Journal of Statistical Physics, Vol. 55, Nos. 1/2, 1989

A Renormalization Group Analysis of Lattice Models of Two-Dimensional Membranes

Jan Ambjørn,[1] Bergfinnur Durhuus,[2] Jürg Fröhlich,[3] and Thórdur Jónsson[4]

Received November 7, 1988

We study lattice models of two-dimensional membranes of interest in statistical physics. The energy functional of a membrane is expressed as a sum of terms proportional to the surface area of the membrane, an extrinsic curvature and an intrinsic curvature quantity, respectively, but we neglect excluded volume effects. We introduce a renormalization transformation for these models which preserves the form of the energy functional, up to nonlocal terms. Our renormalization group construction is used to analyze the phase diagram and the different critical regimes of our models. We find evidence for a crumpling transition, separating a regime where surfaces are "crystalline" from one where the surfaces collapse to branched polymers, and we find a third genuine random-surface regime.

KEY WORDS: Two-dimensional lattice membranes; random surfaces; renormalization group analysis; crystalline surfaces; crumpling transition; collapse to branched polymers.

1. INTRODUCTION

1.1. Motivation, Physical Problems, Perspective

The statistical mechanics of random surfaces (RS) has been the object of much recent study in condensed matter physics (see, e.g., ref. 1) and quantum field theory.[2] The motivation behind the analysis of a considerable number of different RS models has emerged from investigating the following topics.

[1] Niels Bohr Institute, 2100 Copenhagen Ø, Denmark.
[2] Mathematics Institute, 2100 Copenhagen Ø, Denmark.
[3] Theoretical Physics, ETH-Hönggerberg, 8093 Zürich, Switzerland.
[4] Nordita, 2100 Copenhagen Ø, Denmark.

0022-4715/89/0400-0029$06.00/0

1. Crystal growth, properties of crystalline surfaces.
2. Interfaces separating different phases of a physical system, such as domain walls in a ferromagnet; domain wall wandering in uniaxial commensurate–incommensurate transitions; wetting; etc.
3. Soap layers separating oil and water regions in microemulsions.
4. Lipid bilayers, theory of red blood cells.
5. Membranelike structures arising in systems of polymers.
6. Membranes with crystalline ("tethered surfaces") or hexatic order.
7. Statistical properties of foam.

It should be mentioned that there are, in addition, more formal, mathematical reasons for the study of RS models: Many three-dimensional models of classical statistical mechanics, in particular spin systems, are secretly theories of random surfaces. The best known example is the three-dimensional Ising model. Another example is percolation theory. Important properties of bond percolation in three dimensions have been established by studying the dual theory of plaquette percolation, which leads to interesting random surface problems.

Studies of RS models in connection with quantum field theory and particle physics were motivated by the following problems.

(a) The quark confinement problem in Yang–Mills theory: As an attempt to understand permanent confinement, various random surface representations of lattice gauge theories have been proposed and exploited. Physically, the random surfaces arise, roughly speaking, as world sheets of narrowly focused chromoelectric flux tubes.

(b) Various random surface theories have been proposed as low-energy approximations to confining gauge theories (large-N gauge theory as a surface theory, "smooth strings," ...).

(c) Formulations of string and superstring theories as genuine random surface theories; discrete approximations to string theory which could be studied numerically.

Our own work on random surfaces has been motivated by most of the topics described above, although among our main motivations were the confinement problem in lattice gauge theory and the questions of whether there is a quantization of the bosonic string which avoids the tachyon and whether string theory admits a discrete approximation, comparable to the lattice approximation in quantum field theory, well suited for rigorous analytical and numerical investigations. Our and many other attempts to settle these questions have only been moderately successful—they are still

largely open, although they may have clarified some of the obstacles encountered in the search for a bosonic string theory without a tachyon, in approximating string theory by discrete RS models, and in posing the problem of unitarity in string theory.

One consequence of this work may be that, in contrast to conventional field theory, which generally admits a decent lattice approximation, it does not appear to be possible to approximate string theory by models of random surfaces embedded in regular lattices (it appears to be impossible to find simple models of single surfaces which *scale correctly*, as a critical point is approached, *and* which have a Euclidean-invariant scaling limit).

In this paper we nevertheless return to our study of models of random surfaces embedded in a regular lattice. Physically, this is motivated by topics 3–7 described above. There *are* idealized situations in condensed matter physics which one expects can be described reasonably well by models of lattice random surfaces. But our main motivation is more mathematical: In order to guess what to expect from RS theory, it might be good to have some simple models at hand which can be understood in rather much detail and which have features that one may expect to find again in more complicated models. In a companion paper[3,4] we have studied triangulated RS models.

There now exists a good theory of critical phenomena in spin systems, lattice gases, and many other physical systems which admit a field-theoretic formulation or a representation as gases of interacting random paths. This theory is the renormalization group. However, critical behavior in systems involving statistical fluctuations of surfaces is not yet well understood. In general, not even the phase diagrams of such systems are known. We therefore think that it may be of interest to describe and study a class of models of random surfaces whose phase diagram can be mapped out with some precision and whose critical behavior appears to be quite accessible to analysis. That is the main subject of this paper. Our models are cooked up in such a way that renormalization group ideas can be used to study them. We do not pretend that we have a very detailed picture of the renormalization flow in our models or that the specific renormalization procedure that we shall devise in this paper is applicable to truly realistic models. But we hope our work may trigger some further developments in the direction of a renormalization group analysis of a general class of RS models.

