \rho$. A slightly more sophisticated version of this measure can be obtained using Eq. (3.17),

$$\begin{aligned} \left\langle \frac{a[\mathbf{r}]}{R_g^2}\right\rangle &= \frac{1}{Z(0)}\int_0^\infty dy \mathcal{L}\frac{\partial}{\partial\rho} Z(y) \\ &= 2\mathcal{L}^2\sin\rho\int_0^\infty dy y \frac{\frac{1}{w}\sinh w - \frac{1}{\rho}\sin\rho}{(\cosh w - \cos\rho)^2}. \end{aligned} \tag{3.33}$$

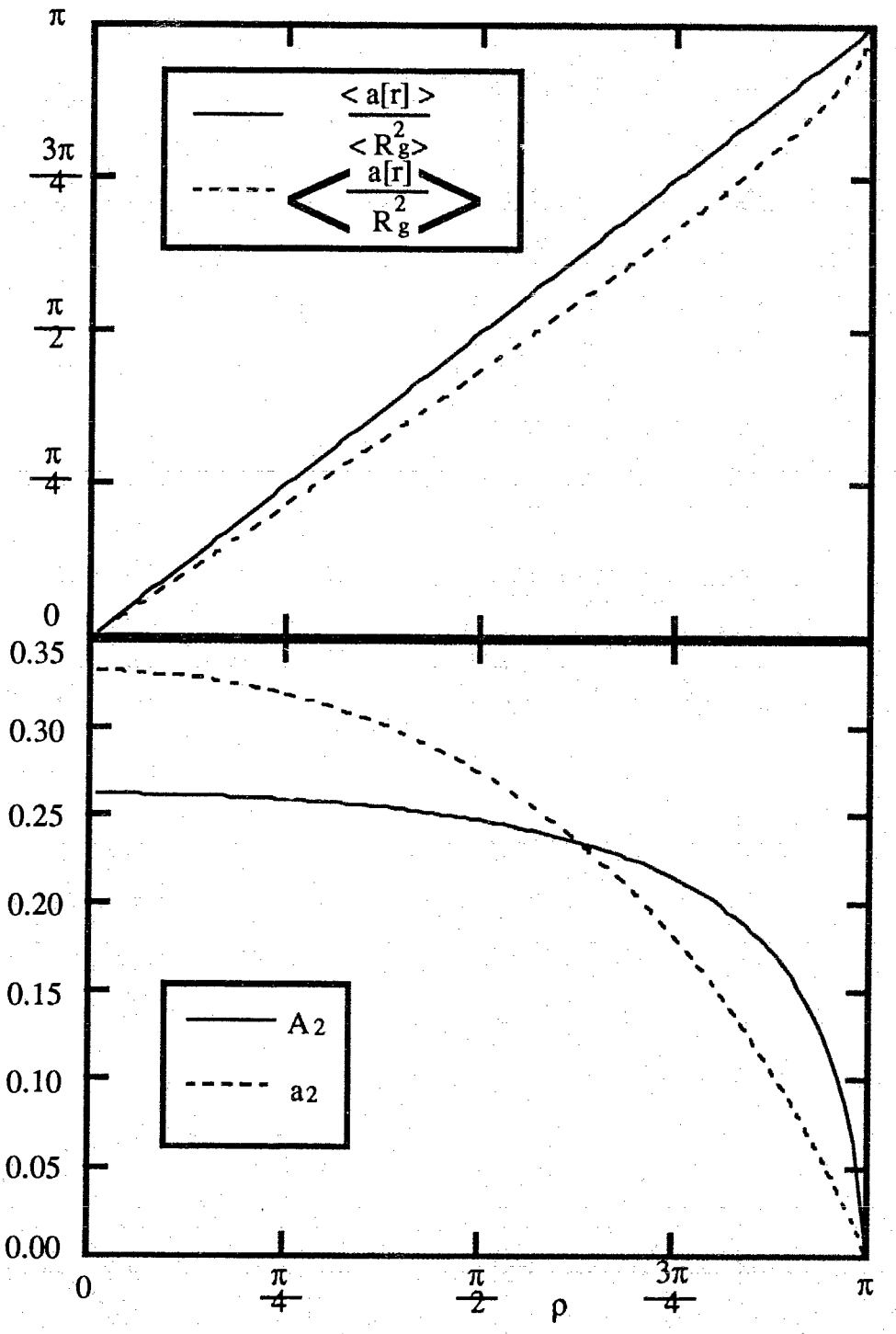


Figure 3.1: (a) A comparison of the shape measures $\mathcal{A}/\mathcal{R}$ and $\left\langle a[\mathrm{r}]/R_g^2 \right\rangle$ as a function of ρ. The difference between them is never more than 20%; (b) A comparison of the asphericities A_2 (Eq. (3.43)) and a_2 (Eq. (15) of Ref. [28]; but see the footnote on page 49).

This integral can be done numerically. The result is plotted in Fig. 3.1a along with $\mathcal{A}/\mathcal{R}^2$ for comparison. The two curves give a quantitative indication of the size of the difference between taking the ratio of the averages and the average of the ratio, which, in this case, is as large as 20%. Also, note that both measures of the shape have the value π for $\rho = \rho_c$. This is consistent with the simple picture that the inflated regime ($\rho \lesssim \rho_c$) consists of circular rings of radius $\mathcal{R}$.

To calculate the asphericity A_2 requires two basic steps (See Appendix A for details.). The first is to compute the generating functional,[5]

$$\begin{aligned} W(y;\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) &= \int D[\mathbf{r}] \exp\left[-\beta H(y) + \boldsymbol{\eta}_1 \cdot \mathbf{r}(s_1) + \boldsymbol{\eta}_2 \cdot \mathbf{r}(s_2)\right] \\ &= \frac{y}{2\pi\Delta} \exp\left\{\frac{L}{w\Delta}\left[\alpha_1 \eta_1^2 + \alpha_2 \eta_2^2 + (\beta\delta_{jk} - \phi\epsilon_{jk})\,\eta_1^j \eta_2^k\right]\right\}, \end{aligned} \tag{3.34}$$

where

$$\Delta = \cosh w - \cos\rho \tag{3.35}$$

$$\alpha_j = \sinh w - \sinh w\sigma_j \cos\rho(1-\sigma_j) - \sinh w(1-\sigma_j)\cos\rho\sigma_j, \tag{3.36}$$

$$\begin{aligned} \beta &= \sinh w + \sinh w(1-\tau)\cos\rho\tau + \sinh w\tau\cos\rho(1-\tau) \\ &- \left[\sinh w\sigma_1 \cos\rho(1-\sigma_1) + \cos\rho\sigma_1 \sinh w(1-\sigma_1) + (\sigma_1 \to \sigma_2)\right], \end{aligned} \tag{3.37}$$

$$\begin{aligned} \phi &= \varepsilon(\sigma_2 - \sigma_1)\left[\sinh w\tau \sin\rho(1-\tau) - \sinh w(1-\tau)\sin\rho\tau\right] \\ &+ \left[\sinh w\sigma_1 \sin\rho(1-\sigma_1) - \sinh w(1-\sigma_1)\sin\rho\sigma_1 + (\sigma_1 \to \sigma_2)\right], \end{aligned} \tag{3.38}$$

$$\sigma_j \equiv \frac{s_j}{\mathcal{L}}, \tag{3.39}$$

$$\tau \equiv |\sigma_2 - \sigma_1|, \tag{3.40}$$

and

$$\varepsilon(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ -1 & \text{for } x < 0 \end{cases}. \tag{3.41}$$

The second step is to use $W(y;\boldsymbol{\eta}_1,\boldsymbol{\eta}_2)$ to calculate two-point and four-point correlation functions such as $\langle r_i(s_1)\, r_j(s_2)\rangle$, which can then be used to construct $\left\langle \mathrm{Tr}\hat{Q}^2 \right\rangle_{H(y)}$. The

[5]There is a typographical error in Eq. (39) of Ref. [29]. The correct expression is given in Eq. (3.34).

result is[6]

$$\left\langle \mathrm{Tr}\hat{Q}^2 \right\rangle_{H(y)} = \frac{\mathcal{L}}{y}\left[\frac{\frac{1}{w}\sinh w + \frac{1}{\rho}\sin\rho}{(\cosh w - \cos\rho)} - \frac{1}{y\mathcal{L}}\right]. \tag{3.42}$$

Inserting Eq. (3.42) into Eq. (3.18) gives

$$A_2 = \frac{2\sin\rho}{\rho}\int_0^\infty \frac{x\,dx}{\left(\cosh\sqrt{(x^2-\rho^2)} - \cos\rho\right)^2} \times \left[\frac{x^2}{4}\left(\frac{\sinh\sqrt{(x^2-\rho^2)}}{\sqrt{(x^2-\rho^2)}} + \frac{\sin\rho}{\rho}\right) - \cosh\sqrt{(x^2-\rho^2)} - \cos\rho\right], \tag{3.43}$$

which can be integrated numerically. The result is plotted in Fig. 3.1b along with a_2 obtained[7] from Ref. [28]. Although the two parameters differ from one another by as much as 20%, both indicate that the shapes are never very far from being isotropic, and both parameters vanish for $\rho = \rho_c$. Furthermore, since both parameters are intrinsically nonnegative, all fluctuations in A_2 and a_2 must also vanish for $\rho = \rho_c$. The $\rho = 0$ results for A_2 and a_2 are in very good agreement with those obtained earlier by analytic means [66, 69] and MC simulations [64].

As mentioned in Sec. 3.2, a disadvantage of analytic methods is that they do not readily yield the details of typical shapes. Nonetheless, these results do suggest a certain picture. The simplest interpretation is that for small ρ the system has the shape and size of an ordinary closed random walk. As ρ increases, the number of self intersections decreases until finally for $\rho = \rho_c$ there are no intersections, and the shape is circular. This is signalled by the vanishing of A_2 and its fluctuations. A somewhat different argument is used in Ref. [70] to obtain the same conclusion. In addition, Eqs. (3.30) and (3.31) show that, for $\rho = \rho_c$, the ensemble contains circles of all sizes.

Finally, it should be noted that all of the above results could equally have been studied using the conjugate, fixed-area ensemble. The partition functions for the two ensembles are

[6] Although the result in Eq. (3.42) is simple, the actual calculation involves doing approximately a thousand integrals of the form $\int_0^{x_1} dx \sinh ax \cos bx$. The individual integrals were generated and evaluated using the Maple symbolic mathematics program.

[7] There is an error in Eq. (15) of Ref. [28]. It should read (using the notation of this chapter) $a_2 = \left(1 - \rho^2/3 - \rho\cot\rho\right) / \left(-1 + 2\rho^2/3 + \rho^2\cot^2\rho\right)$.

related by

$$Z_{\text{area}}(V) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dg e^{igV} Z_{\text{pressure}}(y; p \to -ig). \tag{3.44}$$

The fixed-area ensemble is usually preferable to the fixed pressure one because experiments on real vesicles often control the volume or surface area instead of the pressure. The decision to use the fixed-pressure ensemble here was dictated by simplicity—the integral in Eq. (3.44) cannot usually be done analytically. It turns out that the two ensembles are not completely equivalent [28], but the essential physics is the same.

3.4 Problems with the PRW model

The connection between the PRW model and the more realistic LSF model is not as close as one might hope. There are two serious problems that must be overcome if the program alluded to in Sec. 3.1 of developing a field theory that complements the LSF model is to be realized. These are now discussed.

3.4.1 Effects of the algebraic area

Sec. 3.1.2 discussed the problems arising from replacing the true area with the algebraic area. These problems are now illustrated by a specific example. Consider a system modelled by Eq. (3.19) (with $\kappa = 0$) but with the y-dependent term now representing a real interaction in the system. Furthermore, assume that $y < 0$, so that the interaction between links is repulsive. Such an interaction is not very realistic, since it is quadratic in the link separation, but it is adequate for this illustration. A straightforward calculation of the average algebraic area gives

$$\mathcal{A}_y = \frac{\sin\rho - \frac{\rho}{w}\sin w}{\cos\rho - \cos w}. \tag{3.45}$$

The $\rho > 0$ regime of this model is not very interesting. The repulsive interaction simply generates an effective pressure that leads to the same instability discussed in Sec. 3.3.1 but with a smaller ρ_c, determined by the transcendental equation $\rho_c(y) = \cos^{-1}\left(\cos\sqrt{\rho_c^2(y) - 4y\mathcal{L}}\right)$.

One might imagine that the $\rho < 0$ case would be more interesting, since there is now a competition between the y-dependent interaction, which wants to inflate the ring, and the pressure term, which wants to inhibit such inflation, possibly leading to the branched-polymer configurations seen in Ref. [2]. However, inspection of Eq. (3.45) shows that $\mathcal{A}_y$ is anti-symmetric in ρ as before, so this does not happen. The reason for this is that the negative pressure causes the negative-area rings to further decrease (i. e., make more negative) their area in precisely the same way that a positive pressure inflates positive-area rings. Since these two situations are physically indistinguishable, $\mathcal{A}_y$ has the resulting anti-symmetry in ρ. Furthermore, this anti-symmetry will remain even in the presence of self avoidance. For Hamiltonians using H_p, $\rho < 0$ always implies $\mathcal{A} < 0$.

3.4.2 "Bending rigidity"

The meaning of H_κ is clear for the rigid-rod model, where Eq. (3.9) is enforced [48], but things are less clear for the PRW model. To begin with, since the parameter s is no longer the true arc-length, $d^2\mathbf{r}/ds^2$ will have, in addition to the component parallel to $\mathbf{r}$ (which measures the curvature), a tangential component, which has the effect of suppressing variations in the length of adjacent links. Although this feature is certainly undesirable, it does not have as large an impact on the physics contained in the model as other features.

Of much greater concern is the absence of a constraint fixing the perimeter length. This allows the partition function to be heavily weighted by configurations that would otherwise be absent, and it gives the distribution a qualitatively different character. To illustrate this, it is sufficient to calculate $\mathcal{A}_\kappa$ and $\mathcal{R}_\kappa^2$. These are obtained from $Z(y)$ (Eq. (3.20) with $\kappa \neq 0$) using the definitions in Eqs. (3.29) and (3.32), respectively. The results are

$$\mathcal{A}_\kappa(u) = \frac{2u\rho_c}{1+8\pi^2\frac{\kappa}{L^2}} \sum_{\mu=1}^{\frac{N-1}{2}} \left[N^2 f^2(\kappa)\tan^2 n_\mu - u^2\rho_c^2\right]^{-1} \tag{3.46}$$

and

$$\mathcal{R}_\kappa^2(u) = \frac{2}{1+8\pi^2\frac{\kappa}{\mathcal{L}^2}} \sum_{\mu=1}^{\frac{N-1}{2}} \frac{f(\kappa)\sec^2 n_\mu}{N^2 f^2(\kappa)\tan^2 n_\mu - u^2\rho_c^2}, \tag{3.47}$$

where

$$n_\mu = \frac{\pi\mu}{N}, \tag{3.48}$$

$$f(\kappa) = \frac{1+8\frac{N^2\kappa}{\mathcal{L}^2}\sin^2 n_\mu}{1+8\pi^2\frac{\kappa}{\mathcal{L}^2}}, \tag{3.49}$$

and u is defined through $\rho = u\rho_c$ $(0 \le u \le 1)$.

Explicit numerical calculation of the sums in Eqs. (3.46) and (3.47) for a wide range of the parameters N, ℓ, and ρ show that making the replacement $f(\kappa) \to 1$ always overestimates the sum. Thus, for practical purposes,

$$\mathcal{A}_\kappa(u) \lesssim \frac{\mathcal{A}(u)}{1+8\pi^2\frac{\kappa}{\mathcal{L}^2}}, \tag{3.50}$$

and

$$\mathcal{R}_\kappa^2(u) \lesssim \frac{\mathcal{R}^2(u)}{1+8\pi^2\frac{\kappa}{\mathcal{L}^2}}. \tag{3.51}$$

Increasing κ thus causes both $\mathcal{A}_\kappa$ and $\mathcal{R}_\kappa^2$ to decrease, which is *opposite* to the effect a "true" bending rigidity would have [48].

We can understand the source of this difference by once again resorting to mean-field arguments. First consider the rigid-rod model, with $p = 0$ for simplicity. The minimum-energy configuration for this system is a circle with energy

$$\begin{aligned} H_\kappa &= 8\pi^4\kappa\frac{R^2}{\mathcal{L}_0^3} \\ &= \frac{\pi\kappa}{R}. \end{aligned} \tag{3.52}$$

The last line follows from the constraint, which requires $\mathcal{L}_0 = 2\pi R$. However, if we now consider a circular configuration for the PRW model, we find $H_\kappa = \kappa 4\pi R^2/\mathcal{L}^3$. Although superficially similar to Eq. (3.52), it is quite different because there is now no constraint linking R to L. Circles are still minimum energy configurations, but R is free to vary, so the

ensemble will select the smallest value consistent with the other terms in the Hamiltonian. As κ increases, this becomes an increasingly dominant effect, and $\mathcal{A}_\kappa$ and $\mathcal{R}_\kappa^2$ decrease, as Eqs. (3.50) and (3.51) show. We are left, then, with the disappointing conclusion that, while H_κ can be included in the PRW model, its content is rather unphysical and it does not represent a "true" bending rigidity.

After ten o'clock at night those left round the tables are the genuine, desperate gamblers

Fyodor Dostoyevsky
The Gambler

Chapter 4

Self-Avoiding Tethered Membranes

This chapter discusses MC simulations of SATM's embedded in four and five dimensions that were done mainly in 1989. This chapter is written so as to preserve the historical order of development. Sections 4.1 and 4.2 give the historical background to these 1989 studies. The work itself is presented in Secs. 4.3 and 4.4 and includes subsequent simulations and analyses. Since this work was published [45] there have been several developments. Most notably, the first experiments on crystalline membranes have been performed [17] and other $d > 3$ SATM simulations have been done [1]. These developments will be covered in Sec. 4.5. The final section on outstanding problems also takes into account recent work.

4.1 Motivation

There are several reasons for the current widespread interest in tethered membranes.From a practical point of view, SATM's represent a natural generalization of polymers, which are of enormous technological value. If real membranes turn out to have a similar importance, it will be necessary to have a detailed understanding of their thermodynamics. A good

place to start is with a toy model such as the SATM. There is also interest in SATM's in biophysics, because the spectrin network sometimes attached to the otherwise fluid cell plasma membrane can be realistically modelled as a SATM. Indeed, in studying human red blood cells, inclusion of the spectrin network in the model is probably required for a good understanding of membrane thermal fluctuations [11, 71].

Further motivation for studying SATM's can be gained by comparing the state of knowledge about polymers (or SAW's) with that for SATM's. Polymers have been studied seriously for over fifty years [4, 6, 55] by a wide variety of theoretical techniques. These include ϵ-expansion RG analyses of the Edwards model [51], Monte Carlo simulations [72], exact enumeration of SAW's [73], real-space renormalization-group techniques [74], simple Flory theory [4], and, for $d = 2$, exact results [75]. The results for scaling exponents and shape characteristics obtained from these methods are all in good agreement with one another and with experimental values. The situation for SATM's is quite different and is discussed in detail in Sec. 4.2.2 below. To summarize that discussion here: Most of the techniques mentioned above have been applied to SATM's. They yield results indicating that SATM's are crumpled for $d > 2$, and they usually give $\nu(d = 3) \approx 0.8$. (An exception is the ϵ-expansion calculation based on the Edwards model, which gives $\nu(d = 3) \approx 0.536$.) In particular, early MC simulations found $\nu(d = 3) \approx 0.8$; however, more recent simulations on larger systems and using better data-analysis techniques find $\nu(d = 3) \approx 1.0$, suggesting that SATM's are flat. Although the results of these last simulations were initially controversial [30], there is now general agreement that they are correct [76].

Thus, theoretical approaches that work quite well for polymers fail to describe even the qualitative features of membranes. Consequently, SATM's present an interesting challenge to theoretical techniques that have been successful in handling polymers. Tethered membranes are also considered intrinsically interesting because they belong to new universality classes. In fact, based on our understanding of polymers, it is reasonable to expect that the

SATM is the "minimal model" [8] for the universality class of crystalline membranes.[1]

4.2 History

4.2.1 "Prehistory"

Before the introduction of the TM idea, workers in both the biophysics and particle physics communities were interested in random surfaces.

In the biophysics community this interest arose from the study of human red-blood-cell (RBC) shapes. Although the membrane of a real RBC is a quite complicated object, involving structures on several different length scales, many workers believe (justified by a certain amount of theoretical and experimental evidence) that, at least for the shape problem, this complex membrane structure can be replaced by a uniform lipid bilayer (fluid).

The Helfrich model [56, 77] embodies this idea by describing a RBC as a phantom fluid membrane whose energetics are completely determined by the local principle radii of curvature. Even with these drastic simplifications, the model can be solved exactly only at $T = 0$. The effects of thermal fluctuations about the $T = 0$ equilibrium shape can be described only at the Gaussian level. Nonetheless, many of the shapes resulting from minimizing the Helfrich Hamiltonian [56, 77, 78] accurately describe real RBC shapes.