Roughly speaking, the models of lattice random surfaces that we study in this paper are defined as follows: We specify an ensemble $\mathscr{E}$ of lattice random surfaces. We shall focus our attention on the ensembles of *planar* and *self-avoiding* RS [these ensembles are carefully defined in Section 2,

examples (iv) and (v)]. To each surface $S \in \mathscr{E}$, we then assign a statistical weight

$$\rho_{\underline{\beta}}(S) = \exp[-A_{\underline{\beta}}(S)] > 0 \tag{1.1}$$

where $A_{\underline{\beta}}(S)$ is an "action" or "*energy function*" on $\mathscr{E}$, and $\underline{\beta}$ is a set of "coupling constants." Our choice of $A_{\underline{\beta}}$ will be as follows:

$$A_{\underline{\beta}}(S) = A_{\text{int}}(S) + A_{\text{ext}}(S) \tag{1.2}$$

where $A_{\text{int}}(S)$ depends only on the *intrinsic* geometry of S, while $A_{\text{ext}}(S)$ depends on the way S is arranged inside the lattice $\mathbb{Z}^d$. Typically,

$$A_{\text{int}}(S) = \beta_2 |S| + \beta_0' \Delta(S) \tag{1.3}$$

where $|S|$ is the area of S, and $\Delta(S)$ is a sum over absolute values (or squares) of intrinsic curvature quantities (defect angles) associated with the vertices of S. Such terms in the action of surfaces arise naturally in the study of polymerized or crystalline membranes. They were first discussed in ref. 5.

It was shown in ref. 6, *under some hypothesis only checked numerically*, that if $\beta_0' = 0$ and $A_{\text{ext}}(S) \equiv 0$, the scaling limit of planar RS models with action given by the area is equivalent to the theory of a scalar free field. In the continuum limit, the surface tension diverges to $+\infty$, while the physical mass stays finite. The same result is likely to be true if a critical regime is approached where β_0' and the coupling constants of $A_{\text{ext}}(S)$ stay bounded.[5] The intuitive reason is that the surfaces in these models collapse into branched polymers. Highly branched, thin surfaces have little area, hence small action, but enormous entropy. (This phenomen is probably related to the tachyon problem in bosonic string theory.) In order to suppress this "collapse to triviality," it was proposed in ref. 6 and discussed in more detail in refs. 4 and 5 to add a term $\beta_0' \Delta(S)$ to the action $A_0(S) \equiv \beta_2 |S|$ and to approach a critical regime, where $\beta_0' \to \infty$, while $\beta_2 = \beta_2(\beta_0') \to 0$. In perturbation theory for continuum RS theories, there is no term of the form $\Delta(S)$ which is renormalizable by power counting. This may explain why our proposal was largely ignored. However, in ref. 4 we have shown that, in the context of triangulated RS models, our proposal is well conceived and appears to lead to a meaningful RS theory. For models of random surfaces embedded in the lattice, the proposal made in refs. 4–6 is physically well motivated and viable, provided the dimension of the lattice is $d = 3$. In higher dimensional lattice models, the action must involve extrinsic curvature terms for the physical mass and string tension to stay bounded in the continuum limit.

In perturbative continuum RS theory, it has become popular to study

the effects of an extrinsic curvature term in the action.[10] This term can be chosen to be perturbatively renormalizable, and its coupling constant is asymptotically free [i.e., the coupling constant in front of $A_{\text{ext}}(S)$ grows as one moves to larger distance scales]. A nonperturbative analysis of the effects of this term in triangulated RS models has been carried out in ref. 4. For lattice RS models, it will be studied in this paper.

We choose the action to have the form

$$\begin{aligned} A_{\text{int}}(S) &= \beta_2 \,\text{area}(S) \\ A_{\text{ext}}(S) &= \beta_1 \,|E(S)| + \frac{\beta_0}{2} \sum_{j \in S} n_S(j) \end{aligned} \tag{1.4}$$

where $E(S)$ is the set of "edges" of S (see Section 2) and $|E(S)|$ is the total length of $E(S)$, j is a vertex of S, and $n_S(j)$ is the number of edges of S meeting at the vertex j. In three dimensions, the second term in $A_{\text{ext}}(S)$ is essentially equivalent to $\Lambda(S)$, but in higher dimensions this is not so. The choice of the action (1.2), (1.4) can be well motivated by studying its flow under renormalization group transformations (Section 4) and by steric considerations and rigidity against bending.

One of the main purposes of this paper is to map out the *phase diagrams* of some lattice RS models in the three-dimensional half-space of coupling constants $\{\underline{\beta} = (\beta_0, \beta_1, \beta_2)\colon \beta_2 > 0\}$, to study different critical regimes, and to investigate the flow of the action (1.4) under renormalization transformations. Briefly, our *main results* for the *planar* RS model (see Section 2) are as follows:

1. There exists a convex set $\mathscr{B} \subset \{\underline{\beta}\colon \beta_2 > 0\}$ such that the planar RS model is well defined and has strictly positive mass and surface tension in the interior of $\mathscr{B}$. (Some of these results extend to self-avoiding RS models.)

2. The surface tension tends to zero as $\partial\mathscr{B}$ is approached along a trajectory for which $\beta_0 + \beta_1 \to \infty$.

3. The mass tends to 0 as $\partial\mathscr{B}$ is approached with β_1 sufficiently small and $\beta_2 \to 0$ (hence $\beta_0 \to \infty$).

4. The effective action on an arbitrary length scale has approximately the form (1.4), up to nonlocal, irrelevant terms.

It follows that a sensible scaling limit of the planar RS model, where mass *and* string tension scale to zero, can be reached only by letting $\underline{\beta}$ tend to $\partial\mathscr{B}$, with β_1 sufficiently small and $\beta_0 \to \infty$. However, the scaling limit is most likely *not* Euclidean invariant (so that such models are of little interest in connection with string theory, although they are "unitary" in the sense established in ref. 5; but see also ref. 4). The curve $\mathscr{C}$ in $\partial\mathscr{B}$ separating the part of $\partial\mathscr{B}$ where the mass vanishes from the part where it is positive

is of physical significance: $\mathscr{C}$ separates two regimes where typical random surfaces exhibit very different behavior. Above $\mathscr{C}$, i.e., where the mass is positive, random surfaces are very flat, and the normals to such surfaces exhibit long-range order. Below $\mathscr{C}$, typical random surfaces are crumpled, behaving like branched polymers. We shall see that, below $\mathscr{C}$, branched-polymer behavior most likely disappears only in the limit where $\beta_0 \to \infty$.

Some of these results extend to models of self-avoiding random surfaces. We also conjecture the existence of a phase boundary $\mathscr{C}'$, similar to the curve $\mathscr{C}$ described above, for *triangulated* RS models. (The relation of the lattice planar RS models to the triangulated RS models is expected to be somewhat analogous to the relation between the two-dimensional Ising and two-dimensional XY or Heisenberg models.) A renormalization group analysis for triangulated RS models has, however, not been performed. It would be considerably more difficult, just like a rigorous version of the droplet argument for a transition in a classical spin system with continuous symmetry would be much more intricate than the Peierls argument in the Ising model.

The most important new technical tool introduced in this paper is a real-space renormalization-group method for lattice surfaces. The basic idea is to regard a surface as a hierarchical structure of excitations of an underlying surface that consists of flat pieces. We inductively sum over all levels of this hierarchy and discover that an action of the form (1.4) reproduces itself with renormalized coupling constants up to terms which are expected to be irrelevant. The details of this procedure are described in Sections 4 and 5.

In the next section we introduce the ensembles of surfaces we wish to consider and various definitions, while in Section 3 we give a precise statement of our results. Sections 5 and 8 contain a technical proof of the convergence of the loop functions for arbitrarily small values of β_2, provided β_0 and/or β_1 is sufficiently large. In Section 6 we prove (under some assumptions) that the mass is positive if β_1 is large. The arguments in Sections 4–6 are for a restricted class of surfaces and in Section 7 we show how to extend them.

Section 8 contains a combinatorial estimate of the number of surfaces with a given area and a given number of corners. This estimate gives the strongest bounds on the critical surface $\partial\mathscr{B}$ and does not rely on the simplifying assumptions of Sections 4–6.

The main technical ideas of this paper are described in Sections 4 and 5. Sections 6 and 7 contain auxiliary material, while Section 8 contains a fairly remarkable, but complicated combinatorial construction, whose purpose it is to improve the results obtained by the methods of Sections 4 and 5.

2. DEFINITIONS OF RANDOM SURFACE ENSEMBLES

A model describing the statistical mechanics of a single lattice random surface is defined as follows.

(I) We choose an infinite, regular lattice, in this paper $= \mathbb{Z}^d$, and specify a countable ensemble $\mathscr{E}$ of lattice random surfaces. More precisely, $\mathscr{E}$ is a countable family of two-dimensional, connected cell complexes, built by gluing together elementary lattice 2-cells, in our case plaquettes, along common edges. Elements of $\mathscr{E}$, i.e., random "surfaces," are denoted by S.

(II) To every $S \in \mathscr{E}$ one assigns a positive statistical weight $\rho_{\underline{\beta}}(S)$, where $\underline{\beta}$ is a family of parameters ("coupling constants") on which the weights $\rho_{\underline{\beta}}(S)$ depend. Usually, $\rho_{\underline{\beta}}(\cdot)$ is defined to be invariant under Euclidean motions which leave the lattice invariant; (lattice translations and lattice rotations). The weights $\rho_{\underline{\beta}}(\cdot)$ define a measure on $\mathscr{E}$, $P_{\underline{\beta}}$, which usually is infinite. Information about typical behavior of random "surfaces" in $\mathscr{E}$ with respect to $P_{\underline{\beta}}$ is retrieved by calculating the $P_{\underline{\beta}}$ measure of appropriate subsets of $\mathscr{E}$. Examples will be given below.

Examples of ensembles $\mathscr{E}$ are:

(i) Graphs of integer-valued functions on connected subsets Λ of the dual of the two-dimensional lattice $\mathbb{Z}^2$ viewed as two-dimensional, connected surfaces embedded in $\mathbb{Z}^3$. This ensemble is useful for the description of surfaces of crystals or of interfaces in the solid-on-solid approximation. It has been used widely in studies of the roughening and wetting transitions[11,12] (for recent rigorous results see ref. 13).

(ii) Connected clusters of (singly) occupied plaquettes in $\mathbb{Z}^d$. They are studied in connection with Bernoulli plaquette percolation and q-states Potts gauge theories in the Fortuin–Kasteleyn representation.[22]

(iii) Connected clusters of multiply occupied plaquettes in $\mathbb{Z}^d$ with the property that every link (= bond = pair of neighboring sites) of every plaquette of such a cluster is also a link of at least one further plaquette of the cluster. Two plaquettes in such a cluster may, geometrically, represent the same plaquette of $\mathbb{Z}^d$, which is then at least doubly occupied. Thus, a cluster in this ensemble is an assignment of a nonnegative integer $n(p)$ to each plaquette in $\mathbb{Z}^d$, with the property that if $n(p) > 0$, then either $n(p) \geqslant 2$, or, if $n(p) = 1$, there is a plaquette $p' \neq p$, with $n(p') \geqslant 1$, sharing a link with p. We also require that

$$\sum_{p \subset \mathbb{Z}^d} n(p) < \infty \tag{2.1}$$

for all clusters in the ensemble.

This ensemble is of interest in statistical theories of certain ternary mixtures and foam structures.

(iv) Self-advoiding random surfaces. By a self-avoiding random surface we mean an assignment of a number $n(p) \in \{0, 1\}$ to every plaquette p of $\mathbb{Z}^d$ such that either $n(p_0) = 1$ for some $p_0 \subset \mathbb{Z}^d$ and $n(p) = 0$ for all $p \neq p_0$ (surface consisting of a single plaquette, p_0), or, if $n(p) = 1$, then there exists a plaquette $p' \neq p$, with $n(p') = 1$, sharing a link with p. Moreover, for any link $b \subset \mathbb{Z}^d$ (i.e., a nearest-neighbor pair of sites in $\mathbb{Z}^d$),

$$\sum_{p: \partial p \ni b} n(p) \leqslant 2 \tag{2.2}$$

where ∂p denotes the set of boundary links of the plaquette p. The boundary of a self-avoiding random surface S with occupation numbers $\{n_S(p)\}_{p \subset \mathbb{Z}^d}$ is defined as

$$\partial S = \left\{ b \subset \mathbb{Z}^d : \sum_{p: \partial p \ni b} n_S(p) = 1 \right\} \tag{2.3}$$