Peliti and Leibler [79] took the first steps towards understanding finite-temperature effects on the Helfrich model by doing a one-loop renormalization-group analysis of thermal fluctuations. Their results show that the bending rigidity of a large planar fluid membrane decreases as short-wavelength degrees of freedom are integrated out. If this renormalization drives the bending rigidity to zero at large distances, then, in the thermodynamic limit,

[1] This expectation is reasonable only if bending rigidity turns out to be irrelevant for SATM's, as it is for polymers. The term "irrelevant" is used here in a sense different from how it is used for polymers. It is well established that the presence of bending rigidity affects the infrared fixed point of a TM (see Sec. 4.2.2). The issue is whether or not the SA interaction induces an effective bending rigidity. If it does, then the addition of an *explicit* bending-rigidity term to the SATM Hamiltonian is not likely to have any effect on the SATM infrared fixed point. The question is currently open (see Sec. 4.5.1).

there is no energy cost associated with bending a fluid membrane, so entropy drives it into a crumpled configuration, believed now to be equivalent to a branched polymer phase.[2] In addition to introducing the idea of a crumpled state, Ref. [79] is important because it discussed for the first time the possibility of a crumpling transition between a high-temperature crumpled phase and a low-temperature flat phase[3] of the system.

Within the elementary particle physics community, interest in random surfaces arose from attempts to study the long-wavelength behavior of quantum chromodynamics (QCD). There has been speculation for many years [14] that the bound states of QCD can be represented by the excitations of a one-dimensional object. The propagation of such a "string" in space-time generates a world-sheet (as opposed to a world-line for a particle), which can be regarded as a random surface. QCD transition-probability amplitudes can be calculated in numerical simulations on a four-dimensional space-time lattice [82] by averaging over typical world-sheets weighted by an appropriate action. Initial interest in this idea soon led to disappointment, when it was realized that the natural choice for the action, the surface area of the world-sheet, favored surfaces with divergent fluctuations [38]. The situation changed dramatically when Polyakov [83] discovered that adding an extrinsic curvature term to the action tamed the fluctuation problem in a gauge invariant way, and that the resulting action provided a unified description of a wide range of statistical mechanical and field theory problems. For example, the Helfrich model is a special case of the Polyakov model. Shortly after Polyakov's discovery, several papers were published containing both theoretical and numerical results, mostly dealing with the question of the fractal dimension of a random surface.

Parisi [84] made the first contribution along these lines. He constructed an heuristic argument, by analogy with random walks, which led to the result that $d_{fc} = 4$ for a random

[2]For a discussion of the relationship between fluid membranes with no rigidity and branched polymers see Refs. [59, 80, 81].

[3]The flat phase would be seen only if long range forces are present. The Helfrich model does not contain such forces.

surface. He then used an argument due to Mandelbrot [85] to deduce that the upper critical dimension d_{uc} should obey the formula $d_{uc} = 2d_{lc}$, thus finding $d_{uc} = 8$. A subsequent paper by Gross [86] showed that a careful analysis of the embeddings of a triangulated torus lead to the conclusion that $d_{uc} = \infty$. Subsequent work by several authors gave conflicting results that supported $d_{uc} = 8$ [87], $d_{uc} = \infty$ [38, 88] as well as other values [89, 90]. This confusion was finally sorted out by Cates [91]. He pointed out that Parisi, on the one hand, had studied surfaces with a fixed average number of plaquettes, but without regard for the connectivity of the surface. In fact, these surfaces turn out to have a fractal connectivity. Gross, on the other hand, had studied surfaces of fixed homogeneous connectivity. Hence, there is no reason to expect that these two systems should belong to the same universality class, and, indeed, they do not.

This is an especially important point here because it clearly distinguishes the universality class of random surfaces appropriate to the study of crystalline membranes (fixed homogeneous connectivity and $d_{uc} = \infty$) from those of interest in gauge theories, fluid membranes and branched polymers[4] (fractal connectivity and $d_{uc} = 8$). Although a great deal of work continues to be done on models in both of these universality classes, we will henceforth focus on the fixed-connectivity class, since it has provided the inspiration for the study of SATM's.

4.2.2 Recent history

The above-described studies of random surfaces led Kantor, Kardar, and Nelson [31] to introduce the tethered membrane (see Fig. 1.1b) as a model of a crystalline random surface. In this ground-breaking paper, the authors performed exact calculations for a Gaussian model of a phantom tethered membrane (PTM[5]) and studied other phantom membrane

[4]One should be cautious about including gauge theories in the same universality class as fluid membranes and branched polymers because of the issue of self-avoidance, which arises naturally in the study of the latter two. Self-avoidance has been used as a technique for studying the $N \to 0$ limit of SU(N) gauge theories, but its usefulness for this purpose is now in doubt [89].

[5]A PTM is a model with the same connectivity as in Fig. 1.1b, but without self avoidance.

models using Migdal-Kadanoff RG techniques. They also developed a Flory theory and did the first MC simulations for SATM's. The most important results from this work were: (1) For the purposes of designing a Hamiltonian, the tethering constraint between two vertices could be reliably replaced by a "Gaussian spring" potential (see Sec. 2.1.2); (2) for a PTM, $R_g \sim \sqrt{\ln L}$ (exact), independent of d; (3) $d = 3$ simulations indicated that a SATM crumples with $\nu \approx 0.8$, which agrees with the Flory theory prediction, $\nu_F = \frac{4}{5}$. Subsequent analysis by Bouchard and Bouchard [74] using real-space RG techniques also gave the result $\nu \approx 0.8$.

A short time later, Kardar and Nelson [92], Aronovitz and Lubensky [93], and Duplantier [94] all calculated the scaling exponents ν and γ as functions of D and d for the generalized Edwards model (see Sec. 2.4.2) using a RG ϵ-expansion. The main results from these calculations were that the value of γ depends on the shape of the SATM boundary when D is not an integer and that, for $D = 2$ and $d = 3$, $\nu = 0.536 + O(\epsilon^2)$. Although the calculation itself is technically impressive, the result for ν is not very encouraging since it violates the $\nu \geq \frac{2}{3}$ requirement set by the close-packing argument in Sec. 2.2.3. The accuracy of the calculation for γ is still unknown.

The next important development was a stability analysis by Nelson and Peliti [95] of the flat phase of a PTM model that contains an in-plane elastic energy (which gives an energy cost to local membrane stretching and a bending energy term similar to Eq. (3.3)). Both of these energy terms were studied in a Gaussian approximation. They showed that in-plane oscillations of an elastic membrane cause the bare bending rigidity to diverge at long wavelengths under renormalization. The resulting increased energy cost in folding the membrane leads to a stable flat phase.

The realization that crystalline membranes with a non-zero bare bending rigidity can be flat, combined with the results on crumpled membranes [31], led Kantor, Kardar and Nelson [32] to predict a crumpling transition much like the one discussed in Ref. [79] for fluid membranes. This prediction was tested by performing MC simulations on a PTM model

that included a bending energy. This model will subsequently be referred to as the PTM$_\kappa$ model. The results of these simulations indicated that there was, in fact, a second-order phase transition between a flat state and a crumpled state. Ambjorn *et al.* [37] also did MC simulations on the PTM$_\kappa$ model and discovered a second-order crumpling transition. Self-avoidance was not included in any of these simulations because SATM simulations are much more CPU-intensive than comparably-sized PTM ones, and because earlier simulations on SATM's [31] showed that they crumple in the absence of any bending rigidity. Thus, the transition from a low-temperature flat phase to a high-temperature crumpled phase was still expected for a SATM with bending rigidity, albeit with critical exponents different from those of a PTM$_\kappa$.

David and Guitter [36] added non-linear terms to the model studied in Ref. [95] and performed a RG analysis using a $1/d$ saddle-point expansion. They found a non-trivial UV-stable fixed point, which they were able to interpret as a signal for a second-order crumpling transition. Aronovitz and Lubensky [96] studied a D-dimensional generalization of this model. They found that, for $D < D_{uc} = 4$, conventional (harmonic) elastic theory [97] breaks down, and all elastic constants are renormalized either to zero or infinity in the infrared. They also calculated the scaling exponents for the elastic constants using an ϵ-expansion, where $\epsilon = 4 - D$.

Paczuski, Kardar and Nelson [98, 99] subsequently described a Landau theory for the crumpling transition using the d-component tangent vectors of the surface $\partial\mathbf{r}/\partial s_i$ as a set of order parameters. When combined with the requirement of rotational invariance, the resulting Hamiltonian is very similar to an n-component ϕ^4 model. At low temperatures the Hamiltonian studied in Ref. [95] is recovered, while at high temperatures it reduces to H_0 of Eq. (2.6). They did a mean-field analysis of the model that included SA in the Flory approximation, and found, as expected, that it does not destroy the phase transition but does change the scaling exponents. They also included critical fluctuations in the the context of an ϵ-expansion and found that the transition becomes weakly first order for

$d < d_{uc} \approx 219$. Aronovitz, Golubović, and Lubensky [100] further analyzed the Landau theory, using a $1/d$-expansion,[6] and were able to recover results of Ref. [36]. They also calculated the lower critical dimension, below which the crumpling transition disappears, and found $D_{lc} = 2 - \frac{2}{d} + O(1/d^2)$.

As a result of these investigations, a consistent picture of SATM's embedded in $\mathbb{R}^3$ began to emerge: Thermal fluctuations create an anomalous increase in the bending rigidity, leading to a low-temperature flat phase. Increasing the temperature was expected to result in a second-order (or very weakly first-order) phase transition to a crumpled state with $\nu \approx 0.8$. It was, therefore, somewhat surprising when Boal and Plischke reported [33] that they found no indication of either the crumpled phase or the phase transition in MC studies on SATM's embedded in $\mathbb{R}^3$. In fact, they found $\nu \approx 1.0$, independent of temperature. This last point was soon confirmed by Abraham, Plischke and Rudge [34], who did molecular dynamics simulations on a similar model, but with the tethers replaced by an attractive potential. Ho and Baumgartner [101] also did MC simulation studies on a SA triangular plaquette model[7] and the SATM model, both embedded in $\mathbb{R}^3$, and found $\nu \approx 1.0$. This discrepancy between the simulations of SATM's and the previously mentioned theoretical work has yet to be resolved satisfactorily (see Sec. 4.5).

4.3 Monte Carlo studies of TM's in $d = 4, 5$

4.3.1 Motivation for higher dimensional studies

Because there has been so much controversy about the persistence of the flat phase at high temperature and because it is such a surprise, it has been considered important to obtain a more detailed understanding of this issue. One possible area of investigation is the role

[6] This $1/d$-expansion exploits the analogy with the n-component ϕ^4 model and is different from the one used in Ref. [36].

[7] In this model self avoidance is enforced by forbidding each triangular plaquette from intersecting any other plaquette (see Sec. 4.5.1).

of the embedding dimension d in determining the phase of a SATM. Recall that a SAW changes from a Flory-crumpled state to a Gaussian-crumpled state when $d \nearrow d_{uc} = 4$ [4]. Indeed, the study of polymers in $\mathbb{R}^4$ is a useful starting point for understanding the scaling behavior of polymers in $\mathbb{R}^3$. It seems natural to ask whether something similar occurs for SATM's. Of course, this analogy cannot be taken too literally, since we know from the discussion in Sec. 2.4.2 that $d_{uc} = \infty$ for SATM's; nonetheless significant effects due to a change in d are possible.

More generally, it is desirable to have a more complete understanding of the (D, d) phase diagram shown in Fig. 2.4 and, particularly, to locate its phase boundaries. The non-local nature of the SA interaction makes analytical attempts to do this difficult and unreliable, but many points in the phase diagram can be investigated using simulations. Although the SA interaction makes simulations more time consuming, it creates no difficulty of principle. The remainder of this section describes MC simulations done for two such points in the (D, d) phase diagram — SATM's embedded in $d = 4$ and $d = 5$. Aside from simply measuring the scaling exponents $\nu_i^{(L)}$, I will attempt to address two more-general questions: Does the SATM model have a rough phase?; and, what effect does increasing d have on the SATM phase?

4.3.2 Simulation details

Although all the simulations described in this chapter were done on SATM's, many of the simulation details are equally applicable to PTM's. The parts of the following discussion that refer to a TM are meant to apply to both.

A TM is "constructed" on a computer by creating a labelled list of d-dimensional vectors (one for each vertex), which represent the locations of the vertices in $\mathbb{R}^d$. The initial locations are chosen so that the vertices form a flat, nearly stretched configuration.[8] A

[8] If the simulation is started from a completely stretched configuration, it is almost impossible for the vertices to move since practically every trial move will violate the tethering constraint.

second list contains the intrinsic nearest neighbors for each vertex, and is used for enforcing the tethering constraint Eq. (2.1). In all the TM simulations discussed in this thesis, the tethers form a triangular lattice with an hexagonal boundary, like the one shown in Fig. 1.1. For a TM of linear size L, there are $N = (3L^2 + 1)/4$ vertices.

The are are two problems that arise from starting the simulation in a specially chosen flat configuration. The first is that such a configuration is far from equilibrium, so data taken during the initial relaxation process will not give information about the equilibrium properties of the system. Once the initial relaxation time is known, this problem can be handled by simply ignoring the data from that part of the simulation. The second problem is that typical equilibrium configurations might not be easily accessible to the SATM on the time scales available in the simulation. In particular, crumpled configurations may not be readily accessible from an initial flat phase. This problem has been addressed by studying the effects of crumpled initial configurations on the long-time behavior of the simulation [45]. If configurations characteristic of the true equilibrium state are not equally accessible from both types of initial configurations, we might expect to find measurable differences in the "equilibrium" states obtained from them. The crumpled configurations are formed by adding an attractive central potential to the simulation. This produces very compact configurations with a value of ν near the close-packing limit. The potential is then turned off and the SATM is allowed to relax to an "equilibrium" state. Measurements on such systems show that they are thermodynamically indistinguishable from configurations obtained from the flat phase. Of course, it is still possible that the true equilibrium state is not readily accessible from either type of initial configuration; but, the above procedure increases our confidence that the configurations obtained in the simulation are representative of the true equilibrium.

Temporal evolution of the simulations proceeds by a standard Metropolis algorithm [102]. With the central vertex held fixed to prevent a global translation of the TM, one of the remaining vertices is chosen at random and displaced by a random d-dimensional vector.

If the trial move does not violate the tethering constraint, Eq. (2.1), or, for SATM's, the additional constraint Eq. (2.2), the move is accepted. $N-1$ such trial moves constitute a MC time step of the network, or, more simply, a MC step. A natural time scale (measured in MC steps) is the Rouse time [32],

$$\tau_R = \frac{2dN}{\pi^2\rho^2}, \tag{4.1}$$

where ρ is the maximum length of the random-move vector. The Rouse time is an estimate of how long it takes a point in the membrane to drift a macroscopic ($\gtrsim L$) distance. In practice, ρ is chosen to have a value that causes one-half of the trial moves to be accepted. Since there is no intrinsic energy scale in the model, acceptance of a trial move is determined solely by the constraints, and in this sense the simulations were done at infinite temperature. Data, consisting of the eigenvalues $\lambda_i^{(L)}$ of the inertia tensor $Q_{ij}^{(L)}$, are taken every $\tau_R/10$ MC steps.

After a run of approximately $1000\tau_R$, the simulation is interrupted to determine the initial relaxation time, which is typically about $100\tau_R$. The data from this part of the simulation are discarded. The simulation is then restarted and allowed to run until sufficient data has been acquired so that the statistical error in R_g is less than 1% at the 2σ level. This ensures that the statistical error for the $\nu_i^{(L)}$'s is about 5%[9].

An important consideration in SA simulations is using an efficient method to test for violations of the SA constraint. The naive approach of simply calculating the distance between all pairs of vertices is undesirable, since the resulting algorithm is $O(N^2)$ in CPU time. Breaking up the embedding space into d-dimensional cubes (typically the same size as the hard-sphere radius of a vertex) and keeping track of which vertices are in each cell means that the SA test need only be applied to vertices in adjacent cells. The resulting algorithm is $O(N)$. This increase in computational speed is achieved at the expense of increased CPU memory requirements, which are now $O(L^d)$. In practice, $d \geq 4$ simulations

[9]This criterion was used for the original simulations. Other, more stringent criteria, which lead to longer simulations, will be discussed in Sec. 4.4.2

require so much memory that the cell algorithm is impractical. This problem is solved by changing the cells to "hyper-tubes" that are square in two directions and infinite in the other $d-2$ directions.[10] The algorithm is most efficient when the maximum number of vertices in any hyper-tube is minimized. This is achieved by orienting the hyper-tubes so that their infinite-length directions are parallel to the eigenvectors belonging to the $d-2$ smallest $\lambda_i^{(L)}$'s. Because the SATM rotates in the embedding space with time, it must be periodically realigned with the hyper-tubes. Realigning the SATM about once per Rouse time keeps the vertex occupancy of the hyper-tubes down to a reasonable number.

4.4 Analysis and results

4.4.1 Data analysis

The raw eigenvalue data are used to determine three things: (1) the eigenvalue autocorrelation function, (2) the shape parameter S_d, and (3) the scaling exponents $\nu_i^{(L)}$. The asphericity could also be calculated from the data; but it is of little interest, since we are mainly concerned with whether or not SATM's crumple.

The autocorrelation function. The eigenvalue autocorrelation function,

$$C_i^{(L)}(t) = \frac{\left\langle \left[\lambda_i^{(L)}(t+t') - \left\langle \lambda_i^{(L)}(t+t') \right\rangle_{t'}\right] \left[\lambda_i^{(L)}(t') - \left\langle \lambda_i^{(L)}(t') \right\rangle\right] \right\rangle_{t'}}{\left\langle \left[\lambda_i^{(L)}(t') - \left\langle \lambda_i^{(L)}(t') \right\rangle_{t'}\right]^2 \right\rangle_{t'}}, \tag{4.2}$$

provides a measure of the statistical independence of configurations obtained by the time evolution of the simulation. The angle brackets $\langle \cdots \rangle_{t'}$ denote an average over the MC time t' (measured in units of the Rouse time), i. e.,

$$\langle \mathcal{O}(t'+t)\mathcal{P}(t') \rangle_{t'} = \frac{1}{N} \sum_{t'=1}^{N-t} \mathcal{O}(t'+t)\mathcal{P}(t'). \tag{4.3}$$

Provided that $N \gg T$, where T is the longest correlation time for the system, Eq. (4.3) effectively becomes an ensemble average.

[10] This idea is originally due to M. Plischke.

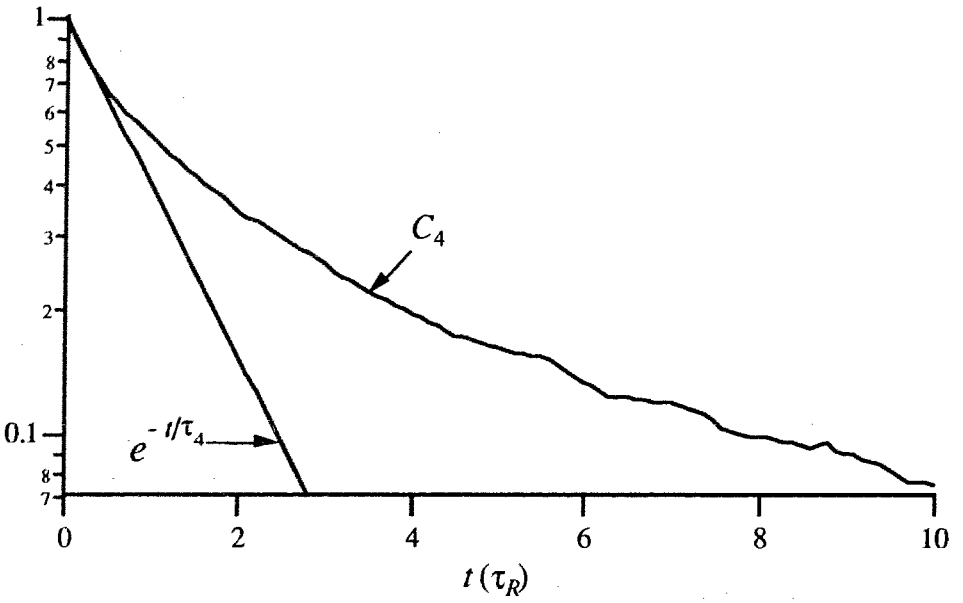


Figure 4.1: The autocorrelation function for λ_4 obtained from a $d = 4$, $L = 17$ simulation of a SATM. The total length of the run is $1800\tau_R$. The solid line is a plot of Eq. (4.4) with $\tau_4^{(17)} \approx 0.42$. Note that the solid line increasingly underestimates $C_4^{(17)}(t)$ as t increases, as will any choice of $\tau_4^{(17)}$ at sufficiently long times.