(v) Connected orientable random surfaces of genus 0, which are called "planar RS."[6] A planar random surface (PRS) S is defined as follows: It is a two dimensional, connected, orientable cell complex $\sum$, where the elementary 2-cells are squares, together with an imbedding

$$i: \sum \to \mathbb{Z}^d \tag{2.4}$$

such that the elementary two cells in $\sum$ *are mapped into plaquettes in* $\mathbb{Z}^d$, the elementary 1-cells are mapped into links in $\mathbb{Z}^d$, and the elementary 0-cells are mapped into points in $\mathbb{Z}^d$. Furthermore, the cell complex $\sum$ is required to have the topology of a sphere or that of a sphere with a number of holes. The statement that $\sum$ is connected means that for any pair (p, p') of 2-cells in $\sum$, there is a sequence of 2-cells in $\sum$, $p_1, \ldots, p_n$ such that p_1 and p share a 1-cell in their boundaries and the same applies to the pairs (p_n, p') and (p_i, p_{i+1}), $i = 1, \ldots, n-1$. The boundary of S, denoted ∂S, is the image of the boundary of $\sum$ under the mapping i, i.e., ∂S consists of those 1-cells in S that belong to the boundary of a single 2-cell.

This definition of planar random surfaces may look complicated. In more cavalier jargon, it goes as follows: A connected planar RS is a collection of copies of plaquettes in $\mathbb{Z}^d$ whose links are glued together pairwise in such a way that the resulting complex is a connected, orientable surface with the topology of a sphere, or of a sphere with holes.

We now proceed to discuss examples of weights $\rho_{\beta}(S)$ of "surfaces" in one of the ensembles specified above. For this purpose we introduce some further notions and notation.

Let $P(S)$ denote the set of all two-cells (copies of plaquettes in $\mathbb{Z}^d$) which belong to a random surface S, $L(S)$ the set of all one-cells (copies of links in $\mathbb{Z}^d$) belonging to S, and $V(S)$ the set of all vertices (copies of sites of $\mathbb{Z}^d$) in S. Every element in $L(S)$ belongs to the boundary of one or more elements of $P(S)$, and every element in $V(S)$ belongs to the boundary of two or more elements of $L(S)$. Two adjacent links of a plaquette in $P(S)$ share one common vertex in $V(S)$. If two plaquettes of $P(S)$ are glued together along an edge $\tilde{b}$, then corresponding endpoints of the copies of $\tilde{b}$ in the boundaries of those two plaquettes are identified as elements of $V(S)$, under the gluing. [However, if $\tilde{b}_1$ and $\tilde{b}_2$ are two *distinct* elements of $L(S)$ which are copies of the *same* link in $\mathbb{Z}^d$, their corresponding endpoints may (or may not) be distinct elements of $V(S)$.]

The notions which we shall now discuss have a precise mathematical meaning for ensembles of types (i) (solid-on-solid surfaces), (iv) (self-avoiding random surfaces), and (v) (planar random surfaces), but require some modifications for the other ensembles, (ii) and (iii), discussed above. Given a vertex $j \in V(S)$, let $\partial^* j$ be the collection of all those one-cells which are distinct elements of $L(S)$ and contain j in their boundary. We define the *intrinsic curvature* (deficiency index) $\delta(j)$ of a random surface S at a vertex $j \in V(S)$ by

$$\delta(j) = 4 - \#\partial^* j \tag{2.5}$$

where $\#A$ denotes the cardinality (number of elements) of a set A. The number 4 on the rhs of (1.5) is due to the fact that in a two-dimensional coordinate plane of $\mathbb{Z}^d$ at most four plaquettes can share a common vertex. The Euler characteristic of S is given by

$$\begin{aligned} \chi(S) &= \#P(S) - \#L(S) + \#V(S) \\ &= 2 - 2h(S) - c(\partial S) \end{aligned} \tag{2.6}$$

where $h(S)$ is the number of handles (or genus) of S, and $c(\partial S)$ is the number of connected components of ∂S. [For planar surfaces, $h(S) = 0$.] It is not hard to prove[5]

$$\sum_{j \in V(S)} \delta_S(j) = 4\chi(S) + |\partial S| \tag{2.7}$$

where $|\partial S|$ is the number of elements in $L(S)$ that belong to ∂S, i.e., the length of ∂S. This is the Gauss–Bonnet identity for lattice surfaces.

The cardinalities of $P(S)$, $L(S)$, and $V(S)$ and the curvature $\delta_j(S)$ of a random surface S at a site $j \in V(S)$ are *intrinsic* notions, defined for abstract, two-dimensional cell complexes without reference to how they lie

in the lattice $\mathbb{Z}^d$. Next, we define some *extrinsic* notions referring to how S is arranged inside $\mathbb{Z}^d$.

Let $\tilde{b} \in L(S)$. Since we are working with self-avoiding or planar RS, $\tilde{b}$ belongs to the boundary of only one or two plaquettes in $P(S)$. If it belongs to only one plaquette of $P(S)$, it is in the boundary ∂S of S. If it belongs to two plaquettes $\tilde{p}$ and $\tilde{p}'$ of $P(S)$, then we may distinguish three cases:

(a) $\tilde{p}$ and $\tilde{p}'$ are copies of two distinct plaquettes p and p' which lie in a *common*, two-dimensional coordinate plane of $\mathbb{Z}^d$.

(b) $\tilde{p}$ and $\tilde{p}'$ are copies of two distinct plaquettes p and $\tilde{p}'$ which lie in *different*, two-dimensional coordinate planes of $\mathbb{Z}^d$.

(c) $\tilde{p}$ and $\tilde{p}'$ are oppositely oriented copies of the same plaquette of $\mathbb{Z}^d$.

In cases (b) and (c) we call $\tilde{b}$ an *edge link* of S. The edge links of S make up a graph in $\mathbb{Z}^d$ which we denote by $E(S)$. More precisely, $E(S)$ is the collection of all edge links in S together with the connectedness properties they inherit from S, i.e., two edge links in $E(S)$ are connected if and only if they share a vertex in S. It is also convenient to adopt the convention that boundary links are edge links.

A vertex $j \in V(S)$ is called a *corner* if more than two edge links meet at j or if two edge links meet at j at an angle different from π. We leave it as an easy exercise for the reader to check that $\partial^* j$ never contains a single edge link.

An *edge* of S is a maximal connected subset of $E(S)$ containing no corners in its interior. If follows that an edge is contained in a lattice line and terminates at a corner.

An *edge network* is a maximal connected subset of $E(S)$. We frequently abbreviate edge network by *network*.

We denote by $C(S)$ the collection of all corners in S, including the corners of ∂S. If $j \in C(S)$, there meet $n_S(j)$ edges at j, where $n_S(j) \geqslant 2$ is called the *order* of the corner j.

We are now prepared to define the statistical weights of the random surfaces in the ensembles $\mathscr{E}$ described in (i), (iv), and (v) above. We introduce the following notation:

$$|S| = \# P(S) = \text{area of } S \tag{2.8}$$

$$|E(S)| = \# E(S) \tag{2.9}$$

$$N(S) = \frac{1}{2} \sum_{j \in C(S)} n_S(j) \tag{2.10}$$

Thus, $|E(S)|$ is the number of edge links, and $N(S)$ is the number of edges or alternatively the number of corners in S counted with multiplicity. We shall also use the notation $N(E)$ for $N(S)$ if $E = E(S)$.

Let $\underline{\beta} = (\beta_0, \beta_1, \beta_2) \in \mathbb{R}^3$. If $S \in \mathscr{E}$, we define the action (or bare action) $A_{\underline{\beta}}(S)$ of S by

$$A_{\underline{\beta}}(S) = \beta_2 |S| + \beta_1 |E(S)| + \beta_0 N(S) \tag{2.11}$$

The coupling constant β_2 is always nonnegative, while β_0 and β_1 may become negative; see Theorem 1 below. The statistical weight of S is given by

$$\rho_{\underline{\beta}}(S) = \exp[-A_{\underline{\beta}}(S)] \tag{2.12}$$

Our purpose is to study the physical properties of the RS model with the ensemble $\mathscr{E}$ specified in (v), with the statistical weight defined in (2.8)–(2.12). Some of our methods also apply to the other ensembles described above. This will be briefly described in Section 7.

Rather than working with the ensemble (v) of planar random surfaces, it is convenient, for the purpose of explaining the renormalization group method, to introduce a somewhat smaller ensemble of surfaces, which we now describe. Let S be a planar random surface. S is said to be *flat* if it contains no edge links in its interior. A *face* of a planar random surface is a maximal connected flat subsurface. If F is a face of S, it follows that $\partial F \subseteq E(S)$ and F lies in a coordinate plane.

It is not hard to see that a face of a PRS has the topology of a disc or that of a disc with a number of holes. The boundary of any face is a union of loops. A face of a surface is *simple* if all its boundary components are simple loops, i.e., self-avoiding. A surface is said to be *simple* if and only if all its faces are simple.

We define ensemble (vi) of *simple PRS* to consist of all the simple surfaces in ensemble (v).

All the results described in the next section are valid for ensemble (v), while the proofs given in Sections 4–6 are for ensemble (vi). In Section 7 we describe the modifications of the RG method required for nonsimple surfaces.

We conclude this section with a comment on the geometrical meaning of the term $N(S)$. (This is of interest in connection with the proposal made in refs. 5 and 6.)

It is not hard to see that

$$2 \leqslant n(j) \leqslant 4 + |\delta(j)| \tag{2.13}$$

But while in *three* dimensions, i.e., for surfaces embedded in $\mathbf{Z}^3$, $|\delta(j)|$ is *positive* for every corner $j \in C(S)$, this is *not* so in *four or more* dimensions:

For $d \geqslant 4$, there are, for example, corners j of order $n(j) = 4$ for which $\delta(j) = 0$. [Consider the origin in $\mathbb{Z}^d$, $d \geqslant 4$, and glue together the plaquettes (12), (13), (24), and (34) attached at the origin. Then $n(0) = 4$, while $\delta(0) = 0$.] Therefore, a term $\beta_0' \Delta(S)$, with $\Delta(S) = \sum_{j \in V(S)} |\delta(j)|$ (ree ref. 5), in the action is much less coercive than then term $\beta_0 N(S)$ chosen in (2.11) and is presumably *not* sufficient to avoid a "collapse to triviality" of the model if $d \geqslant 4$. (This is in contrast to the situation met in triangulated RS models.[3,4]) However, in *three dimensions*, the terms $\beta_0' \Delta(S)$ and $\beta_0 N(S)$ play equivalent roles. The results proven in this paper for the action defined in (2.11) are also valid for an action where $\beta_0 N(S)$ is replaced by $\beta_0' \Delta(S)$ if $d = 3$. In order not to have to distinguish between $d = 3$ and $d \geqslant 4$, we shall work with the action (2.11), but we emphasize that, in the physical dimension $d = 3$, defect angles are covenient and physically plausible parameters.

3. RESULTS

In order to study physical properties of an RS model one must ask appropriate questions about the behavior of typical random surfaces. The most elementary questions are of combinatorial nature.