The information in Eq. (4.2) is important for estimating the size of the statistical errors in the eigenvalue data and for determining the length of the initial relaxation time. Although $C_i^{(L)}(t)$ is characterized by several time scales, it is nonetheless useful to give a one-parameter characterization, which is determined by assuming that, for short times,

$$C_i^{(L)}(t) = \exp\left(-t/\tau_i^{(L)}\right). \tag{4.4}$$

The eigenvalue relaxation time (also measured in Rouse times) is, then,

$$\frac{1}{\tau_i^{(L)}} \equiv \lim_{t\to 0} \frac{-1}{t} \ln C_i^{(L)}(t). \tag{4.5}$$

Figure 4.1 shows a plot of a typical $C_i^{(L)}(t)$ obtained from one of the simulations along with the estimated value of τ_i. Although the fit is reasonably good for short times ($t \lesssim 1$), the form

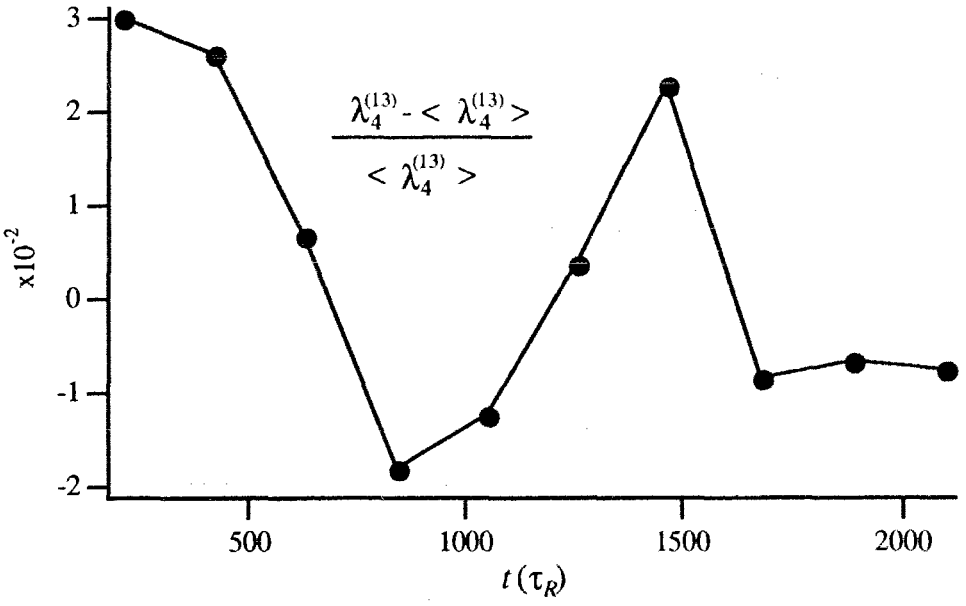


Figure 4.2: Evidence for the presence of long-wavelength oscillations in the simulations. The curve shows $\frac{\lambda_4^{(13)}-\left\langle\lambda_4^{(13)}\right\rangle_{t'}}{\left\langle\lambda_4^{(13)}\right\rangle_{t'}}$ versus t' for $d = 5$. Although the oscillations have an amplitude of only a few percent, they can create systematic errors in the estimates of the $\nu_i^{(L)}$'s that are as large as 15% (see Sec. 4.4.2).

in Eq. (4.4) clearly underestimates the correlation at longer times. Choosing a larger value for τ_i does not really solve the problem, because $d \ln C_i(t)/dt \to 0$ as $t \to \infty$, so any fit to the data of the form in Eq. (4.4) will underestimate the correlation for sufficiently large values of t. The reason that $\tau_i^{(L)}$ becomes an increasingly poor characterization of $C_i^{(L)}(t)$ with increasing t is because Eq. (4.5) is only sensitive to short time-scale correlations, in contrast to $C_i^{(L)}(t)$, which is better described by a spectrum of relaxation times. Furthermore, this spectrum contains some very long-period relaxation times. For example, the oscillations in Fig. 4.2 induce additional correlations that have very long relaxation times and are not accounted for by Eq. (4.5). The long relaxation times are also L-dependent; in general $\tau_{\mathbf{max}}^{(L)} \to \infty$ as $L \to \infty$.

The shape parameter. It is straightforward to apply the definition given in Eq. (2.27) to calculate $S_d^{(L)}$ from the raw eigenvalue data. If SATM's do not crumple, then we expect $S_d \sim L^{-2\nu_S} \to 0$ as $L \to \infty$ ($\nu_S \equiv (\nu_d - \nu_1)$). Therefore, plotting $\ln\left(S_d^{(L)}\right)$ versus $\ln\left(\frac{1}{L}\right)$ and obtaining a straight line with positive slope is evidence for no crumpling.

The scaling exponent. There are several ways to determine ν_j. One method, inspired by similar techniques used in polymer studies [103], is to calculate the structure function [33],

$$\mathcal{S}_j^{(L)}(k) \equiv \frac{1}{N^2}\left\langle \sum_{\alpha,\beta} \exp\left[i\mathbf{k}_j \cdot \left(\mathbf{r}^{(\alpha)}(t') - \mathbf{r}^{(\beta)}(t')\right)\right]\right\rangle_{t'}, \tag{4.6}$$

where

$$\mathbf{k}_j \equiv k\hat{\mathbf{e}}\{\mathbf{r}\}_j, \tag{4.7}$$

and $\hat{\mathbf{e}}\{\mathbf{r}\}_j$ is the $\mathbb{R}^d$ eigenvector corresponding to the eigenvalue $\lambda_j^{(L)}$. The $\hat{\mathbf{e}}_j$'s, therefore, define a coordinate system attached to membrane. Note that, in addition to the time average denoted by $\langle\cdots\rangle_{t'}$, Eq. (4.6) also contains an average over the membrane coordinates, which is taken with respect to a frame of reference fixed to the membrane rather than the "laboratory" frame. Rescaling the coordinates, as in Eq. (2.39), naturally leads to a scaling form,

$$\mathcal{S}_j^{(L)}(k) \sim \phi_j(x_j), \tag{4.8}$$

where

$$x_j \equiv k\ell L^{\nu_j}. \tag{4.9}$$

For the proper choice of ν_j, plots of $\mathcal{S}^{(L)}$ versus x_j for different values of L will collapse to a single curve. Determining ν_j in this way has the advantage that it allows one to test whether or not $\mathcal{S}^{(L)}(k\hat{\mathbf{e}}_1) \sim x_1^{2/\nu_1}$ when x_1 is sufficiently large. Abraham and Nelson [76] have argued that simulated SATM's that pass this test are large enough for scaling laws to be valid. A disadvantage of this method is that it entails a certain amount of subjectivity in deciding what value of ν_j gives the best fit to the data. This subjectivity also makes it difficult to assign meaningful confidence levels to the exponent error estimates.

Another method for determining ν_j is to plot $\lambda_j^{(L)}$ versus L, and then to do a least squares fit to the data assuming the scaling form $\lambda_j^{(L)} \sim L^{2\nu_j}$ [34]. This method does not rely on any subjective judgments to determine ν_j, so a meaningful error analysis can be done, but it suffers from the disadvantage that it treats all data points, regardless of their L value, as equally important. This is clearly undesirable, since we expect the large-L data to reflect the long-distance scaling behavior better than the small-L data. Of course, it is possible to weight the data points as a function of L; but, since there is currently no theoretical understanding of finite-size corrections to scaling for SATM's, there is no way to choose between various weighting schemes.

The last method for calculating ν_j is to use the eigenvalue data from two different-sized simulations to calculate a series of L-dependent scaling exponents,

$$\nu_j^{(L)} = \frac{1}{2} \frac{\ln\left(\lambda_j^{(L)}/\lambda_j^{(L')}\right)}{\ln\left(L/L'\right)}, \tag{4.10}$$

where L' refers to the SATM next smaller in size than the L-sized one. The $\nu_j^{(L)}$'s are then plotted versus $1/L$, and ν_j estimated by extrapolating the data to $1/L = 0$. Calculating ν_j in this way makes the importance of any finite-size effects more evident than in the previous methods; however, it does produce a larger error estimate for ν_j. Another advantage of this method is that it provides an internal consistency check for the statistical error estimates on the $\nu_j^{(L)}$. For those SATM's with L values large enough to be in the scaling regime, we expect the derived $\nu_j^{(L)}$'s to be ordered properly, i. e., $\lambda_j^{(L)} < \lambda_k^{(L)} \Rightarrow \nu_j^{(L)} \leq \nu_k^{(L)}$ (within experimental error). Thus, for sufficiently large L,

$$\nu_j^{(L)} - \delta\nu_j^{(L)} \leq \nu_k^{(L)} + \delta\nu_k^{(L)}, \forall\, j < k, \tag{4.11}$$

where $\delta\nu_j^{(L)}$ is an estimate of the statistical error of $\nu_j^{(L)}$. As we shall see in Sec. 4.4.2, Eq. (4.11) turns out to be the most stringent test of the data.

4.4.2 Error analysis

There are two important sources of error in the determination of the $\lambda_i^{(L)}$'s: statistical errors and systematic errors. In addition, the determination of the ν_i's and S_d from the $\lambda_i^{(L)}$'s generate finite-size effects.

Statistical errors. Determining the size of the statistical errors for an uncorrelated data set is straightforward [104]. However, as Fig. 4.1 illustrates, the data obtained from successive configurations are highly correlated, and these correlations are not well-described by a characteristic relaxation time. Several methods are used for handling this problem. They are (in order of increasing sophistication):

1. The data obtained from each configuration is regarded as independent, and a straightforward [104] error analysis done. This treatment provides a lower bound on the error estimate.

2. The data are binned, with the bin size given by $\tau_{\rm bin} = -\ln(.01)\tau_i^{(L)}$, where $\tau_i^{(L)}$ is defined by Eq. (4.5). Assuming that $C_i^{(L)}$ has the form given in Eq. (4.4), this gives $C_i^{(L)}(\tau_{\rm bin}) = 0.01$. Numerical experiments [105] suggest that, if the correlation between successive bins is less than 0.01, then the correlations can be safely ignored for the purposes of error estimates.

3. This method is the same as the one above, except that $\tau_{\rm bin}$ is determined by measuring the $C_i^{(L)}$ directly from the simulation (i. e., without assuming that the correlations are of the form in Eq. (4.4)) and requiring that $C_i^{(L)}(\tau_{\rm bin}) \leq 0.01$.

4. A direct calculation of the error for a series of correlated measurements is done [105]. This method also ignores correlations less than 0.01 .

The data are analyzed using all four methods. In practice, methods 2–4 give similar results, but in all cases the largest value obtained is used as the error estimate. All statistical errors are calculated at the 2σ level.

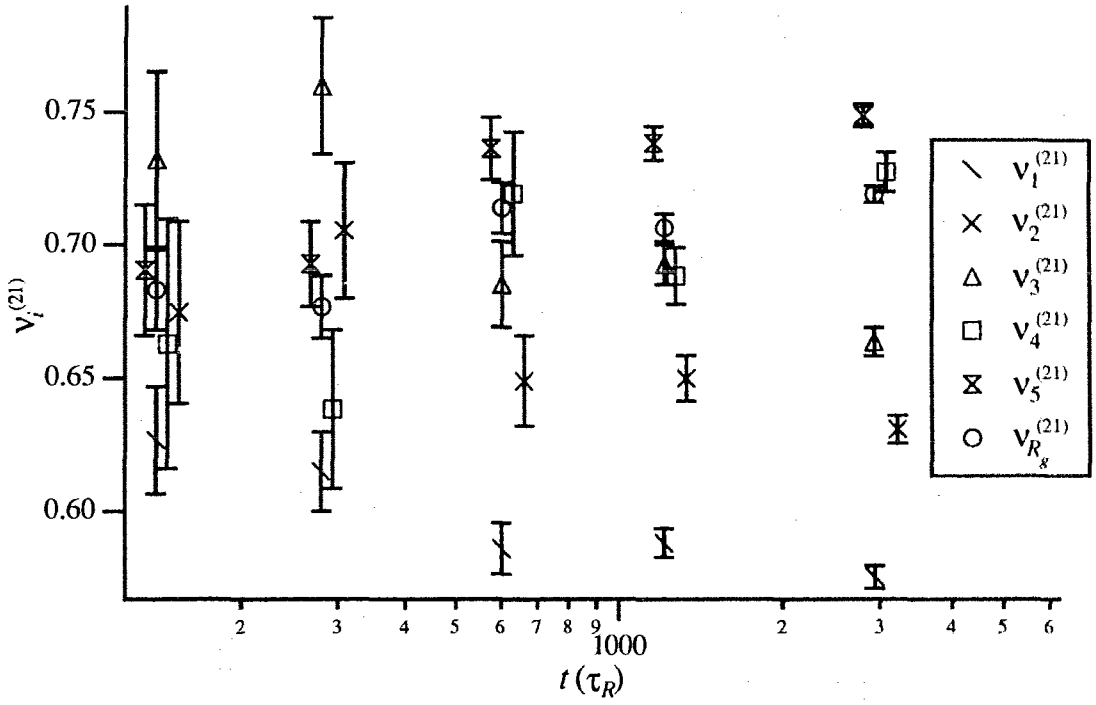


Figure 4.3: An example of how systematic errors caused by long-wavelength oscillations in the $\lambda_i^{(L)}$'s affect the ordering of the $\nu_i^{(L)}$'s. For short simulation times, the systematic errors are relatively important and the $\nu_i^{(21)}$'s do not satisfy Eq. (4.11). As the simulation time increases, the systematic errors become relatively unimportant and the $\nu_i^{(21)}$'s do satisfy Eq. (4.11). The error bars represent 1σ statistical errors. The data points have been spread out slightly along the t-axis for visual clarity.

Systematic errors. In addition to the statistical errors, there are systematic errors that arise from the long-wavelength oscillations in $\lambda_i^{(L)}$'s. If the simulation time is only a fraction of the longest oscillation period $T^{(L)}$, then the average will be biased by this effect. The existence of such long-wavelength oscillations is evident from Fig. 4.2, where $T^{(L)} \approx 1000\tau_R$. Although the eigenvalue oscillation amplitudes are typically only a few percent, they can lead to systematic errors in the estimates of the $\nu_i^{(L)}$'s (calculated using Eq. (4.10)) that are as large as 15%. For $t < T^{(L)}$, the systematic error typically dominates the statistical error. This can be seen in Fig. 4.3, where the time evolution of $\nu_i^{(21)}$ is plotted. The late-time

$\nu_i^{(21)}$'s are frequently outside the range predicted by the statistical error estimates of the early-time values, thus indicating the presence of significant systematic errors. The $t \lesssim T^{(L)}$ region also does not satisfy Eq. (4.11).

Fortunately, the systematic error is proportional to $1/T^{(L)}$ (as opposed to the statistical error, which is proportional to $1/\sqrt{T^{(L)}}$), so at sufficiently large times it becomes small relative to the statistical error. There is no way to know *a priori* how large $T^{(L)}$ must be in order to insure that the systematic errors are negligible. However, it seems reasonable to assume that the systematic errors are responsible for the failure of the $\nu_i^{(L)}$'s to satisfy Eq. (4.11). This certainly appears to be true for the data in Fig. 4.3, where $\nu_i^{(21)}$ does satisfy Eq. (4.11) at sufficiently large times. In all subsequent discussions the systematic error will be assumed to be unimportant whenever the $\nu_i^{(L)}$'s satisfy Eq. (4.11).

Finite-size effects. Once the errors for the $\lambda_i^{(L)}$'s have been determined, it is straightforward to calculate the errors for $\nu_i^{(L)}$. However, the determination of the ν_i's from the $\nu_i^{(L)}$'s introduces another source of error. Because the simulations were done for finite (and rather small) L-values, the $\nu_i^{(L)}$'s must be extrapolated to the $L \to \infty$ limit to obtain the thermodynamically meaningful ν_i's. The errors generated by this extrapolation are difficult to estimate without a finite-size scaling theory. For the ν_i's they are estimated by simply extrapolating the statistical errors from the two largest L-value measurements to the $L \to \infty$ limit. This method probably overestimates the finite-size errors, but it will not underestimate them provided that the statistical and systematic errors are not so large as to mask the $L \to \infty$ trend in the data.

4.4.3 Results

The results of the simulations are summarized in Tables 4.1 and 4.2 and Figs. 4.4 – 4.9.

$d = 4$ results. From Fig. 4.4 it appears that: (1) the two smaller eigenvalues have similar scaling exponents; (2) the two larger eigenvalues have similar scaling exponents; (3) the scaling exponent for $\lambda_1^{(L)}$ and $\lambda_2^{(L)}$ is different from the scaling exponent for $\lambda_3^{(L)}$ and $\lambda_4^{(L)}$.

	$d = 4$		$d = 5$	
	(a)	(b)	(a)	(b)
ν_S	–	0.20± 0.02	–	0.18±0.02
ν_{R_g}	0.85± 0.04	0.80± 0.01	0.76± 0.06	*0.71±0.01
ν_1	0.66± 0.06	0.62± 0.01	0.66±0.09	0.58±0.02
ν_2	0.77± 0.08	*0.69±0.01	0.70±0.09	*0.63± 0.03
ν_3	0.92± 0.08	0.84± 0.01	0.75±0.11	*0.67± 0.04
ν_4	0.88± 0.05	0.82± 0.01	0.71±0.09	0.72± 0.01
ν_5	–	–	0.79±0.08	*0.73± 0.01

Table 4.1: Scaling-exponent results for simulations of SATM's embedded in $d = 4, 5$. The exponents were determined in two different ways: Method (a) calculates the $\nu_i^{(L)}$'s using Eq. (4.10) and then extrapolates this information to the $L \to \infty$ limit. The uncertainties include both statistical errors and finite-size effects. Method (b) determines the ν_i's from a least-squares fit to the points plotted in Figs. 4.4 and 4.7. For the table entries marked with an asterisk the $L = 5$ data has been omitted from the least-squares fit. The uncertainties contain only the statistical errors. The ν_S's are obtained from the slopes of the curves in Fig. 4.6.

d	L	MC steps ($\times 10^6$)	τ_R	$\tau_{\max}^{(L)}$	$\tau_{\text{bin}}/\ln(100)$	FLOP ($\times 10^9$)
4	5	1.39	1262	0.13	0.26	4.62
	9	5.39	4896	0.27	0.75	66.0
	13	108	10788	0.53	1.38	3116
	17	83.8	19313	0.85	3.04	5464
	21	104	30603	1.13	4.28	19091
5	5	2.80	1332	0.12	0.29	18.9
	9	10.6	5043	0.18	0.41	239
	13	23.8	11349	0.33	0.82	1348
	17	62.2	20063	0.50	1.23	1879
	21	97.7	31531	0.71	1.81	15643
	31	107	71384	1.21	2.78	54369

Table 4.2: Data for MC simulations of SATM's embedded in $d = 4, 5$. τ_R is measured in units of MC steps. $\tau_{\max}^{(L)}$ and τ_{bin} are in units of τ_R. $\tau_{\max}^{(L)}$ is the largest eigenvalue relaxation time obtained from the size-L simulation using Eq. (4.5). τ_{bin} is determined from error analysis method 3 in Sec. 4.4.2. τ_{bin} is scaled by the factor $1/\ln(100)$ in order to provide a comparison with τ_R and $\tau_{\max}^{(L)}$. FLOP is the total number of floating-point operations carried out for each simulation.

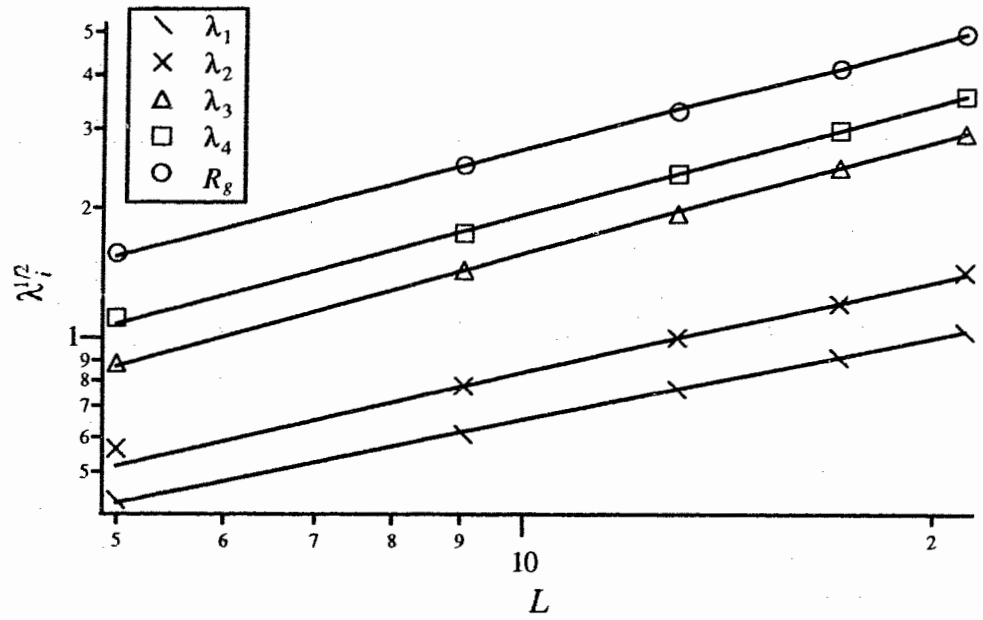


Figure 4.4: The $d = 4$ eigenvalues versus L. The straight lines are obtained from a least-squares fit to the data points. Not all of the fits include the $L = 5$ data (see Table 4.1). The error bars (2σ) have been omitted for clarity. In all cases they are smaller than the size of the symbol.

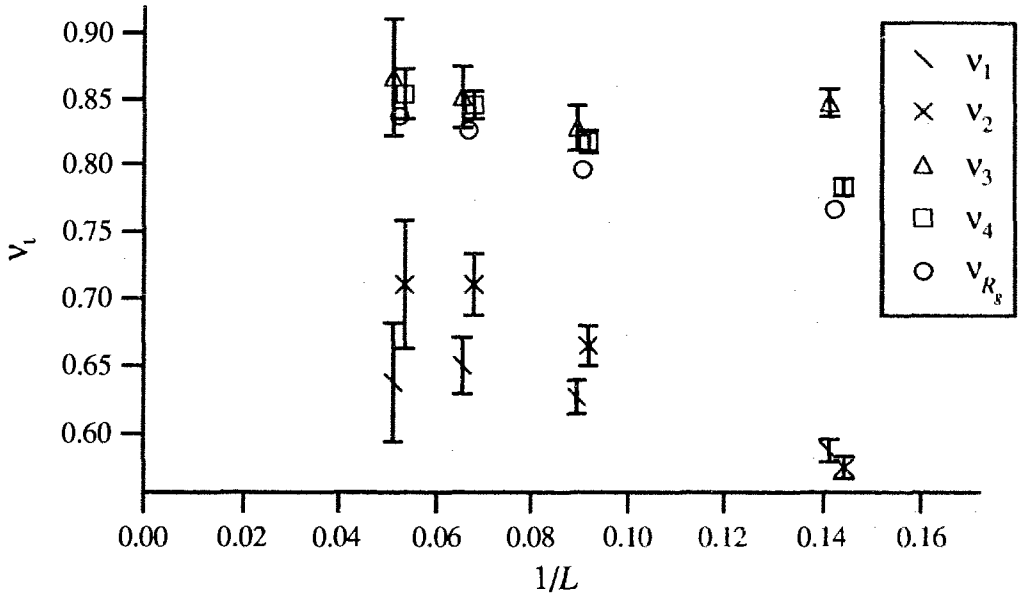


Figure 4.5: The $d = 4$ scaling exponents versus $1/L$. The dashed lines illustrate the procedure for extrapolating the data to the $L \to \infty$ limit. The large separation between ν_1 and ν_4 and the separation between ν_4 and unity suggest that SATM's embedded in $\mathbb{R}^4$ are rough.

Figure 4.4 thus suggests that the SATM is in the rough phase. On the other hand, Fig. 4.5 gives a somewhat different picture. The most obvious feature of this graph is the strong L-dependence of $\nu_2^{(L)}$. If the trend established at small L-values were to continue, it might suggest that the SATM has a phase with $\nu_4 = \nu_3 = \nu_2 \neq \nu_1$. Although the evidence is not conclusive that $\nu_3 = \nu_4$, this statement is certainly consistent with the data.

In order to appreciate better the meaning of Fig. 4.5, it is worthwhile to consider what a graph of the $\nu_i^{(L)}$'s should look like for each of the phases. If the SATM were flat in the sense of Sec. 2.2.3, then the graph would show, for sufficiently large L, the $\nu_i^{(L)}$'s in two clusters. The first cluster, consisting of $\nu_{R_g}^{(L)}$, $\nu_d^{(L)}$, and $\nu_{d-1}^{(L)}$, would have a value of unity. The second cluster, consisting of the remaining $\nu_i^{(L)}$'s, would have a smaller value. Furthermore, we would expect that within each cluster the $\nu_i^{(L)}$'s would be properly ordered as in Eq. (4.11). If the SATM were rough, the $\nu_i^{(L)}$'s would be ordered and clustered in the same way as the

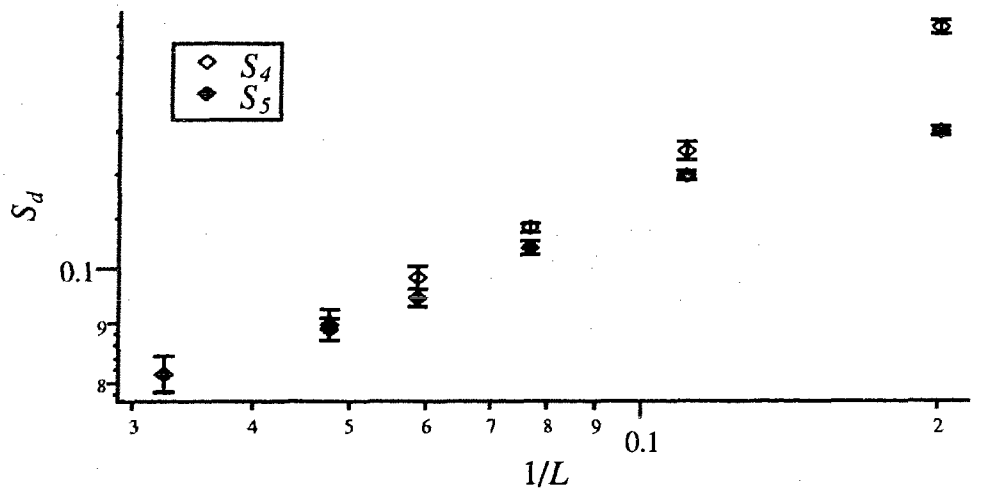


Figure 4.6: The shape parameters S_4 and S_5 versus $1/L$.

flat phase, except that $\nu_{R_g}^{(L)}$, $\nu_d^{(L)}$, and $\nu_{d-1}^{(L)}$ would have a value less than unity. If the SATM were crumpled, all the $\nu_i^{(L)}$'s and $\nu_{R_g}^{(L)}$ would form a single cluster with $\nu < 1$.

Figure 4.5 thus indicates that SATM's embedded in $\mathbb{R}^4$ are in some kind of rough phase, i. e., they are not flat or crumpled. The possibility that the SATM is crumpled is ruled out by the separation between ν_1 and ν_4, which is at the 5σ level. Further evidence of at least two distinct scaling exponents is given in Fig. 4.6, where apparently $S_4^{(L)} \to 0$ as $L \to \infty$. The data in this figure also give a value for ν_S consistent with $2\,(\nu_d - \nu_1)$.

Figure 4.5 also shows an example of the extrapolation procedure used to obtain the ν_i's and $\delta\nu_i$'s shown in Table 4.1. Note that the estimates for the ν_i's have been chosen to be midway between their upper and lower error estimates, as can be seen in the example. This procedure has not been used on ν_2 because it exhibits such strong finite-size effects. All that can be said is that $\nu_1 \leq \nu_2 \leq \nu_4$. The extrapolations for ν_1 and ν_3 has been done using the two intermediate-L data points, instead of the two large-L data points, because the combination of a large statistical errors and significant sytematic errors make $\nu_1^{(21)}$ and $\nu_3^{(21)}$ too unreliable for the extrapolation procedure.

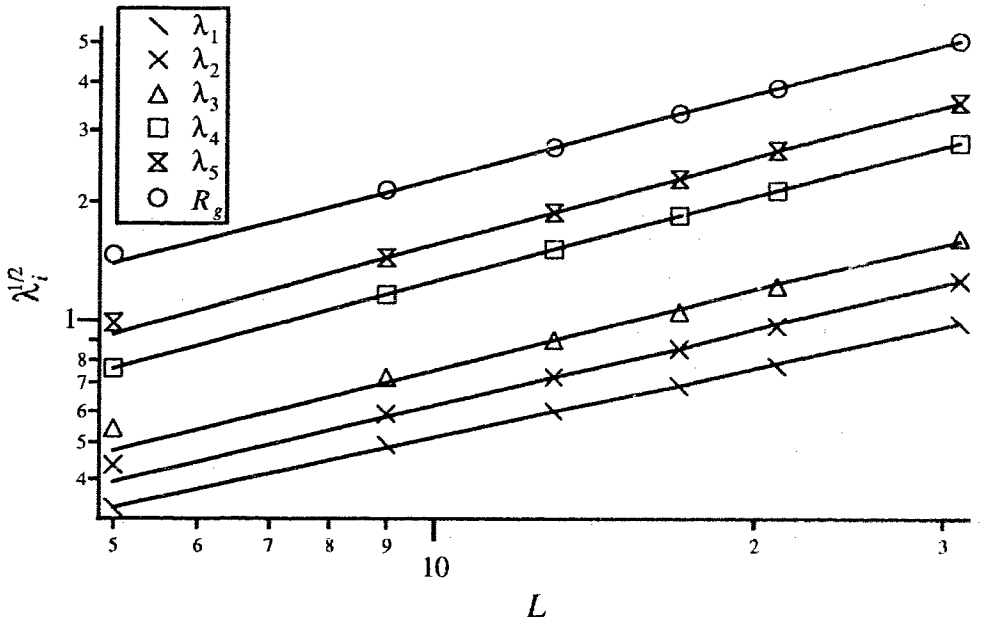


Figure 4.7: The $d = 5$ eigenvalues versus L. The straight lines are obtained from a least-squares fit to the data points. Not all of the fits include the $L = 5$ data (see Table 4.1).

$d = 5$ results. Figure 4.7 appears to suggest that SATM's embedded in $\mathbb{R}^5$ are also rough, but less so than those embedded in $\mathbb{R}^4$. However, Fig. 4.8 shows that the $d = 5$ results are even more ambiguous than the $d = 4$ results. At intermediate values of $1/L$ the exponents are apparently separated into two distinct clusters, suggesting the presence of the rough phase. However, the data for smaller values of $1/L$ indicate that the clustering is greatly reduced there. It is impossible to say, on the basis of this data, whether this trend continues for even smaller values of $1/L$ and eventually leads to the crumpled phase, or if the trend simply leads to two (or perhaps three) clusters that are somewhat closer together but still separate. In either event, it is clear that the data are not yet in the scaling regime. Figure 4.6 is also inconclusive, since it indicates that S_5 is decreasing with decreasing $1/L$, but the $L = 31$ data point suggests that S_5 may be approaching a constant value of approximately 0.08, which would be a signal for crumpling.

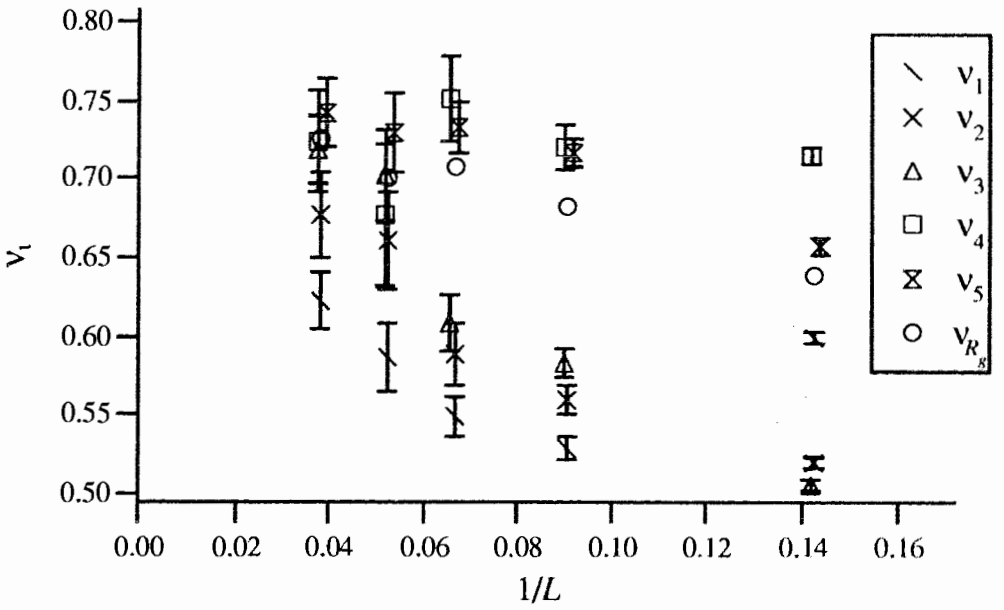


Figure 4.8: The $d = 5$ scaling exponents versus $1/L$. The data points are plotted at $1/L = 2/(L_1 + L_2)$ (L_1 and L_2 are the L-values of two adjacent-sized simulations). The clustering of the exponents at each value of L and the fact that they are correctly ordered (to within experimental value) is taken as further evidence that the statistical errors have been reliably estimated and that there are no significant systematic errors. Note the good separation of the exponent clusters at small L and the tendency of this gap to decrease as L increases. Because of this trend, it is impossible to tell from the data whether the thermodynamic phase of this system is rough or crumpled.

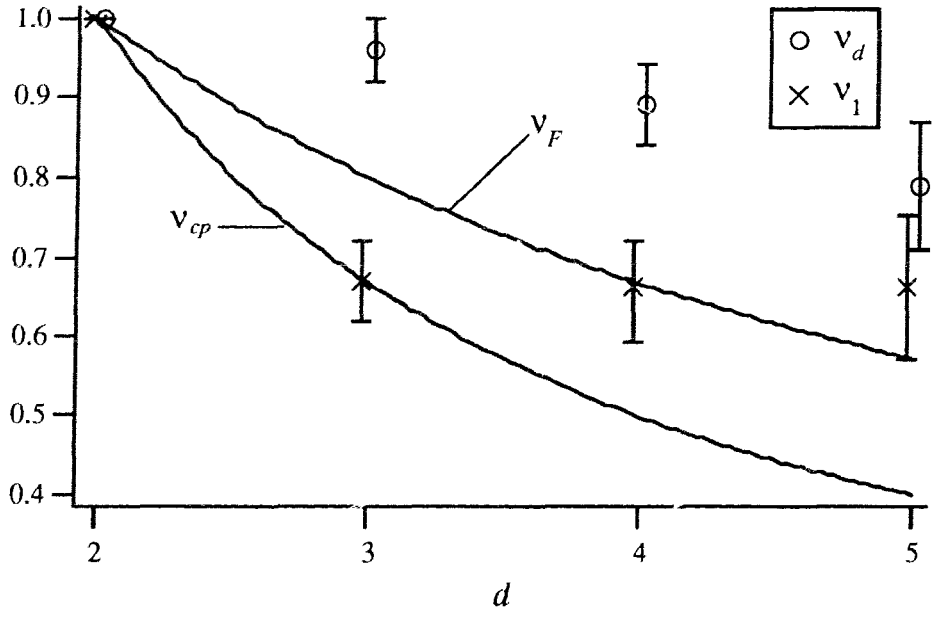


Figure 4.9: The largest (ν_d) and smallest (ν_1) scaling exponents as a function of embedding dimension. The $d = 3$ data is taken from Ref. [45]. The large uncertainty for the $d = 5$ ν_1 value is caused by finite size effects. The error bars for the $d = 4$ and $d = 5$ data points are calculated using Method (a) in Table 4.1. The $d = 3$ error bars reflect only statistical errors. The data points have been slightly offset for clarity. The Flory value (ν_F) and the close-packing value (ν_{cp}) for ν are also shown for comparison.

This conclusion is somewhat different from the one published earlier in Ref. [45] (1989), which concluded that SATM's embedded in $\mathbb{R}^5$ were rough. The results discussed here are based on substantially longer simulation runs (typically two to three times longer than those in Ref. [45]) as well as larger L values. Taken together, Ref. [45] and the work described here offer a cautionary tale on the wisdom of doing simulations in the absence of experimental or theoretical guides. The prohibitive CPU requirements needed for reliably simulating significantly larger L-values make it unlikely that the identity of the $d = 5$ SATM phase will be resolved in the near future.

The exponent results for both simulations are summarized in Fig. 4.9.

4.5 Recent work

There have been a series of recent developments in the field that are directly relevant to the issues and results discussed in Sec. 4.4.3: (1) recent experiments [17, 18] and MC simulations [35, 106] have found $\nu(d = 3) \approx 0.8$; (2) recent molecular dynamics simulations on SATM's embedded in $d > 3$ [1] find no rough phase for $d = 4$ and a crumpled phase for $d \geq 5$. Both of these developments are discussed below.

4.5.1 Recent results for $d = 3$

Although the following discussion concerns $d = 3$ SATM's, it is relevant to the work in this chapter because it calls into question the appropriateness of modelling real crystalline membranes using the SATM model studied here.

Because the existence of the high temperature flat phase is such a surprise, there has been a certain amount of speculation as to its cause. Most of these conjectures focus on the idea that the SA interaction somehow induces an effective bending rigidity, even if there is none explicitly present in the model. Abraham and Nelson [76] have recently offered a simple picture of how this might occur: The hard spheres that are used to enforce the SA constraint have the additional effect of restricting the range of allowable angles defined by the normals of two adjacent triangles. As a result of this restriction, this angle has a non-zero expectation, which depends only on the ratio of the ball size and the tether length. A similar effect can be achieved by adding a bending energy term similar to Eq. (3.3) in a Hamiltonian *without* self avoidance. The conclusion drawn from this argument is that self avoidance induces an effective "entropic bending rigidity".

Abraham and Nelson also support this idea with a simulation study of a TM with first- and second-nearest-neighbor SA interactions only. The absence of long-range SA causes the TM to crumple; but, as the tether lengths are shortened, the TM is observed to undergo an apparent "phase transition" to a flat state. If the entropic bending rigidity idea is

correct, then shortening the tethers has the effect of increasing the bending rigidity. Thus, this simulation is supposed to demonstrate that TM's without long-range SA can still be made flat by increasing the entropic bending rigidity arising from second-nearest-neighbor interactions between hard spheres.

The calculation and simulation data given do not provide a very convincing argument for the finite ball size as the source of the bending rigidity. For example, this explanation does not seem to account for the flat phase observed in simulations [34, 45] employing very small hard spheres (as small as one-tenth the size used in the simulations here). Another objection is that identical simulations could be used to "prove" that second-nearest-neighbor hard core interactions are sufficient to produce "flat" polymers, which is clearly wrong. It seems quite likely that all that is accomplished by reducing the tether length is to increase the effective plaquette size (much as reducing the tether length in a polymer increases the effective segment size), and that the crumpled behavior would be recovered on a much larger length scale.

Although the evidence provided in support of this idea is not convincing, the idea itself has attracted some interest. Baumgartner [35] has suggested a model of SATM's that does not induce an effective entropic bending rigidity. This model, known as the self-avoiding plaquette model (SAPM), consists of a network of tethered vertices similar to the PTM described in Sec 4.2.2. The model differs from the SATM in the method used to impose the SA constraint. Rather than using hard spheres, the triangular plaquettes (defined by the triplets of mutual nearest neighbors) are forbidden from intersecting each other. Since the plaquettes in the SAPM have zero thickness, the range of angles for the plaquette normals is not restricted. If the Abraham and Nelson idea is correct, then the SAPM will have zero effective bending rigidity and will crumple.

Baumgartner and collaborators have done three studies of the SAPM embedded in $\mathbb{R}^3$. The first study [101] (published before Ref. [76]) concluded that SAPM's are flat. More recent studies [35, 106] (published after Ref. [76]) have found that SAPM's crumple with

$\nu \approx 0.8$. Although there are some minor differences in the methods of data analysis and the system sizes studied, these differences do not provide a satisfactory explanation for the discrepancies observed [107]. Clearly more work is needed to obtain definitive answers about the properties of the SAPM.

The issue of whether or not SAPM's crumple has implications beyond testing the entropic bending-rigidity hypothesis. If the later SAPM simulations turn out to be correct, then the SATM belongs in different universality class than the SAPM. This naturally leads to a question about which model provides a better description of real crystalline surfaces. This is not an easy question to answer *a priori*. However, if experiments turn out to favor the SAPM, then much of the motivation for studying SATM's (especially for $d > 3$) will be lost.

The first laboratory experiments on crystalline membranes [17, 18] have been performed recently, as well. The membranes are made of graphite oxide (layers of graphite bonded together by oxygen atoms) with a typical size of a few (3 – 8) micrometers and a thickness $50\text{Å} \leq d \leq 100\text{Å}$. The interatomic spacing of the carbon atoms is $a_0 \approx 2.5\text{Å}$; thus $1.2 \times 10^4 \leq L \leq 3.2 \times 10^4$. After the membranes are synthesized they are suspended in a solvent. The resulting solution contains approximately 10^{-3} membranes/$(\mu\text{m})^3$. Information about the membrane structure factor is obtained by performing light scattering experiments on the solution.

If these membranes are crumpled, then their structure factor will be governed by the theory of scattering for fractal objects, which predicts [17]

$$\mathcal{S}(k, R_g) = k^{-d_f} F(kR_g), \tag{4.12}$$

where $d_f = 2/\nu$ is the Hausdorff or fractal dimension of the membrane. In the long-wavelength limit ($kR_g \ll 1$) the membranes act like point scatterers, and we should expect $\mathcal{S}$ to be k-independent. Thus, $\lim_{k\to\infty} F(kR_g) = (kR_g)^{d_f}$. In the short-wavelength limit, only the internal structure of the membrane is probed, so there should be no dependence of $\mathcal{S}$ on R_g. Thus, $\lim_{k\to 0} F(kR_g) = \textit{constant}$ and $\mathcal{S} \sim k^{-d_f}$.

Abraham and Nelson [108] have calculated the directionally averaged structure factor for a flat membrane, which is necessary for comparing the theory for flat membranes with scattering experiments. They find[11]

$$\mathcal{S}(k,L) \sim k^{-2} \quad \text{for} \quad 2\pi < kR_g < \frac{2\pi R_g}{c_1 a_0 L^{\nu_1}}$$
$$\text{and} \tag{4.13}$$
$$\mathcal{S}(k,L) \sim k^{-2/\nu_1} \quad \text{for} \quad \frac{2\pi R_g}{c_1 a_0 L^{\nu_1}} < kR_g < \frac{2\pi R_g}{d}.$$

$c_1 a_0 L^{\nu_1}$ is the membrane size in the direction perpendicular to the plane defined by the membrane itself. Note that the structure-factor scaling exponents 2 and $2/\nu_1$ are the Hausdorff dimensions for a flat membrane. After inserting the experimental parameters and the values $\nu_1 = 0.65$ and $c_1 = 0.25$ obtained from simulations, Eq. (4.13) becomes

$$\mathcal{S}(k,L) \sim k^{-2} \quad \text{for} \quad 6.28 < kR_g < 648$$
$$\text{and} \tag{4.14}$$
$$\mathcal{S}(k,L) \sim k^{-2/\nu_1} \quad \text{for} \quad 648 < kR_g < 2990.$$

The quantity of interest in these experiments is d_f, because it can distinguish between the flat state and the crumpled state. Thus, using wavelengths small enough to probe only the internal structure of a membrane will lead to $\mathcal{S} \sim k^{-d_f}$. d_f can then be readily extracted from a log-log plot of $\mathcal{S}$ versus kR_g. For membranes with a minimum size of $3\mu m$, the internal structure will be probed for $k \gg 2(\mu m)^{-1}$. In the actual experiments

$$2.1(\mu m)^{-1} < k < 31.42(\mu m)^{-1}$$
$$\text{or} \tag{4.15}$$
$$6.3 < kR_g < 94.$$

The upper bound on k is fixed by the wavelength (514 nm) of the light source (an argon cw laser). Comparison of Eqs. (4.14) and (4.15) shows that, if the membranes are flat, then the experiments are well within the $\mathcal{S} \sim k^{-2}$ regime. Thus $d_f = 2$ is evidence for a flat phase, and $d_f > 2$ is evidence for a crumpled phase.