(a) Consider, for example, the quantity

$$n_{\mathscr{L}}(A) = \#\{S \in \mathscr{E};\ \ |S| = A,\ \partial S = \mathscr{L}\} \tag{3.1}$$

where $\mathscr{L} = \mathscr{L}_1 \cup \cdots \cup \mathscr{L}_n$ is a finite union of lattice loops. For the ensembles (ii)–(vi) described in Section 2, one may show that

$$n_{\mathscr{L}}(A) = \exp[\beta^{(2)} A + o(A)] \tag{3.2}$$

for some $\beta^{(2)}$ *independent* of $\mathscr{L}$. Numerical results[14] indicate that, for planar random surfaces and $\mathscr{L} = \partial p$, the boundary of a plaquette,

$$n_{\partial p}(A) \underset{A \to \infty}{\sim} A^{\varepsilon_0} e^{\beta^{(2)} A} \tag{3.3}$$

with $\varepsilon_0 \simeq -1.5$ (within 10%, for $d = 3$). It has been shown in ref. 6 that if $\varepsilon_0 \geqslant -2$, then $\varepsilon_0 = -1.5$ for planar random surfaces in arbitrary dimension d. [Throughout this article, $f(A) \sim_{A \to \infty} g(A)$ means that there are finite, positive constants c_1 and c_2 such that, for A large enough, $c_1 g(A) \leqslant f(A) \leqslant c_2 g(A)$.] The equation $\varepsilon_0 = -1.5$ plays a key role in our analysis of the collapse to triviality, for $\beta_0 = \beta_1 = 0$, presented in ref. 6. The exponent ε_0 is not known exactly for any of the other ensembles defined in Section 2.

(b) In this paper we shall require bounds on the quantities

$$n_{\mathscr{L}}(A, E, N) = \#\{S \in \mathscr{E}: \quad |S| = A, |E(S)| = E, |N(S)| \leqslant N, \partial S = \mathscr{L}\} \tag{3.4}$$

for the ensemble of planar RS.

We shall prove that, for every $\varepsilon > 0$, there exists a convex region $\mathscr{R}_\varepsilon$ in the (β^0, β^1) plane containing the intervals

$$I_0 = \{(\beta^0, \beta^1): \quad \beta^1 = 0, B^0(\varepsilon) < \beta^0 < \infty\}$$
$$I_1 = \{(\beta^0, \beta^1): \quad \beta^0 = 0, B^1(\varepsilon) < \beta^1 < \infty\}$$

for some finite constants $B^0(\varepsilon)$ and $B^1(\varepsilon)$, such that for $(\beta^0, \beta^1) \in \mathscr{R}_\varepsilon$,

$$n_{\mathscr{L}}(A, E, N) \leqslant e^{\varepsilon A + \beta^1 E + \beta^0 N} \tag{3.5}$$

This combinatorial result is proven in Section 8.

(c) *Loop correlations* are defined by

$$G_{\underline{\beta}}(\mathscr{L}_1, ..., \mathscr{L}_n) = \sum_{\substack{S \in \mathscr{E} \\ \partial S = \mathscr{L}_1 \cup \cdots \cup \mathscr{L}_n}} \exp[-A_{\underline{\beta}}(S)] \tag{3.6}$$

where $\mathscr{L}_1, ..., \mathscr{L}_n$ are loops in $\mathbb{Z}^d$.

Let $\gamma_{L,T}$ be a rectangular loop in a lattice plane with sides of length L and T. We define the "string potential" $V_{\underline{\beta}}(L)$ by

$$V_{\underline{\beta}}(L) = \lim_{T \to \infty} -\frac{1}{T} \log G_{\underline{\beta}}(\gamma_{L,T}) \tag{3.7}$$

and the *surface tension* $\tau(\underline{\beta})$ by

$$\tau(\underline{\beta}) = \lim_{L \to \infty} \frac{1}{L} V_{\underline{\beta}}(L) \tag{3.8}$$

An inverse "linear extension" or *mass* $m(\underline{\beta})$ is given by

$$m(\underline{\beta}) = \lim_{a \to \infty} -\frac{1}{a} \log G_{\underline{\beta}}(\partial p, \partial p_a) \tag{3.9}$$

where ∂p_a is a copy of ∂p translated by a lattice units in a lattice direction. It is easy to see that $m(\underline{\beta})^{-1}$ is a measure for the *mean linear extension* of random surfaces (or random "bubbles"). The existence of the limits in (3.7)–(3.9) for planar random surfaces is proven in ref. 16 for a special model ($\beta_0 = \beta_1 = 0$), but the arguments in ref. 16 extend to the more general models studied in this paper.

We define a *susceptibility* (or better: *specific heat* or *mean area*) $\chi(\underline{\beta})$ by setting (see ref. 6 for details)

$$\chi(\underline{\beta}) = \sum_{p' \in \mathbb{Z}^d} G_{\underline{\beta}}(\partial p, \partial p')$$

$$\sim \sum_{\substack{S \in \mathscr{E} \\ \partial S = \partial p}} |S| \exp[-A_{\underline{\beta}}(S)] \tag{3.10}$$

All the results summarized below are results for the *simple planar random surface model whose action is given by* (2.11). Many of our results extend to self-avoiding random surfaces (SAS), some also to foam models [ensemble (iii)]; see Section 7.

We now state our basic results.

Theorem 1. There exists a convex, open region $\mathscr{B} \subset \{\underline{\beta}: \beta_2 > 0\}$ such that, for all $\underline{\beta} \in \mathscr{B}$, the loop correlations $G_{\underline{\beta}}(\mathscr{L}_1, ..., \mathscr{L}_n)$ defined in (3.6) are well-defined, finite quantities.

Let $(\beta_0, \beta_1, \beta_2(\beta_0, \beta_1)) \in \partial\mathscr{B}$. Then $\beta_2(\beta_0, \beta_1)$ is a positive, monotone decreasing, convex function of β_0 and β_1, and

$$\beta_2(\beta_0, \beta_1) \to 0$$

if β_0 or β_1 tends to ∞ with $(\beta_0, \beta_1, \beta_2) \in \partial\mathscr{B}$.

This theorem—which extends to some SAS models—is obviously closely related to the bound (3.5), and in fact follows from it. Its proof is also a consequence of the renormalization group analysis performed in Sections 4 and 5.

Our next result says that, for the PRS model, critical behavior can only occur on $\partial\mathscr{B}$.

Theorem 2. For $\underline{\beta} \in \mathscr{B}$,

$$\tau(\underline{\beta}) > 0, \qquad m(\underline{\beta}) > 0, \qquad \chi(\underline{\beta}) < \infty$$

and all the moments

$$M_n(\underline{\beta}) = \sum_{\substack{S \in \mathscr{E} \\ \partial S = \partial p}} |S|^n \exp[-A_{\underline{\beta}}(S)], \qquad n = 0, 1, 2, ... \tag{3.11}$$

are finite.

Except possibly for the positivity of τ, this result also holds for SAS models. Precise statements and proofs have been given in ref. 16 for the case $\beta_0 = \beta_1 = 0$. Those arguments extend without difficulty to the present situation. Some of these extensions are described in ref. 8.

(d) We consider the notion of *critical exponents*. In order to discuss critical properties and continuum limits of the PRS models, we must, according to Theorem 2, study the behavior of these models as $\underline{\beta}$ approaches $\partial\mathscr{B}$. We say that a *trajectory* $\underline{\beta}(t) \subset \mathscr{B}$, with $t \in (0, 1]$ and $\underline{\beta}(0) \in \partial\mathscr{B}$, approaches $\partial\mathscr{B}$ *transversally* iff

$$|\dot{\beta}_i(t)| \leqslant K\dot{\beta}_2(t), \qquad i = 0, 1 \tag{3.12}$$

and

$$K' \leqslant \dot{\beta}_2(t) \leqslant K'' \tag{3.13}$$

for some finite constants K, K', and K'' independent of t. [Here $\dot{\beta}(t) \equiv (d/dt)\,\beta(t)$.]

Our next result is the following theorem.

Theorem 3. (i) For $\underline{\beta} \in \partial\mathscr{B}$ with β_0, β_1 finite

$$G_{\underline{\beta}}(\mathscr{L}) < \infty$$

for an arbitrary loop $\mathscr{L}$ (e.g., $\mathscr{L} = \partial p$) in $\mathbb{Z}^d$.

(ii) If $\underline{\beta}(t)$ approaches $\partial\mathscr{B}$ transversally, then

$$\chi(\underline{\beta}(t)) \leqslant \text{const}/\sqrt{t}$$

Theorem 3, part (i), follows by the arguments used in the proof of Lemma 2.1 in ref. 6, and part (ii) follows by the arguments developed in Section 3 of ref. 5. It is worthwhile remarking that (i) follows from (ii) (see also ref. 15).

In order to study the critical behavior of the PRS model in *bounded* regions of $\partial\mathscr{B}$, it is convenient to introduce the following quantity:

$$n_{\mathscr{L}}(A; \beta_0, \beta_1) = \sum_{\substack{S \in \mathscr{S} \\ \partial S = \mathscr{L} \\ |S| = A}} \exp[-\beta_1 |E(S)| - \beta_0 N(S)] \tag{3.14}$$

By Theorem 1 there exists a finite, positive constant $\beta^{(2)} = \beta^{(2)}(\beta_0, \beta_1)$ such that

$$(\beta_0, \beta_1, \beta^{(2)}) \in \partial\mathscr{B}$$

and

$$n_{\mathscr{L}}(A; \beta_0, \beta_1) \underset{A \to \infty}{\sim} \exp[\beta^{(2)} A + o(A)] \tag{3.15}$$

For the details of a subadditivity argument of this type, see ref. 16.

Under appropriate assumptions on $n_{\mathscr{L}}(A; \beta_0, \beta_1)$, one can define an exponent $\varepsilon = \varepsilon(\beta_0, \beta_1)$ by

$$n_{\mathscr{L}}(A; \beta_0, \beta_1) \underset{A \to \infty}{\sim} A^{\varepsilon} e^{\beta^{(2)} A} \tag{3.16}$$

This exponent is the analogue of ε_0 and plays an important role in the analysis of the models with β_0 and β_1 bounded.

We define critical exponents γ, ν, μ, and η as follows:

$$\chi(\underline{\beta}(t)) \sim t^{-\gamma}, \qquad m(\underline{\beta}(t)) \sim t^{\nu} \tag{3.17}$$

$$d\tau(\underline{\beta}(t))/dt \sim t^{\mu - 1} \tag{3.18}$$

and

$$G_{\underline{\beta}(t)}(\partial p, \partial p_a) \sim a^{-(d-2+\eta)} \tag{3.19}$$

for $1 \ll a \ll m(\underline{\beta}(t))$, $t \to 0$.

If these exponents exist, then, under some standard scaling hypotheses, one can show that

$$\gamma = \nu(2 - \eta), \qquad \nu = \tfrac{1}{2}\mu \tag{3.20}$$

Moreover, if $\underline{\beta}(t)$ approaches $\partial\mathscr{B}$ transversally, then

$$\gamma = 2 + \varepsilon \tag{3.21}$$

See refs. 6 and 7 for proofs.

Let $\mathscr{E}_0$ be the subensemble of the ensemble $\mathscr{E}$ of planar random surfaces, consisting of planar surfaces which do not contain any loops of length 2 made of two distinct copies $\tilde{b}$, $\tilde{b}'$ of a link $b \subset \mathbb{Z}^d$ glued at the endpoints (i.e., with endpoints identified). The results discussed above for the simple PRS model extend to the model with ensemble $\mathscr{E}_0$, called the PRS_0 model. In particular, the loop correlations of the PRS_0 model are finite in a convex, open region $\mathscr{B}_0 \supset \mathscr{B}$. Given β_0 and β_1, let $\beta_0^{(2)}(\beta_0, \beta_1)$ be such that $(\beta_0, \beta_1, \beta_0^{(2)}(\beta_0, \beta_1)) \in \partial\mathscr{B}_0$. Clearly $\beta_0^{(2)} \leqslant \beta^{(2)}$, and it is easy to see that $\beta_0^{(2)} > \frac{1}{4}\beta^{(2)}$.[6] Quantities referring to the PRS_0 model receive a subscript zero in the following. It is plausible that the PRS and the PRS_0 models belong to the same universality class; see ref. 6. The proof of the following theorem follows from the methods developed in refs. 6, 8, and 9 (for a review see also ref. 15).