[11] The value given in Ref. [108] for the crossover from $\mathcal{S} \sim k^{-2}$ behavior to $\mathcal{S}(k,L) \sim k^{-2/\nu_1}$ behavior is incorrect. Equation (4.13) contains the corrected value.

Two experiments have been done so far: the first one [17] finds $d_f = 2.4 \pm 0.1$; the second one [18] finds $d_f = 2.54 \pm 0.05$. In addition, a compact phase with $d_f = 3$ (created by reducing the solvent quality) was observed. The authors argue that the observation of a compact phase rules out the possibility that the $d_f = 2.5$ phase is actually a metastable state on the way to forming a compact phase. The experiments thus seem to indicate that membranes are crumpled with ν very close to the Flory value of 0.8 .

As the authors point out, these experiments cannot yet be regarded as definitive. Other explanations of the results, such as crumpling induced by internal membrane defects, need to be explored. A more convincing demonstration that graphite oxide crumples would be to observe the flat-to-crumpled transition. This should be relatively easy to do, since it requires only that the effective rigidity be increased, which can be accomplished either by cooling the system or by increasing the membrane thickness.

4.5.2 Recent results for $d > 3$

In a recent article [1] Grest has studied relatively large SATM's ($13 \leq L \leq 57$) embedded in $d = 4, 5, 6, 8$ using molecular dynamics simulations. His results are rather different from the ones described in Sec. 4.4.3: Grest finds that SATM's embedded in $d = 4$ are flat and that SATM's embedded in $d \geq 5$ are crumpled with ν somewhat larger than the corresponding ν_F predicted by Eq. (2.53). The numerical results of Ref. [1] are summarized in Table 4.3 and should be compared with the results in Table 4.1.

Grest provides the following argument to explain these results: Abraham and Nelson's [76] explanation for the entropic origin of the bending rigidity is assumed to be correct. Grest, then argues that because this explanation invokes the interactions between next-nearest neighbors, it is natural to assume that the bending rigidity is caused by three-body interactions. Eq. (2.51) shows that if $D = 2$, then all n-body interactions are relevant near the Gaussian fixed point. However, SATM phases are clearly controlled by a non-Gaussian fixed point, and we expect that near this other fixed point $\nu > \nu_0 (= 0$ when $D = 2)$. A

d	ζ	$\nu_{\|\|}$	state
3	0.64±0.04	1.0	flat
4	0.77±0.04	1.0	flat
5	0.77±0.03	0.82±0.05	crumpled
6	0.69±0.03	0.69±0.05	crumpled
8	0.60±0.03	0.60±0.03	crumpled

Table 4.3: The scaling exponent results for the SATM simulations done in Ref. [1]. ζ is defined to be the scaling exponent for the $d-2$ smallest eigenvalues. $\nu_{||}$ is defined to be the scaling exponent for the two largest eigenvalues.

naive scaling analysis shows that (cf. Eq. (2.51))

$$d_{ucn} = \left(\frac{n}{n-1}\right)\frac{D}{\nu}. \tag{4.16}$$

Hence $d_{uc3} < d_{uc2}$ whenever ν is positive. Thus, when $d_{uc2} > d > d_{uc3}$, the irrelevancy of the three-body interaction will, in turn, make the bending rigidity irrelevant, and the SATM will crumple.

This argument has a certain appeal, especially since the results in Table 4.3 support it, but it is open to criticism:

1. As has already been discussed in Sec. 4.5.1 the explanation for the origin of the bending rigidity given in Ref. [76] is not very convincing.

2. Even if one accepts the idea that bending rigidity is induced by next-nearest neighbor interactions, these interactions still involve repulsion between only two (albeit distant) vertices. The n-body interaction in Eq. (2.50) contributes only when n vertices have the *same* position vector. Next-nearest neighbors apparently interact via two-body interactions — not three-body interactions.

3. Even if one accepts that three-body interactions are somehow responsible for inducing a bending rigidity, they are still unimportant in any model that contains two-body interactions. Recall from the discussion in Sec. 2.4.2 that the effect of the interaction terms is to enforce the self-avoidance constraint, and that in the thermodynamic limit

$H_{\rm SA}^{(n)}$ reproduces this constraint exactly. Clearly, any configuration forbidden by $H_{\rm SA}^{(3)}$ will be forbidden by $H_{\rm SA}^{(2)}$ as well. The converse statement is, of course, not true.

Aside from the issue of whether or not Grest's explanation of his results is correct, there remains the more basic question of why his results disagree with the ones described in Sec. 4.4.3. It is easy to dismiss two explanations: the difference in simulation techniques and differences due to simulation size.

Although the two sets of simulations used different simulation techniques (MC versus molecular dynamics), it is not likely that this is the source of the difference. On general grounds, we expect the SATM model to belong to a particular universality class, with critical exponents that are independent of the details of the model. This has certainly been the case for simulations of SATM's embedded in $\mathbb{R}^3$, where the results for the critical exponents appear to be independent of the simulation technique. It seems unlikely that a change of d would affect this result. Of course, universality will be applicable only for simulations that are large enough to be in the scaling regime, and it is possible that one simulation technique reaches the scaling regime much more readily (i. e., for smaller L values) than the other [109]. Once again, there is no indication that this is true for $d = 3$ simulations, so it is hard to understand why the situation should be different for $d > 3$.

The most obvious difference between the two sets of simulations is in the L-values used. A potentially simple explanation is that my simulations suffer from significant finite-size effects, and that, if they were extended to much larger L-values, the results for the scaling exponents would agree with those obtained in Ref. [1]. There are two arguments against this idea. The first argument is that although the $\nu_i^{(L)}$'s for the smaller $\lambda_i^{(L)}$'s generally exhibit strong finite-size effects in my simulations, $\nu_d^{(L)}$ and $\nu_{d-1}^{(L)}$ exhibit a rather weak dependence on L, suggesting that the scaling regime for $\lambda_{d-1}^{(L)}$ and $\lambda_d^{(L)}$ begins at $L \approx 13$. The second argument against the finite-size-effect explanation is that there is some overlap in L-values for the two sets of simulations, and even here the two simulation sets give very different scaling exponents.

Finally, I discuss a potential source of the discrepancy. Figure 4.10 shows the $\nu_i^{(L)}$'s obtained from Ref. [1][12]. It is apparent that the Ref. [1] exponents are not well-ordered, and this suggests the presence of significant systematic errors in the data (see Sec. 4.4.2). There is no way to know *a priori* whether such systematic errors (if they are present) are responsible for the difference in the conclusions reached in the two studies, but reducing them by increasing the run-time of the Ref. [1] simulations would seem to be a reasonable first step towards resolving the discrepancy.

[12]I wish to thank G. Grest for graciously providing me with a copy of his raw data for analysis.

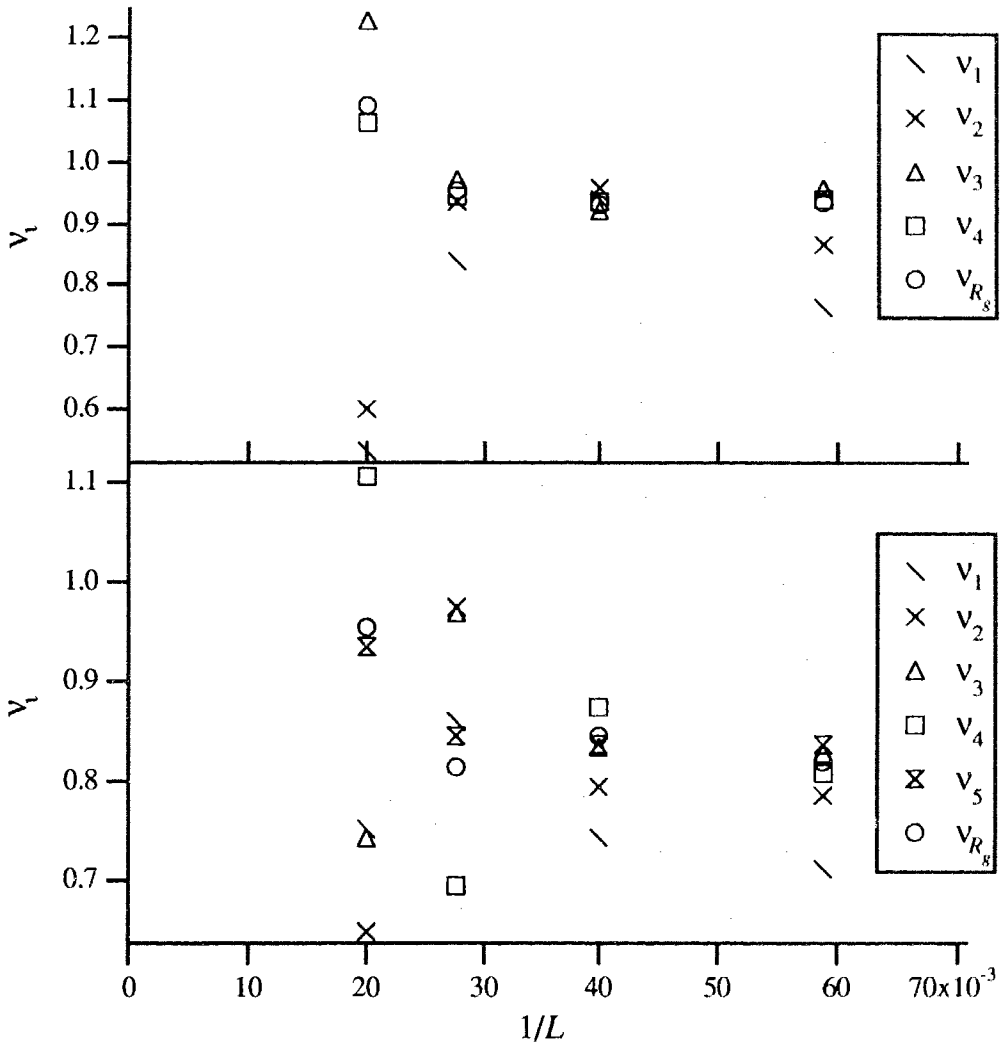


Figure 4.10: The scaling exponents obtained from the SATM simulations described in Ref. [1]. The upper figure is the $d = 4$ data; the lower figure is the $d = 5$ data. Note that the scaling exponents are not well-ordered and that some of the ν_i's are larger than 1, which violates Eq. (2.25). Together, these two observations suggest the presence of significant systematic errors. Such errors may explain the discrepancy between the scaling exponents obtained from the two sets of simulations.

He got out two thin butter-spades from the shelf and put them down into the little chest and pulled out something that seemed to me remarkably like another chest. I went over to it and gave it a close examination with my hand, feeling the same identical wrinkles, the same proportions and the same completely perfect brasswork on a smaller scale. It was so faultless and delightful that it reminded me forcibly, strange and foolish as it may seem, of something I did not understand and had never heard of.

Flann O'Brien
The Third Policeman

Chapter 5

Sierpiński Gaskets

5.1 Introduction

In Chap. 4 it was pointed out that many theoretical approaches that are useful for studying polymers, such as Flory theory and an ϵ-expansion analysis of the Edwards model, do not correctly describe even the qualitative features of SATM's. In the context of field theory, both polymers and membranes are seen as special cases of a more general (but still homogeneous) manifold characterized by its intrinsic connectivity or topological dimension D. It is then apparent that, as D increases, somewhere in the regime $1 < D \leq 2$, three things must occur:

1. Renormalization group arguments based on the ϵ-expansion about the Gaussian fixed point break down. Alternatively, the Edwards model may cease to be a good description of SATN's;

2. The Flory theory must cease to contain all of the essential physics needed to describe the thermodynamics properly;

3. The thermodynamic behavior of the manifold must change from crumpled to flat.

It seems worthwhile to study the connections between these three points and the reasons why the theory breaks down in the first place. Ideally, one would like to compare the results of simulations with theoretical calculations for arbitrary D and to look for discrepancies. Unfortunately, it is not clear how to simulate such manifolds. To overcome this difficulty, a model is needed that is amenable to simulation and theoretical calculations and interpolates between polymers and membranes. In this chapter I propose regularly connected fractal networks for this purpose. Although not homogeneous, these networks are characterized by a *spectral dimension*[1] D_s, which is completely determined by the intrinsic connectivity of the network, and, in many ways, plays the role of D (see Sec. 5.2). As a result, they are amenable to both Flory theory calculations and ϵ-expansions. They are also straightforward to simulate. The remainder of this chapter is concerned with the comparison between Flory theory calculations, ϵ-expansions and simulations for a particular model: a network with the connectivity of a $b = 2$ Sierpiński gasket [115]. As we shall see, the SA version does, in fact, crumple; the Flory theory is reasonably accurate (within a few percent of the measured value); and the ϵ-expansion gives modest results (within 20% of the measured value) for $d = 3$, but improves as d approaches its upper critical dimension.

5.2 Theory

Figure 5.1 shows the intrinsic connectivity of the model, which is that of a $b = 2$ Sierpiński gasket. For the nth iteration of the network, the number of vertices is

$$N^{(n)} = \frac{3^{n+1} + 3}{2}. \tag{5.1}$$

[1]The notion of the spectral dimension (sometimes referred to as the fracton dimension) of a network originally arose from the study of the density of states for a generalized Laplacian operator on a fractal. See Refs. [110] and [111] for this point of view. More recently, D_s has been used to characterize random walks on fractal lattices, $\langle [\mathbf{r}(s) - \mathbf{r}(0)]^2 \rangle \sim s^{D_s/d_{fl}}$. This is the definition of D_s used here. d_{fl} is the fractal dimension of the lattice; i. e., $N \sim L^{d_{fl}}$, where N is the number of vertices on the lattice. See Refs. [112, 113, 114, 115] for this latter point of view.

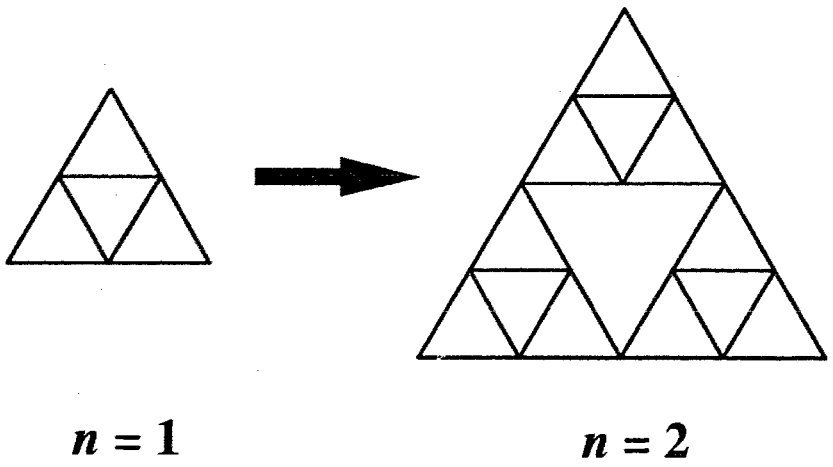


Figure 5.1: Iteration process for a $b = 2$ Sierpiński gasket. The value of b gives the multiplication factor for the increase in the number of connections along an edge after each iteration.

The lattice fractal and spectral dimensions, defined in the $n \to \infty$ limit, are, respectively,

$$d_{fl} = \frac{\ln 3}{\ln 2} \approx 1.585 \tag{5.2}$$

and

$$D_s = 2\frac{\ln 3}{\ln 5} \approx 1.365\,. \tag{5.3}$$

Both phantom Sierpiński gaskets (PSG) and self-avoiding Sierpiński gaskets (SASG) are discussed in this chapter.

5.2.1 Theoretical predictions

There exists for the PSG a rigorous prediction for the scaling exponent ν_0 based on a mapping between the partition functions for tethered networks and resistor networks with the same connectivity [19]. This mapping predicts (subscript zero denotes phantom-network quantities)

$$\nu_0 = \left(\frac{2 - D_s}{2}\right)\frac{d_{fl}}{D_s} \tag{5.4}$$

so

$$d_{f0} = \frac{d_{fl}}{\nu_0} = \frac{2D_s}{2 - D_s}. \tag{5.5}$$

d_{f0} is the fractal dimension, which relates the radius of gyration to the number of vertices,

$$R_{g0} \sim N^{1/d_{f0}}. \tag{5.6}$$

For SA-networks, a Flory theory can be constructed using a line of argument identical to the one in Sec 2.4.2 for homogeneous TN's with the result

$$\nu_F = \left(\frac{D_s + 2}{d + 2}\right)\frac{d_{fl}}{D_s}. \tag{5.7}$$

The upper critical dimension, d_{uc} (above which self-avoidance is irrelevant) can also be determined using the methods of Sec 2.4.2. Recall that d_{f0} measures how much of the embedding space is filled by the fractal object. For embedding spaces with $d > 2d_{f0}$, intersections between two or more portions of the network separated by large intrinsic distances will be unlikely; hence, $d_{uc} = 2d_{f0}$. An alternative, but related, argument is that including self-avoidance always increases R_g, so ν_0 is a lower bound on ν_F. Hence, $\nu_F|_{d=d_{uc}} = \nu_0$. Either way, one finds

$$d_{uc} = \frac{4D_s}{2 - D_s}. \tag{5.8}$$

Substitution of Eq. (5.3) into Eq. (5.8) gives $d_{uc} \approx 8.6$ for the $b = 2$ SASG. This is in marked contrast to the case of homogeneous membranes [20](i. e., $D = 2$ networks) for which $d_{uc} = \infty$. Also, notice that Eqs. (5.4), (5.5), (5.7) and (5.8) reduce to the corresponding homogeneous manifold results by simply substituting $D_s = d_{fl} = D$. In this sense D_s plays the role of an effective topological dimension.

5.2.2 The Edwards model

As mentioned in Sec. 4.2.2 there have been several studies of homogeneous manifolds that have applied ϵ-expansion techniques (see Ref. [20] for a recent review) to the generalized

version of the Edwards model [50] already introduced in Eqs. (2.44), (2.37), and (2.45),

$$\beta H_{\rm E} = \frac{1}{2}\int_0^{\mathcal{L}} d^D s \sum_{i=1}^{D}\left(\frac{\partial \mathbf{r}}{\partial s_i}\right)^2 + \frac{v_2}{2}\int_0^{\mathcal{L}} d^D s\, d^D s'\, \delta^d[\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')]. \tag{5.9}$$

The above discussion suggests that it might be possible to generalize Eq. (5.9) to include fractal networks by simply making the replacement $D \to D_s$. However, this does not work, as can be easily verified. For example, requiring the Gaussian part of $\beta H_{\rm E}(D \to D_s)$ to be scale invariant does not give the same result for ν as Eq. (5.4) gives. The reason for this failure is that, whereas $\int d^D s$ implies an integration over a homogeneous and compact space, $\int d^{D_s} s$ implies neither. Rather than attempt to discuss a measure theory for fractal sets, which would be needed to understand fully the meaning of $\int d^{D_s} s$, I will simply propose an ansatz for establishing a correspondence between a fractal manifold and an homogeneous one that has the same physics. The resulting Hamiltonian can then be analyzed in a straightforward manner.

The only other parameter besides D_s that characterizes the internal space is $\mathcal{L}$. It is not surprising that an intrinsic length scale might be modified for the case of a fractal network, since, unlike homogeneous networks, which are characterized by a single intrinsic large length scale $\mathcal{L}$, fractals have many such scales, and it is not clear which scale is appropriate for integration. Assuming an integration scale of the form, $\mathcal{L} = \ell^{2/(2-D_s)} L^x$, we can fix x by requiring ($\mathbf{s} = \ell^{2/(2-D_s)}\boldsymbol{\sigma}$)

$$\int_0^{L^x} d^{D_s}\sigma = N \tag{5.10}$$

from which we obtain $x = d_{fl}/D_s$. The replacements,

$$\begin{aligned} L &\to L^{d_{fl}/D_s} \\ D &\to D_s, \end{aligned} \tag{5.11}$$

lead to the Hamiltonian,

$$\beta H'_{\rm E} = \frac{1}{2}\int_0^{\mathcal{L}'} d^{D_s} s \sum_{i=1}^{D_s}\left(\frac{\partial \mathbf{r}}{\partial s_i}\right)^2 + \frac{v_2}{2}\int_0^{\mathcal{L}'} d^{D_s} s\, d^{D_s} s'\, \delta^d[\mathbf{r}(\mathbf{s}) - \mathbf{r}(\mathbf{s}')] \tag{5.12}$$

($\mathcal{L}' = \ell^{2/(2-D_s)} L^{d_{fl}/D_s}$), which has the same upper-critical dimension and Gaussian scaling exponent as the original Sierpiński gasket model. It also leads to the same Flory theory prediction.

For example, applying the usual Gaussian integration techniques[41] to the calculation of the mean-squared distance between two points of the phantom network gives

$$\left\langle [\mathbf{r}(\mathbf{s_1}) - \mathbf{r}(\mathbf{s_2})]^2 \right\rangle = \frac{d\Gamma(\frac{1}{2}D_s)\ell^2}{\pi^{D_s/2}(2-D_s)} |\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2|^{2-D_s} \tag{5.13}$$

The rms end-to-end distance, which scales like R_g, is obtained by setting $|\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2| = L^{d_{fl}/D_s}$, from which we obtain a result identical to Eq. (5.4). Application of the usual power counting arguments to Eq. (5.12) gives results identical to Eqs. (5.7) and (5.8). In addition, ϵ-expansion results for ν can be obtained directly from the homogeneous-manifold results (cf. Refs. [93, 94, 116]),

$$\nu_\epsilon = \nu_0 + \frac{d_{fl}}{D_s} \frac{(2-D_s^*)\epsilon}{8\,[D_s^* + 2C(D_s^*)]} + O(\epsilon^2), \tag{5.14}$$

where

$$\epsilon = 4D_s - (2-D_s)\,d, \tag{5.15}$$

$$C(D_s^*) \equiv \frac{\sqrt{\pi}\,\Gamma\left(\frac{2}{2-D_s^*}\right)}{2^{2D_s^*/(2-D_s^*)}\Gamma\left(\frac{2+D_s^*}{2(2-D_s^*)}\right)}, \tag{5.16}$$

and

$$D_s^* = \frac{2d^*}{4+d^*} \tag{5.17}$$

The ϵ-expansion is performed about any point (D_s^*, d^*) that lies on the critical line that separates SA-relevant behavior from SA-irrelevant behavior. If Eq. (5.14) turns out to be consistent with numerical results for the Sierpiński gasket (which, as we shall see, it does), then they provide evidence for the correctness of the ansatz, Eq. (5.12).

5.3 Simulations

In order to perform Monte Carlo simulations, tethered network versions of a Sierpiński gasket were developed. The simulation model, simulation method, data analysis techniques, and

error analysis techniques are identical to those used in Chap. 4, with the following minor modifications:

1. The ratio ℓ/r_0 has been changed from the TM value of $2\sqrt{3}$ to 4 for reasons of computational convenience.[2]

2. The scaling exponents for the eigenvalues are now defined in terms of the generation number n of the Sierpiński gasket instead of the length scale L. Thus, in an obvious notation,

$$\nu_i^{(n)} = \frac{d_{fl}}{2} \frac{\ln\left(\lambda_i^{(n)}/\lambda_i^{(n-1)}\right)}{\ln\left(N^{(n)}/N^{(n-1)}\right)}, \tag{5.18}$$

and

$$\nu_i = \lim_{n\to\infty} \nu_i^{(n)}. \tag{5.19}$$

3. The definition of the shape parameter (Eq. (2.27)) has been changed to

$$S_d = \lim_{n\to\infty} S_d^{(n)}, \tag{5.20}$$

where

$$S_d^{(n)} = \left\langle \frac{\lambda_{d-2}^{(n)}}{\lambda_d^{(n)}} \right\rangle. \tag{5.21}$$

Notice that the definition of $S_d^{(n)}$ uses λ_{d-2} instead of λ_1 (cf. Eq. (2.27)). The motivation for this change is merely that in the simulations the statistical fluctuations in $\lambda_1^{(n)}$ are usually somewhat larger than those in $\lambda_{d-2}^{(n)}$. Thus, Eq. (5.20) provides a more stable measure of Sierpiński gasket shapes than Eq. (2.27). As long as $\lambda_1^{(n)}$ and $\lambda_{d-2}^{(n)}$ have the same scaling behavior, both definitions will provide the same information about the existence of a crumpled state. Of course, we have already observed in Sec. 4.4.3 that $\lambda_1^{(L)}$ and $\lambda_{d-2}^{(L)}$ may have different scaling behaviors, so this assumption may be wrong. However, as we shall see in Sec. 5.4, SG's already crumple in $\mathbb{R}^3$, so the assumption is correct.

[2]With the exception of Ref. [76], numerical studies done so far on TM models show no dependence of ν on the ratio of the ball size to the tether length. See Refs. [34, 45]. See Sec. 4.5.1 for a criticism of Ref. [76].

4. Because SG's crumple, the asphericity A_d (as defined in Eq. (2.28), but with $\lambda_i^{(L)} \to \lambda_i^{(n)}$) provides a meaningful measure of their shape. The simulation results include measurement of A_d.

5. Another consequence of SG's crumpling is that all the $\nu_i^{(n)}$'s are very similar for a given value of n. Consequently, Eq. (4.11) is easily satisfied and, therefore, does not give much insight into the size of the systematic errors. For this reason, systematic errors have been ignored in the error analysis.

5.4 Results

5.4.1 Phantom networks

Figure 5.2 shows the results of simulations of PSG's done for $d = 3$ and $d = 9$. To facilitate extrapolation of the data to the $n \to \infty$ limit, $\nu^{(n)}$, $S_d^{(n)}$, and $A_d^{(n)}$ have been plotted versus $1/n$. Also, the $\nu^{(n)}$s are plotted horizontally midway between the two relevant values of n. With the exception of the $n = 5$, $d = 3$ data point, the $\nu^{(n)}$s are all consistent with the theoretical prediction for ν_0 discussed in Sec. 5.2, and they appear to be independent of d. The failure of this one point to fit the theory probably reflects the presence of systematic errors. From the shape data, one can see that $S_d^{(n)}$ approaches a finite value in the $n \to \infty$ limit, indicating, as one would expect, that phantom networks crumple. Of course, since S_d compares the scaling of $\lambda_d^{(n)}$ and $\lambda_{d-2}^{(n)}$, it is possible that the smaller $\lambda_i^{(n)}$'s scale differently and that the PSG does not crumple. However, direct measurements of the $\nu_i^{(n)}$'s show that the $\lambda_i^{(n)}$'s are all the same (to within experimental error). Comparison of the A_d results with the corresponding polymer values [44] ($A_3 \approx 0.526$ and $A_9 \approx 0.431$) shows that crumpled phantom Sierpiński gaskets are much more spherical than their polymer counterparts.

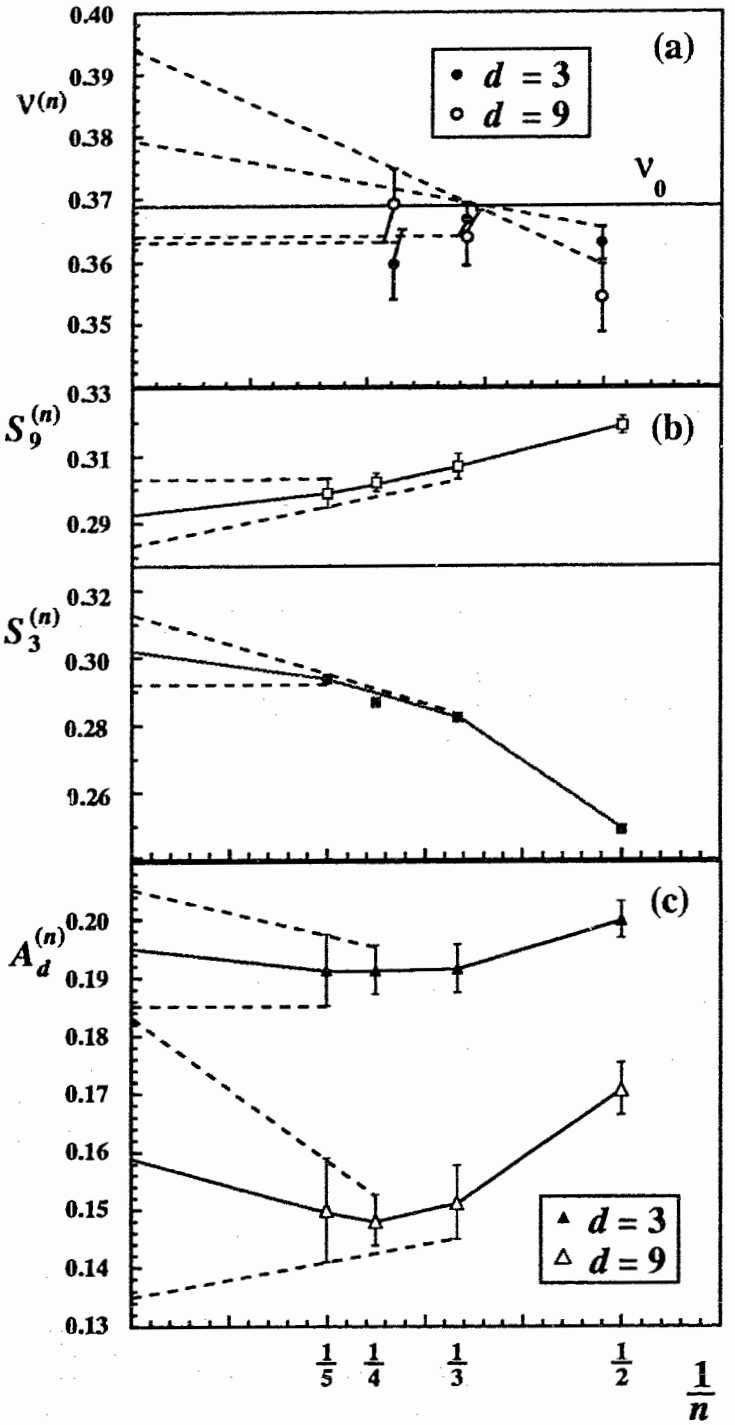


Figure 5.2: Results for the phantom-network simulations: (a) $\nu^{(n)}$ vs. $1/n$; the curves for extrapolating $\nu^{(n)}$ have been omitted for clarity; (b) $S_d^{(n)}$ vs. $1/n$; (c) $A_d^{(n)}$ vs. $1/n$. Error bars on all points reflect statistical errors only. Also note that the points in (a) have been positioned at $1/2\frac{1}{2}$, $1/3\frac{1}{2}$, etc. to reflect the fact that they have been determined using the two simulations whose n values they lie between.

Type	d	Length of run (in $\tau_{\rm bin}$) $n=2$	$n=3$	$n=4$	$n=5$	ν	S_d	A_d
Phantom	3	1452	1884	1808	1411	0.372±0.008	0.302±0.010	0.195±0.010
	9	87	79	232	67	0.379±0.016	0.292±0.010	0.159±0.024
SA	3	693	482	204	47	0.790±0.028	0.176±0.012	0.210±0.025
	7	523	709	163	18	0.451±0.019	0.295±0.018	0.177±0.022
	8	260	320	70	26	0.412±0.023	0.298±0.028	0.167±0.024
	9	442	557	280	17	0.395±0.018	0.291±0.011	0.172±0.032
	10	358	459	32		0.382±0.016	0.283±0.013	0.143±0.017
	11	203	225	67		0.361±0.007	0.289±0.006	0.133±0.010
	14	731	639	611		0.380±0.016	0.272±0.013	0.131±0.007

Table 5.1: Summary of simulation parameters and results. The run lengths do not include the initial discarded data. $\tau_{\rm bin}$ is determined from error analysis method 3 in Sec. 4.4.2. The quoted errors include both statistical and finite-size errors, but no systematic errors.

5.4.2 SA networks

Figure 5.3 shows the results for the $d = 3$ SA network. The Flory prediction for ν is remarkably good — differing from the data by at most a few percent. This is consistent with the $d = 2$ simulation of this model [42]. As with the phantom case, $S_d^{(n)}$ approaches a finite value for $n \to \infty$, indicating that the network is crumpled. Once again, since S_d compares the scaling of $\lambda_d^{(n)}$ and $\lambda_{d-2}^{(n)}$, it is possible that the smaller $\lambda_i^{(n)}$'s scale differently and that the PSG does not crumple. However, as with the phantom case, direct measurements of the $\nu_i^{(n)}$'s show that the $\lambda_i^{(n)}$'s are all the same (to within experimental error). The SASG S_d's are much smaller value than the corresponding phantom ones, which suggests that SASG's is much flatter in shape than PSG's. This conclusion is also borne out, to some extent, by the $A_3^{(n)}$ data. This is different from polymers, where self-avoidance has almost no effect on A_3 [68]. However, as d increases, the distinction between the phantom and SA values for A_d disappears.

Results for all the simulations are summarized in Table 5.1 and Fig. 5.4. The figure shows that the data are consistent with the Flory theory for $d \leq 8$. The first-order ϵ-expansion results are not as good for $d = 3$, but improve, as one would expect, as $d \to d_{uc}$. The

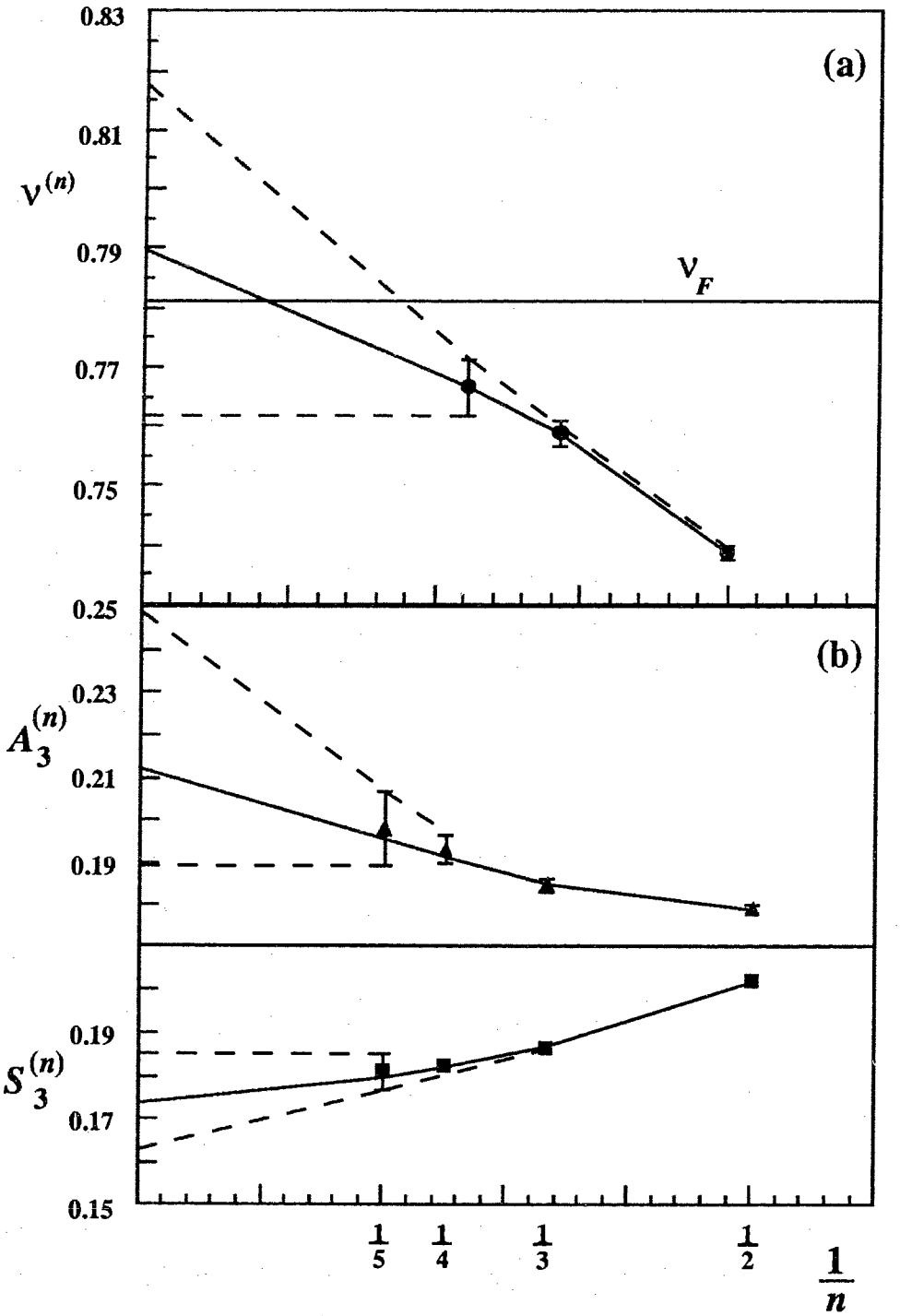


Figure 5.3: Results for the $d = 3$ SA-network simulations: (a) $\nu^{(n)}$ vs. $1/n$; the Flory theory prediction is shown for comparison; (b) $S_d^{(n)}$, $A_d^{(n)}$ vs. $1/n$. Error bars on all points reflect statistical errors only. The meaning of the dashed lines is explained Sec. 4.4.2. Also note that the points in (a) have been positioned at $1/2\frac{1}{2}$, $1/3\frac{1}{2}$, etc. to reflect the fact that they have been determined using the two simulations whose n values they lie between.

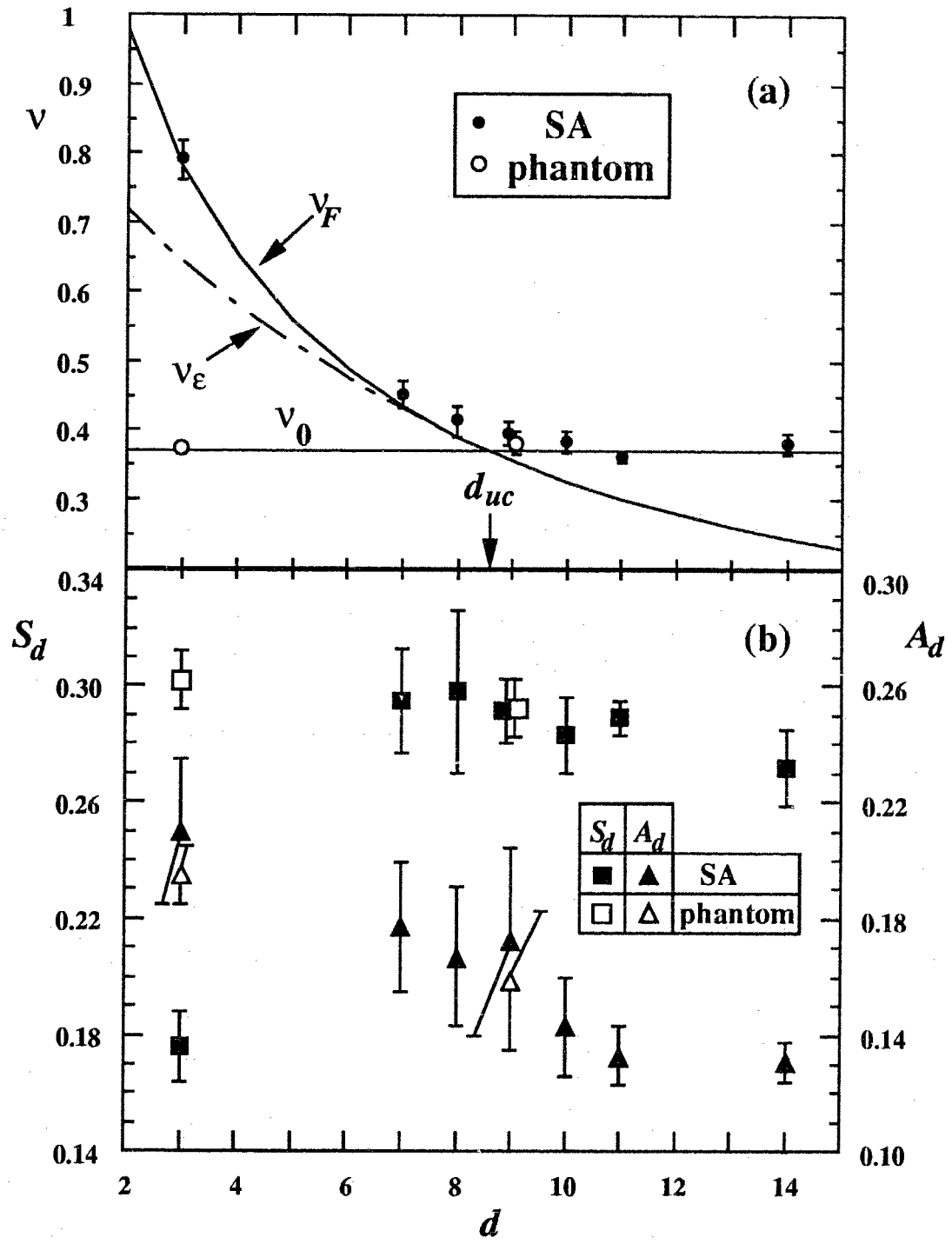


Figure 5.4: Summary of all simulations done: (a)ν vs. d. The Flory prediction (ν_F), obtained from Eq. (5.7), and the ϵ-expansion prediction (ν_ϵ), obtained from Eq. (5.14), are shown for comparison; (b)S_d,$A_d$vs.d. The error bars in both figures include finite-size effects as well as statistical errors.

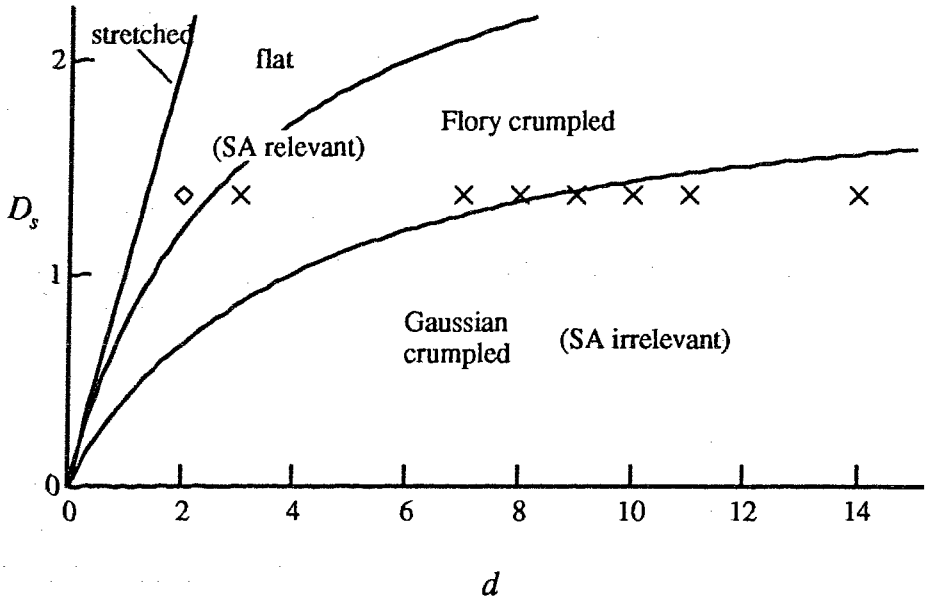


Figure 5.5: Phase diagram for SA regularly connected fractal networks. The crosses (×) represent the data from this simulation. The dashed curve is hypothetical; although the point ◇ (discussed in Ref. [42]) is taken to be evidence for its existence.

data also show a clear crossover to Gaussian behavior in the $8 \leq d \leq 10$ region. The exact location of the crossover is not very well determined, but it is certainly consistent with the prediction $d_{uc} \approx 8.6$.

There are several obvious extensions one could make to the present work, of which I mention only two. The first is to include bending energy in the simulation models and look for a crumpling transition. Currently, there are no known SATN models which exhibit a crumpling transition (they are always flat, always rough, or always crumpled[3], so it would, of course, be interesting to find a counter-example. The second extension is explore more fully the phase diagram in Fig.5.5.

Although not as interesting from a physical point of view as the corresponding diagram for homogeneous systems [12], Fig. 5.5 does have the advantage that the $1 < D_s < 2$ region of

[3]However, if it does turn out that SATM's embedded in $d \geq 5$ are crumpled, it is very likely (based on the PTM simulations of Ref. [37]) that a flat phase could also be found.

the diagram is accessible to simulations, whereas the $1 < D < 2$ region of the homogeneous diagram is not. Furthermore, the results of this chapter suggest that a correspondence between regular fractal manifolds and homogeneous ones may exist, although work remains to be done to justify the treatment of the Edwards model given here. This is important since theoretical analyses [19, 74, 20, 93, 94, 116] of membranes have so far been unsuccessful in discovering the flat phase (but see Ref. [117]). It may be possible, then, to set up a systematic program of simulations of regular fractal systems on the one hand, and theoretical calculations of the corresponding homogeneous manifold on the other. One could then explore in detail their relationship to questions concerning the existence of phase boundaries. For example, if a first-order phase boundary exists between the $D = 2$ and SA-irrelevant lines, ϵ-expansions will not detect it; but, a corresponding boundary in the $D_s - d$ plane could be found through simulation. Indeed, one such phase boundary may have already been discovered [42].

Many possibilities are open to you — work a little harder.
Chinese fortune cookie

Chapter 6

Conclusions

The following is a brief summary of the conclusions obtained from the work in Chaps. 3, 4, and 5 along with some thoughts about extensions to this thesis that are worth pursuing.

6.1 Two-dimensional closed pressurized random walks

The main motivation for constructing the model described in Chap. 3 was to study the shapes of a simplified version of a three-dimensional pressurized fluid vesicle. In particular, the hope was to create an analytic version of a computer model of the two-dimensional SA vesicle introduced by Leibler, Singh, and Fisher (the LSF model [2]), and to calculate the asphericity for this analytic model. Unfortunately, even the LSF model is sufficiently complicated so that simplifications have to be made in order to make analytic calculations possible. These simplifications consist of: (1) dropping the SA constraint; (2) replacing the true area that a closed SAW would enclose with the algebraic area; (3) replacing the fixed-length bonds between monomers with Gaussian springs. The resulting model is called the PRW model. With these simplifications it is possible to calculate exactly the asphericity A_2, as defined in Eq. (2.28). The results are given in Eq. (3.43) and Fig. 3.1.

Removing the SA constraint creates a model that is very different from the LSF model,

but it is, nonetheless, a reasonable first step towards an analytic description of a SA model. On the other hand, simplifications (2) and (3) turn out to be so drastic that the resulting model seems no longer to be related to the LSF model in any useful way. In particular, the PRW model does not distinguish between negative and positive pressure differences (as a result of simplification (2)); and, it does not seem possible to introduce a realistic bending rigidity (as a result of simplification (3)). It therefore seems unlikely that the PRW model, or some variation on it, will lead to a useful analytic description of the LSF model.

6.2 Self-avoiding tethered membranes

In Chap. 4 MC simulations on SATM's embedded in $\mathbb{R}^4$ and $\mathbb{R}^5$ were carried out in an attempt to find a SA crumpled state. No conclusive evidence for a crumpled state was found, but this possibility cannot be ruled out for the $d = 5$ simulation.

More generally, these simulations indicate that the two largest eigenvalues probably have the same scaling exponent, with $\nu_{\max}(d = 4) \approx 0.88$ and $\nu_{\max}(d = 5) \approx 0.73$. The $d = 4$ simulation results also indicate that the scaling exponents for the two smallest eigenvalues may be distinct from each other with $\nu_1(d = 4) \approx 0.65$ and $\nu_2(d = 4) \approx 0.72$. The $d = 5$ results are somewhat more ambiguous. One possibility is that SATM's embedded in $\mathbb{R}^5$ crumple with $\nu \approx 0.73$. Another possibility is that $d = 5$ SATM's are rough with the three smallest exponents ≈ 0.70. Finally, it is also possible that $\nu_2 = \nu_3 \approx 0.70$ and $\nu_1 \approx 0.65$. Unfortunately, $\nu_2(d = 4)$, $\nu_1(d = 5)$, $\nu_2(d = 5)$, and $\nu_3(d = 5)$ all have a strong L-dependence for the lattice sizes we have been able to simulate, so the determination of the thermodynamic phase for SATM's embedded in $\mathbb{R}^4$ and $\mathbb{R}^5$ will likely require simulations with much larger L-values. The large L-values required, combined with the need to run each simulation much longer (typically by a factor of 10) than other workers have assumed necessary (in order to adequately reduce the systematic error), make it nearly impossible to resolve the SATM phases by simulation, given the currently available computers.

These results should be contrasted with those in Ref. [1], which are that $d = 4$ SATM's are flat and $d = 5$ SATM's are crumpled with $\nu \approx 0.82$. Although there is no definitive explanation for the difference between the Ref. [1] results and the Chap. 4 results, the error analysis techniques described in Sec. 4.4.2 suggest that the Ref. [1] data might have significant systematic errors, and that this might be the source of the discrepancy.

6.3 Sierpiński gaskets

In Chap. 5 the results of MC simulations on phantom and SA Sierpiński gaskets were compared with appropriately generalized Flory theory and Edwards-model calculations. The motivation for doing this was to study a TN with an intrinsic connectivity between that of a polymer ($D = 1$) and that of a membrane ($D = 2$). The two main results of that chapter are: (1) replacing a TN's topological dimension with its spectral dimension and a careful generalization of the TN's intrinsic length scale (Eq. (5.11)) leads to an analytic theory that agrees quite well with the results of simulations; (2) simulated SA Sierpiński gaskets have an upper critical dimension ≈ 8.6, in agreement with theoretical predictions. For $d > d_{uc}$, SA is irrelevant and the simulated SASG's crumple with $\nu = \nu_0 \approx 0.368$. For $d < d_{uc}$, SA is relevant, but the simulated SASG's are, nevertheless, crumpled with $\nu \approx \nu_F \approx \nu_\epsilon$. These results are qualitatively similar to the case of polymers.

6.4 The SATN phase diagram

In Sec. 2.4 it was stated that much of the motivation for studying SATM's embedded in $\mathbb{R}^4$ and $\mathbb{R}^5$ and SASG's was to explore the (D, d) phase diagram shown in Fig. 2.4 and subsequently updated in Fig. 5.5. Figure 6.1 summarizes all the known data for SATN's. Unfortunately, the data are, in some cases, inadequate or controversial, so not all the phase boundaries can be drawn with confidence. In particular, the boundary between the flat and rough phases may cross the $D_s = 2$ line at a different value of d, or may not exist at all

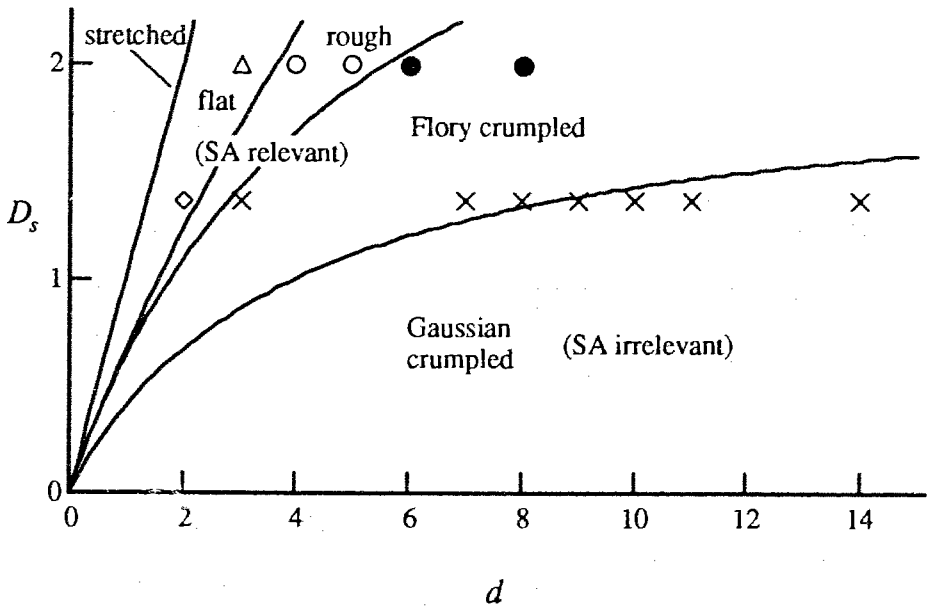


Figure 6.1: The D_s versus d phase diagram. The data points are taken from Chap. 4(o), Chap. 5(×), Ref. [42](◇), Ref. [34](△), and Ref. [1](•). This diagram assumes that SATM's embedded in $\mathbb{R}^4$ and $\mathbb{R}^5$ are rough and that those embedded in $\mathbb{R}^6$ are crumpled.

[1]. Similarly, the boundary between the Flory crumpled phase and the rough phase (or flat phase, if the rough phase does not exist) may or may not intersect the $D_s = 2$ line (although the results from Ref. [1] suggest that it does). Finally the two points in the flat phase ($\diamond$, $\triangle$) are simulations whose results are *consistent* with $\nu = 1$, but much longer runs (like the ones described in Chap. 4) may indicate that these points are, in fact, in a rough phase with $\nu \lesssim 1$.

Clearly, many more simulations will be needed before the SATN phase diagram is adequately understood.

In the thirties, under the demoralizing influence of quantum-theoretic perturbation theory, the mathematics required of a theoretical physicist was reduced to a rudimentary knowledge of the Latin and Greek alphabets.
R. Jost

Appendix A

This appendix contains some of the calculational details needed to derive Eqs. (3.34) and (3.42).

A.1 Derivation of Eq. (3.34)

Starting from the first line of Eq. (3.34),

$$W(y;\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) = \int D[\mathbf{r}]\exp\left[-\beta H(y) + \boldsymbol{\eta}_1\cdot\mathbf{r}(s_1) + \boldsymbol{\eta}_2\cdot\mathbf{r}(s_2)\right], \tag{A.1}$$

where

$$\begin{aligned}\beta H(y) &= \int_0^{\mathcal{L}} ds\left[\frac{1}{4}\left(\frac{d\mathbf{r}}{ds}\right)^2 - \frac{p}{2}\epsilon^{ij}r_i\frac{dr_j}{ds}\right] \\ &\quad + \frac{y}{2\mathcal{L}^2}\int_0^{\mathcal{L}} ds_1 ds_2\left[\mathbf{r}(s_1)-\mathbf{r}(s_2)\right]^2.\end{aligned} \tag{A.2}$$

The last term in the Hamiltonian can be rewritten as

$$\frac{y}{2\mathcal{L}^2}\int_0^{\mathcal{L}} ds_1 ds_2\left[\mathbf{r}(s_1)-\mathbf{r}(s_2)\right]^2 = \frac{y}{\mathcal{L}}\int_0^{\mathcal{L}} ds\, r^2(s) - \frac{y}{\mathcal{L}^2}\left[\int_0^{\mathcal{L}} ds\,\mathbf{r}(\mathbf{s})\right]^2. \tag{A.3}$$

The last term is difficult to handle in its present form, but we can rewrite it in a more convenient form using a trick described in Ref. [39],

$$\exp\left\{\frac{y}{\mathcal{L}^2}\left[\int_0^{\mathcal{L}} ds\,\mathbf{r}(\mathbf{s})\right]^2\right\} = \frac{\mathcal{L}}{\pi}\int d^2u\, e^{-\mathcal{L}u^2 - 2\sqrt{\frac{y}{\mathcal{L}}}\int_0^{\mathcal{L}} ds\,\mathbf{r}\cdot\mathbf{u}}. \tag{A.4}$$

Thus,

$$W(y,\eta_1,\eta_2)=\frac{\mathcal{L}}{\pi}\int d^2u\, e^{-\mathcal{L}u^2}\int D[\mathbf{r}]e^{\eta_1\cdot\mathbf{r}(\mathbf{s_1})+\eta_2\cdot\mathbf{r}(\mathbf{s_2})-\int_0^{\mathcal{L}}ds\,\mathcal{I}}, \tag{A.5}$$

where

$$\mathcal{I}=\frac{1}{4}\left(\frac{d\mathbf{r}}{ds}\right)^2+\frac{y}{\mathcal{L}}r^2+2\sqrt{\frac{y}{\mathcal{L}}}\mathbf{r}\cdot\mathbf{u}-\frac{p}{2}\epsilon^{ij}r_i\frac{dr_j}{ds}. \tag{A.6}$$

Now make the change of variables

$$\boldsymbol{\xi}=\mathbf{r}+x\mathbf{u}, \tag{A.7}$$

$$x\equiv\sqrt{\frac{\mathcal{L}}{y}}, \tag{A.8}$$

so that for the endpoints

$$\mathbf{r}^I=\mathbf{r}^F=0\rightarrow\boldsymbol{\xi}^I=\boldsymbol{\xi}^F=x\mathbf{u}, \tag{A.9}$$

and

$$r_j\frac{\partial r_k}{\partial s}\rightarrow(\xi_j-xu_j)\frac{\partial\xi_k}{\partial s}. \tag{A.10}$$

It follows that

$$\begin{aligned}\int_0^{\mathcal{L}}\mathcal{I} &= -\mathcal{L}u^2-\frac{p}{2}x\epsilon^{jk}u_j\left(\xi_k^F-\xi_k^I\right)\\ &\quad+\int_0^{\mathcal{L}}ds\left[\frac{1}{4}\left(\frac{\partial\boldsymbol{\xi}}{\partial s}\right)^2+\frac{\xi^2}{x^2}+\frac{p}{2}\epsilon^{jk}\xi_j\frac{\partial\xi_k}{\partial s}\right].\end{aligned} \tag{A.11}$$

Thus,

$$W(y;\eta_1,\eta_2)=\int d^2u\, e^{\frac{p}{2}x\epsilon^{jk}u_j(\xi_k^F-\xi_k^I)}\Gamma(u;\eta_1,\eta_2), \tag{A.12}$$

where

$$\begin{aligned}&\Gamma(u;\eta_1,\eta_2)=\\ &\frac{\mathcal{L}}{\pi}e^{-(\eta_1+\eta_2)\cdot\mathbf{u}x}\int d^2\xi_1 d^2\xi_2\, e^{\eta_1\cdot\boldsymbol{\xi}_1+\eta_2\cdot\boldsymbol{\xi}_2}G(\mathcal{L},s_2)G(s_2,s_1)G(s_1,0),\end{aligned} \tag{A.13}$$

$\boldsymbol{\xi}_i\equiv\boldsymbol{\xi}(s_i)$, and

$$G(s_2,s_1)=\int D[\boldsymbol{\xi}]\exp\left\{-\int_{s_1}^{s_2}ds\left[\frac{1}{4}\left(\frac{\partial\boldsymbol{\xi}}{\partial s}\right)^2+x^2\xi^2+\frac{p}{2}\epsilon^{jk}\xi_j\frac{\partial\xi_k}{\partial s}\right]\right\}. \tag{A.14}$$

Using standard path-integral techniques [118] gives

$$G(s_2, s_1) = e^{-S_{12}^{(\mathrm{cl})}} \int D[\mathbf{z}] e^{-S_0[\mathbf{z}]} \tag{A.15}$$

where

$$S_{12}^{(\mathrm{cl})}[\mathbf{q}] = \int_{s_1}^{s_2} ds \left[\frac{1}{4} \left(\frac{\partial \mathbf{q}}{\partial s} \right)^2 + x^2 q^2 + \frac{p}{2} \epsilon^{jk} q_j \frac{\partial q_k}{\partial s} \right] \tag{A.16}$$

is the classical action (q solves the classical equations of motion and $\mathbf{q}^I = \mathbf{q}^F = x\mathbf{u}$ are the fixed endpoints of the motion), and

$$S_0[\mathbf{z}] = \int_0^0 ds \left[\frac{1}{4} \left(\frac{\partial \mathbf{z}}{\partial s} \right)^2 + x^2 z^2 + \frac{p}{2} \epsilon^{jk} q_j \frac{\partial x_k}{\partial s} \right] \tag{A.17}$$

is the quadratic variation in $S_{12}^{(\mathrm{cl})}[\mathbf{q}]$. Straightforward evaluation of the integrals in Eqs. (A.16) and (A.15) gives

$$\begin{aligned} S_{12}^{(\mathrm{cl})}[\mathbf{q}] \;=\; & \frac{w}{4\mathcal{L} \sinh w(\sigma_2 - \sigma_1)} \Big[\left(q^{F2} + q^{I2} \right) \cosh w(\sigma_2 - \sigma_1) \\ & - 2\mathbf{q}^I \cdot \mathbf{q}^F \cosh \rho(\sigma_2 - \sigma_1) + 2\epsilon^{jk} q_j^I q_k^F \sin \rho(\sigma_2 - \sigma_1) \Big] \end{aligned} \tag{A.18}$$

and

$$\int D[\mathbf{z}] e^{-S_0[\mathbf{z}]} = \frac{w}{4\pi\mathcal{L} \sinh w(\sigma_2 - \sigma_1)}, \tag{A.19}$$

where

$$w = \left[4y\mathcal{L} - \rho^2 \right]^{1/2}, \tag{A.20}$$

$$\rho = p\mathcal{L}, \tag{A.21}$$

and

$$\sigma_i \equiv \frac{s_i}{\mathcal{L}}. \tag{A.22}$$

Hence,

$$\begin{aligned} G(\sigma_2, \sigma_1) \;=\; & \frac{w}{4\pi\mathcal{L} \sinh w(\sigma_2 - \sigma_1)} \exp \left\{ \frac{-w}{4\mathcal{L} \sinh w(\sigma_2 - \sigma_1)} \right. \\ & \Big[\left(q^{F2} + q^{I2} \right) \cosh w(\sigma_2 - \sigma_1) - 2\mathbf{q}^I \cdot \mathbf{q}^F \cosh \rho(\sigma_2 - \sigma_1) \\ & \left. + 2\epsilon^{jk} q_j^I q_k^F \sin \rho(\sigma_2 - \sigma_1) \Big] \right\}. \end{aligned} \tag{A.23}$$

Now evaluate Eq. (A.13). First note that

$$\Phi(\sigma_1;\sigma,\sigma') = \int d^2\xi\,(\sigma_1) e^{\boldsymbol{\eta}_1\cdot\boldsymbol{\xi}(\sigma_1)} G(\sigma',\sigma_1)G(\sigma_1,\sigma) \tag{A.24}$$

is a Gaussian integral. Direct evaluation gives

$$\begin{aligned}\Phi(\sigma_1;\sigma,\sigma') \;=\; & \exp\frac{1}{\sinh wT}\Big\{\eta_1^j\Big[\xi_j^I\cosh\rho t\sinh wt' + \xi_j^F\cosh\rho t'\sinh wt \\ & +\epsilon_{jk}\left(\xi_k^I\sinh\rho t\sinh wt' - \xi_k^F\sinh\rho t'\sinh wt\right)\Big] \\ & +\eta_1^2\mathcal{L}\frac{\sinh wt\sinh wt'}{w}\Big\}\,G(\sigma',\sigma),\end{aligned} \tag{A.25}$$

where $T=\sigma'-\sigma$, $t=\sigma_1-\sigma$, and $t'=\sigma'-\sigma_1$.

Inspection of Eq. (A.23) shows that $G(\sigma',\sigma)$ is only well-defined when $\sigma'\geq\sigma$, so we must be careful to write

$$\Gamma(\mathbf{u};\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) = \theta(\sigma_2-\sigma_1)\Gamma_+(\mathbf{u};\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) + \theta(\sigma_1-\sigma_2)\Gamma_-(\mathbf{u};\boldsymbol{\eta}_1,\boldsymbol{\eta}_2). \tag{A.26}$$

Explicitly,

$$\Gamma_+(\mathbf{u};\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) = \frac{L}{\pi}e^{-(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)\cdot\mathbf{u}x}\int d^2\xi(\sigma)\,\Phi(\sigma_2;\mathcal{L},\sigma)\Phi(\sigma_1;\sigma,0) \tag{A.27}$$

and

$$\Gamma_-(\mathbf{u};\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) = \Gamma_+(\mathbf{u};\boldsymbol{\eta}_1\leftrightarrow\boldsymbol{\eta}_2,\sigma_1\leftrightarrow\sigma_2). \tag{A.28}$$

Eqs. (A.27) and (A.28) are Gaussian integrals. After a lot of arithmetic, I find

$$\begin{aligned}\Gamma_+(\mathbf{u};\boldsymbol{\eta}_1,\boldsymbol{\eta}_2) = & \frac{w}{4\pi^2\sinh w}\exp\left\{\frac{w}{\mathcal{L}\sinh w}\left[-\frac{x^2u^2}{2}(\cosh w-\cosh\rho)\right.\right. \\ & + \frac{\mathcal{L}^2\eta_1^2}{w^2}\sinh w\sigma_1\sinh w(1-\sigma_1) + \frac{\mathcal{L}^2\eta_2^2}{w^2}\sinh w\sigma_2\sinh w(1-\sigma_2) \\ & + \frac{x\mathcal{L}}{w}\boldsymbol{\eta}_1\cdot\mathbf{u}\left(\sinh w\sigma_1\cosh\rho(1-\sigma_1)+\cosh\rho\sigma_1\sinh w(1-\sigma_1)-\sinh w\right) \\ & + \frac{x\mathcal{L}}{w}\boldsymbol{\eta}_2\cdot\mathbf{u}\left(\sinh w\sigma_2\cosh\rho(1-\sigma_2)+\cosh\rho\sigma_2\sinh w(1-\sigma_2)-\sinh w\right) \\ & + \epsilon^{jk}\left(\frac{x\mathcal{L}}{w}\eta_{1j}u_k\left[\sin\rho\sigma_1\sinh w(1-\sigma_1)-\sinh w\sigma_1\sin\rho(1-\sigma_1)\right]\right. \\ & \quad +\frac{x\mathcal{L}}{w}\eta_{2j}u_k\left[\sin\rho\sigma_2\sinh w(1-\sigma_2)-\sinh w\sigma_2\sin\rho(1-\sigma_2)\right] \\ & \quad \left.\left.\left.-\frac{2\mathcal{L}^2\eta_{1j}\eta_{2k}}{w^2}\sinh w\sigma_1\sinh w(1-\sigma_2)\sin\rho(\sigma_2-\sigma_1)\right)\right]\right\}\end{aligned} \tag{A.29}$$

and a similar expression for $\Gamma_-(\mathbf{u};\eta_1,\eta_2)$. Finally, I calculate

$$W(y;\eta_1,\eta_2) = \theta(\sigma_2-\sigma_1)W_+(\eta_1,\eta_2) + \theta(\sigma_1-\sigma_2)W_-(\eta_1,\eta_2), \tag{A.30}$$

where

$$W_\pm(\eta_1,\eta_2) = \int d^2u\, \Gamma_\pm(\mathbf{u};\eta_1,\eta_2). \tag{A.31}$$

Inserting Eq. (A.29) and the corresponding expression for $\Gamma_-(\mathbf{u};\eta_1,\eta_2)$ into Eq. (A.31) give, after another set of tedious Gaussian integrations, the result on the second line of Eq. (3.34).

A.2 Derivation of Eq. (3.42)

Start with

$$\begin{aligned}\left\langle \mathrm{Tr}\hat{Q}^2\right\rangle_{H(y)} &= \left\langle \mathrm{Tr}Q^2\right\rangle_{H(y)} - \frac{1}{d}\left\langle (\mathrm{Tr}\,Q)^2\right\rangle_{H(y)} \\ &= T_1 - T_2 - T_3,\end{aligned} \tag{A.32}$$

where,

$$T_1 = \left\langle Q'_{jk}Q'_{kj}\right\rangle - \frac{1}{d}\left\langle \left(Q'_{jj}\right)^2\right\rangle, \tag{A.33}$$

$$T_2 = \left\langle R_{jk}Q'_{kj}\right\rangle - \frac{1}{d}\left\langle Q'_{jj}R_{kk}\right\rangle, \tag{A.34}$$

$$T_3 = \langle R_{jk}R_{kj}\rangle - \frac{1}{d}\left\langle (R_{jj})^2\right\rangle, \tag{A.35}$$

and the definitions,

$$Q'_{jk} = \frac{1}{\mathcal{L}}\int_0^{\mathcal{L}} ds\, r_j(s)r_k(s) \tag{A.36}$$

and

$$R_{jk} = \frac{1}{\mathcal{L}^2}\int_0^{\mathcal{L}} ds\, ds'\, r_j(s)r_k(s') \tag{A.37}$$

have been introduced. The first term in Eq. (A.33) is

$$\left\langle Q'_{jk}Q'_{kj}\right\rangle = \left\langle \frac{1}{\mathcal{L}^2}\int_0^{\mathcal{L}} ds_1 ds_2\, r_j(s_1)r_k(s_1)r_k(s_2)r_j(s_2)\right\rangle \tag{A.38}$$

$$
\begin{aligned}
&= \frac{1}{\mathcal{L}^2}\int_0^{\mathcal{L}} ds_1 ds_2\, G^{(4)}_{jkkj}(s_1,s_1,s_2,s_2) \\
&= \frac{1}{\mathcal{L}^2}\int_0^{\mathcal{L}} ds_1 ds_2 \left[G_{jk}(s_1,s_1)G_{kj}(s_2,s_2)\right. \\
&\quad \left. + G_{jk}(s_1,s_2)G_{kj}(s_1,s_2) + G_{jj}(s_1,s_2)G_{kk}(s_1,s_2)\right],
\end{aligned}
$$

where

$$
G_{jk}(s_1,s_2) = \frac{1}{W}\frac{\partial^2}{\partial\eta_1^j \partial\eta_2^k} W(y;\boldsymbol{\eta}_1,\boldsymbol{\eta}_2). \tag{A.39}
$$

The last line in Eq. (A.38) follows from Wick's theorem. Using the result in Eq. (3.34) gives (recall that $\sigma_i = s_i/\mathcal{L}$),

$$
G_{jk}(\sigma_i,\sigma_i) = \frac{2\alpha_i\mathcal{L}}{w\Delta}\delta_{jk}, \tag{A.40}
$$

and

$$
G_{jk}(\sigma_1,\sigma_2) = \frac{1}{w\Delta}\left[\beta\delta_{jk} - \phi\epsilon_{jk}\right], \tag{A.41}
$$

where, Δ, α_i, β, ϕ, are defined in Eqs. (3.36)–(3.38). Inserting Eq. (A.41) into Eq. (A.38) gives

$$
\left\langle Q'_{jk}Q'_{kj}\right\rangle = \frac{2\mathcal{L}^2}{w^2\Delta^2}\left[4\alpha_1\alpha_2 + 3\beta^2 + \phi\right]. \tag{A.42}
$$

The second term in Eq. (A.33) is

$$
\begin{aligned}
\langle r_j(s_1)r_j(s_1)r_k(s_2)r_k(s_2)\rangle &= G_{jj}(s_1,s_1)G_{kk}(s_2,s_2) \\
&\quad + 2G_{jk}(s_1,s_2)G_{jk}(s_1,s_2) \\
&= \frac{\mathcal{L}^2}{w^2\Delta^2}\left[16\alpha_1\alpha_2 + 4\beta^2 - 4\phi^2\right].
\end{aligned} \tag{A.43}
$$

Hence,

$$
T_1 = \frac{4\mathcal{L}^2}{w^2\Delta^2}\int_0^1 d\sigma_1 d\sigma_2 \left[\beta^2 + \phi^2\right]. \tag{A.44}
$$

This integral can be done using elementary techniques, but the number of terms in the integrand make a computation by hand prohibitive. The integral was actually done using the MAPLE symbolic mathematics program. Maple finds

$$
T_1 = 4\mathcal{L}^2\frac{\sinh^2 w}{w^2\Delta^2} + \mathcal{L}\frac{3w\sinh\rho - \rho\sinh w}{\rho w\Delta y} - \frac{2}{y^2}. \tag{A.45}
$$

A similar treatment of Eqs. (A.34) and (A.35) give

$$T_2 = 4\mathcal{L}^2 \frac{\sinh^2 w}{w^2 \Delta^2} + \mathcal{L} \frac{w \sinh \rho - 3\rho \sinh w}{\rho w \Delta y} \tag{A.46}$$

and

$$T_3 = 4\mathcal{L}^2 \frac{\sinh^2 w}{w^2 \Delta^2} - 4\mathcal{L} \frac{\sinh w}{w \Delta y} + \frac{1}{y^2}. \tag{A.47}$$

Inserting Eqs. (A.45), (A.46) and (A.47) into Eq. (A.32) gives Eq. (3.42).

Bibliography

[1] G. S. Grest, Journal de Physique I (France) **1**, 1695 (1991).

[2] S. Leibler, R. R. P. Singh, and M. E. Fisher, Physical Review Letters **57**, 1989 (1987).

[3] A. L. Kholodenko, Annals of Physics **202**, 186 (1990).

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

[5] *Fluctuation Phenomena*, edited by E. W. Montroll and J. L. Lebowitz (North Holland Physics Publishing, Amsterdam, 1987), especially Chaps. 1 and 2.

[6] M. Barber and B. W. Ninham, *Random and Restricted Walks; Theory and Applications* (Gordon and Breach, New York, 1970).

[7] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (J. Wiley and Sons, New York, 1986).

[8] Y. Oono, Advances in Chemical Physics **61**, 301 (1985).

[9] J. Frölich, in *Applications of Field Theory to Statistical Mechanics*, Vol. 216 of *Lecture Notes in Physics*, edited by L. Garrido (Springer-Verlag, Berlin, 1985), p. 32.

[10] A. Polyakov, Physics Letters **103B**, 207 (1981).

[11] R. Lipowsky, Nature **349**, 475 (1991).

[12] D. Nelson, in *Statistical Mechanics of Membranes and Surfaces*, Vol. 5 of *Jerusalem Winter School for Theoretical Physics*, edited by D. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989), p. 1.

[13] F. David, to appear in *Two Dimensional Gravity and Random Surfaces*, Vol. 8 of *Jerusalem Winter School for Theoretical Physics* (unpublished).

[14] C. Itzykson and J.-M. Drouffe, *Statistical Field Theory* (Cambridge University Press, Cambridge, 1989), Vol. 1.

[15] *Statistical Mechanics of Membranes and Surfaces*, Vol. 5 of *Jerusalem Winter School for Theoretical Physics*, edited by D. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989).

[16] D. Nelson, in *Statistical Mechanics of Membranes and Surfaces*, Vol. 5 of *Jerusalem Winter School for Theoretical Physics*, edited by D. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989), p. 137.

[17] T. Hwa, E. Kokufuta, and T. Tanaka, Physical Review A **44**, R2235 (1991).

[18] X. Wen *et al.* (unpublished).

[19] Y. Kantor, in *Statistical Mechanics of Membranes and Surfaces*, Vol. 5 of *Jerusalem Winter School for Theoretical Physics*, edited by D. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989), p. 115.

[20] B. Duplantier, in *Statistical Mechanics of Membranes and Surfaces*, Vol. 5 of *Jerusalem Winter School for Theoretical Physics*, edited by D. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989), p. 226.

[21] M. Bloom, E. Evans, and O. Mouritsen, Quarterly Reviews of Biophysics **24**, 293 (1991).

[22] S. Leibler, in *Statistical Mechanics of Membranes and Surfaces*, Vol. 5 of *Jerusalem Winter School for Theoretical Physics*, edited by D. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989), p. 46.

[23] M. E. Fisher, Physica D **38**, 112 (1989).

[24] C. J. Camacho and M. E. Fisher, Physical Review Letters **65**, 9 (1990).

[25] A. C. Maggs, S. Leibler, M. E. Fisher, and C. J. Camacho, Physical Review A **42**, 691 (1990).

[26] C. J. Camacho, M. E. Fisher, and R. R. P. Singh, Journal of Chemical Physics **94**, 5693 (1991).

[27] C. J. Camacho and M. E. Fisher, in *Computer Simulation Studies in Condensed Matter Physics IV*, edited by D. P. Landau, K. K. Mon, and H. B. Schuttler (Springer-Verlag, Berlin, 1991).

[28] J. Rudnick and G. Gaspari, Science **252**, 422 (1991).

[29] E. Levinson, Physical Review A **45**, 3629 (1992).

[30] A. Khurana, Physics Today **42**, 17 (August 1989).

[31] Y. Kantor, M. Kardar, and D. R. Nelson, Physical Review Letters **57**, 791 (1986).

[32] Y. Kantor, M. Kardar, and D. R. Nelson, Physical Review A **35**, 3056 (1987).

[33] M. Plischke and D. H. Boal, Physical Review A **38**, 4943 (1988).

[34] F. F. Abraham, W. E. Rudge, and M. Plischke, Physical Review Letters **62**, 1757 (1989).

[35] A. Baumgartner, Journal de Physique I (France) **1**, 1549 (1991).

[36] F. David and E. Guitter, Europhysics Letters **5**, 709 (1988).

[37] J. Ambjorn, B. Durhuus, and T. Jonsson, Nuclear Physics **B316**, 526 (1989).

[38] A. Billoire, D. J. Gross, and E. Marinari, Physics Letters **139B**, 75 (1984).

[39] D. C. Khandekar and F. W. Wiegel, Journal de Physique (Paris) **50**, 263 (1989).

[40] H. Goldstein, *Classical Mechanics* (Addison Wesley, Reading, 1980).

[41] J. des Cloizeaux, Journal de Physique (Paris) **42**, 635 (1981).

[42] E. Duering and Y. Kantor, Physical Review B **40**, 7443 (1989).

[43] E. Levinson, Physical Review A **43**, 5233 (1991).

[44] J. Rudnick and G. Gaspari, Science **237**, 384 (1987).

[45] D. H. Boal, E. Levinson, D. Liu, and M. Plischke, Physical Review A **40**, 3292 (1989).

[46] D. Stauffer, M. Ferer, and M. Wortis, Physical Review Letters **29**, 345 (1972).

[47] F. Family, T. Vicsek, and P. Meakin, Physical Review Letters **55**, 641 (1985).

[48] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics* (World Scientific, Singapore, 1990).

[49] F. W. Wiegel, *Introduction to Path Integrals Methods in Physics and Polymer Science* (World Scientific, Singapore, 1986).

[50] S. F. Edwards, Proceedings of the Physical Society **85**, 613 (1965), a more modern view of this model can be found in Refs. [20] and [51].

[51] B. Duplantier, to appear in *Fundamental Problems in Statistical Mechanics VII*, H. van Beijeren, ed. (unpublished).

[52] A. Malakis, Journal of Physics A **9**, 1283 (1976).

[53] Y. Oono and K. F. Freed, Journal of Chemical Physics **75**, 993 (1981).

[54] Y. Shapir and Y. Oono, Journal of Physics A **17**, L39 (1984).

[55] P. J. Flory, *Statistics of Chain Molecules* (Interscience Publishers, New York, 1969).

[56] W. Helfrich, Z. Naturforsch. **28c**, 693 (1973), see also Ref. [78].

[57] M. A. Peterson, Mol. Cryst. Liq. Cryst. **127**, 159 (1985), see also Ref. [119].

[58] J.-S. Ho and A. Baumgartner, Molecular Simulations **6**, 163 (1991).

[59] D. H. Boal and M. Rao, Physical Review A **45**, R6947 (1992).

[60] A. Baumgartner and J.-S. Ho, Physical Review A **41**, 5747 (1990).

[61] D. H. Boal, Physical Review A **43**, 6771 (1991).

[62] K. Šolc and W. H. Stockmayer, Journal of Chemical Physics **54**, 2756 (1971), see also Ref. [120].

[63] M. Bishop and J. P. J. Michels, Journal of Chemical Physics Letters **82**, 1059 (1985), and references cited therein.

[64] M. Bishop and C. J. Saltiel, Journal of Chemical Physics **85**, 6728 (1986), also see Ref. [121].

[65] J. Rudnick and G. Gaspari, Journal of Physics A **19**, L191 (1986).

[66] J. Rudnick, A. Beldjenna, and G. Gaspari, Journal of Physics A **20**, 971 (1987).

[67] G. Gaspari, J. Rudnick, and A. Beldjenna, Journal of Physics A **20**, 3393 (1987).

[68] J. A. Aronovitz and D. R. Nelson, Journal de Physique (Paris) **47**, 1445 (1986).

[69] H.-W. Diehl and E. Eisenriegler, Journal of Physics A **22**, L87 (1989).

[70] A. D. Verga, Journal of Physics A **24**, L561 (1991).

[71] D. Boal, U. Seifert, and A. Zilker (unpublished).

[72] A. Baumgartner, lectures at the workshop on '25 Years of the Edwards Model', Puri, India, 1991 (unpublished).

[73] H. E. Stanley, P. J. Reynolds, S. Redner, and F. Family, in *Real Space Renormalization*, edited by T. W. Burkhardt and J. M. J. van Leeuwen (Springer-Verlag, New York, 1982), p. 169.

[74] E. Bouchard and J. P. Bouchard, Journal de Physique (Paris) **50**, 829 (1989).

[75] B. Nienhuis, Physical Review Letters **49**, 1062 (1982).

[76] F. F. Abraham and D. R. Nelson, Journal de Physique (Paris) **51**, 2653 (1990).

[77] P. B. Canham, Journal of Theoretical Biology **26**, 61 (1970).

[78] H. J. Dueling and W. Helfrich, Journal de Physique (Paris) **37**, 1335 (1976).

[79] L. Peliti and S. Leibler, Physical Review Letters **54**, 1690 (1985).

[80] H. Tasaki and T. Hara, Physics Letters **121B**, 115 (1985).

[81] U. Glaus, Journal of Statistical Physics **50**, 1141 (1988).

[82] J. B. Kogut, Reviews of Modern Physics **51**, 659 (1979).

[83] A. Polyakov, Nuclear Physics **B268**, 406 (1986).

[84] G. Parisi, Physics Letters **81B**, 357 (1979).

[85] B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco, 1982).

[86] D. J. Gross, Physics Letters **138B**, 185 (1984).

[87] A. Maritan and A. Stella, Physical Review Letters **53**, 123 (1984).

[88] B. Duplantier, Physics Letters **141B**, 239 (1987).

[89] J. F. B. Durhuus and T. Jonsson, Nuclear Physics **B225**, 185 (1983).

[90] J. F. B. Durhuus and T. Jonsson, Nuclear Physics **B240**, 453 (1984).

[91] M. E. Cates, Physics Letters **161B**, 363 (1985).

[92] M. Kardar and D. R. Nelson, Physical Review Letters **58**, 12 (1987).

[93] J. A. Aronovitz and T. C. Lubensky, Europhysics Letters 4, 395 (1987).

[94] B. Duplantier, Physical Review Letters **58**, 2733 (1987).

[95] D. R. Nelson and L. Peliti, Journal de Physique (Paris) **48**, 1085 (1987).

[96] J. A. Aronovitz and T. C. Lubensky, Physical Review Letters **60**, 2634 (1988).

[97] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, Vol. 7 of *Course of Theoretical Physics* (Pergamon Press, Oxford, 1986).

[98] M. Paczuski, M. Kardar, and D. R. Nelson, Physical Review Letters **60**, 2638 (1988).

[99] M. Paczuski and M. Kardar, Physical Review A **39**, 6086 (1989).

[100] J. Aronovitz, L. Golubović, and T. C. Lubensky, Journal de Physique (Paris) **50**, 609 (1989).

[101] J.-S. Ho and A. Baumgartner, Physical Review Letters **63**, 1324 (1989).

[102] N. Metropolis *et al.*, Journal of Chemical Physics **21**, 1087 (1953), see Ref. [122] for a more recent discussion of the Metropolis algorithm.

[103] A. Baumgartner, in *Applications of the Monte Carlo Method in Statistical Physics*, Vol. 216 of *Topics in Current Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1987), p. 145.

[104] L. Lyons, *Statistics for Nuclear and Particle Physicists* (Cambridge University Press, Cambridge, 1986).

[105] P. A. P. Moran, Biometrika **62**, 1 (1975).

[106] A. Baumgartner and W. Renz, Europhysics Letters **17**, 381 (1992).

[107] A. Baumgartner, private communication (unpublished).

[108] F. F. Abraham and D. R. Nelson, Science **249**, 393 (1990).

[109] G. Grest, private communication (unpublished).

[110] S. Alexander and R. Orbach, Journal de Physique (Paris) Letters **43**, L625 (1982).

[111] R. Rammal and G. Toulouse, Journal de Physique (Paris) Letters **44**, L13 (1983).

[112] M. E. Cates, Journal de Physique (Paris) **46**, 1059 (1985).

[113] B. O'Shaunessy and I. Procaccia, Physical Review A **32**, 3073 (1985).

[114] R. A. Guyer, Physical Review A **29**, 2751 (1984).

[115] J. A. Given and B. Mandelbrot, Journal of Physics B **16**, L565 (1983).

[116] M. Kardar and D. R. Nelson, Physical Review A **38**, 966 (1988).

[117] T. Hwa, Physical Review A **41**, 1751 (1990).

[118] L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981).

[119] S. T. Milner and A. Safran, Physical Review A **36**, 4371 (1987).

[120] K. Šolc, Journal of Chemical Physics **55**, 335 (1971).

[121] M. Bishop and C. J. Saltiel, Journal of Chemical Physics **88**, 3976 (1988).

[122] K. Binder, in *Applications of the Monte Carlo Method in Statistical Physics*, Vol. 216 of *Topics in Current Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1987), p. 1.