Theorem 4. Suppose that the susceptibilities $\chi(\underline{\beta}(t))$ and $\chi_0(\underline{\beta}_0(t))$ diverge as $\underline{\beta}(t)$ and $\underline{\beta}_0(t)$ approach $\partial\mathscr{B}$ and $\partial\mathscr{B}_0$ transversally, respectively. Then

$$\varepsilon = -1.5, \qquad \gamma = \tfrac{1}{2} = \mu, \qquad \nu = \tfrac{1}{4}, \qquad \eta = 0 \tag{3.22}$$

$$\tau(\underline{\beta}(0)) > 0 \tag{3.23}$$

and the continuum ($\equiv$ scaling) limit of $G_{\underline{\beta}}(\partial p, \partial p')$ is the propagator of a free, scalar field.

The exponents (3.22) are the mean-field exponents ($d \to \infty$) of *branched polymers*. The hypotheses of Theorem 4 have been tested numerically in ref. 14 for $\beta_0 = \beta_1 = 0$ (see also ref. 17 for related numerical work).

Next, we show that on trajectories $\underline{\beta}(t)$ which tend to ∞ as $t \to 0$, a different critical regime can be reached.

Theorem 5. Let $\underline{\beta}(t) \in \mathscr{B}$, $t \in (0, 1]$ be a path in $\mathscr{B}$ such that $\beta_2(t) \downarrow 0$ as $t \downarrow 0$. Then

$$\lim_{t \to 0} \tau(\underline{\beta}(t)) = 0 \tag{3.24}$$

If $\beta_1(t) = 0$ for $t \in [0, 1)$, then

$$\lim_{t \to 0} m(\underline{\beta}(t)) = 0 \tag{3.25}$$

The existence of paths in $\mathscr{B}$ with $\beta_2(t) \downarrow 0$ follows immediately from Theorem 1. In order to prove (3.24), let S_L be the unique flat surface with boundary $\gamma_{L,L}$. Clearly, $A_{\underline{\beta}}(S_L) = \beta_2 L^2$, so

$$G_{\underline{\beta}}(\gamma_{L,L}) > e^{-\beta_2 L^2} \tag{3.26}$$

which implies (3.24). Similarly, considering a long, thin cylinder,

$$G_{\underline{\beta}}(\partial p, \partial p_a) > \text{const} \cdot e^{-2\beta_2 a - 2\beta_1 a}$$

from which (3.25) follows.

The above theorem tells us that if $\beta_0 \to \infty$, $\beta_1 \to 0$, and $\beta_2 \to 0$ a critical regime is reached where both the mass $m(\underline{\beta})$ *and* the surface tension $\tau(\underline{\beta})$ tend to 0. Our estimates are, however, not sharp enough to show that $m^2(\underline{\beta})/\tau(\underline{\beta})$ tends to a preassigned, finite value as $\underline{\beta}$ tends to $\partial\mathscr{B}$ in a suitable way, with $\beta_0 \to \infty$ and $\beta_2 \to 0$.

The next result will show that it is likely that $\tau(\underline{\beta})/m^2(\underline{\beta})$ can be made to reach a finite value, although possibly $= 0$, if $\underline{\beta}$ tends to $\partial\mathscr{B}$ in a suitable way, with $\beta_0 \to \infty$ and $\beta_2 \to 0$. We call this result a quasitheorem, since we only have a partial proof for it (see Sections 4 and 6).

Quasitheorem 6. There exists a curve $\mathscr{C} \subset \partial\mathscr{B} \cap \{\underline{\beta}\colon \beta_1 \leqslant B_1\}$, where B_1 is a finite constant such that if $\underline{\beta}(t)$ is a trajectory in $\mathscr{B}$ reaching $\partial\mathscr{B}$ above the curve $\mathscr{C}$, as $t \searrow 0$, then

$$\lim_{t \to 0} m(\underline{\beta}(t)) = m(\underline{\beta}(0)) > 0 \tag{3.27}$$

In particular, if $\underline{\beta}$ with $\beta_1 \to \infty$, so that $\beta_2 \to 0$, then

$$\tau(\underline{\beta}) \to 0 \qquad \text{while} \quad m(\underline{\beta}) \to \infty$$

It should be quite straightforward to test the predictions of this quasitheorem numerically.

It can be argued that the part $\partial\mathscr{B}_+$ of $\partial\mathscr{B}$ above the curve $\mathscr{C}$ contains a second phase boundary $\mathscr{C}'$ such that typical surfaces contributing to $G_{\underline{\beta}}(\gamma_{L,L})$ as L becomes large are *rigid* if $\underline{\beta}$ tends to $\partial\mathscr{B}$ *above* $\mathscr{C}'$, while they are *rough* if $\underline{\beta}$ tends to $\partial\mathscr{B}$ *between* $\mathscr{C}$ and $\mathscr{C}'$ (i.e., there is a *roughening transition* in this model).

(e) For the sake of completeness, we briefly recall the definition of the *radius of gyration* of random surfaces and its critical exponent, which is related to their Hausdorff dimension, the *recurrence exponent*, and a correlation which is a measure for *orientational order*. We define

$$\bar{d}(S) = \max\{\operatorname{dist}(p, p')\colon \partial S = \partial p,\ p' \in P(S)\}$$

where $\operatorname{dist}(A, B)$ is the minimal Euclidean distance between elements of a set $A \subset \mathbb{R}^d$ and of a set $B \subset \mathbb{R}^d$. We define quantities

$$R_n(\underline{\beta}) = \sum_{\substack{S \in \mathscr{E} \\ \partial S = \partial p}} \bar{d}(S)^n \exp[-A_{\underline{\beta}}(S)] \tag{3.28}$$

The definition of *Hausdorff dimension* δ of a random surface S (see refs. 6, 7, and 15) suggests that

$$\bar{d}(S) \underset{|S| \to \infty}{\sim} |S|^{1/\delta} \tag{3.29}$$

Hence,

$$R_n(\underline{\beta}) \sim \sum_{\substack{S \in \mathscr{E} \\ \partial S = \partial p}} |S|^{n/\delta} \exp[-A_{\underline{\beta}}(S)] \tag{3.30}$$

By (3.16)

$$R_n(\underline{\beta}) \sim \sum_{A=1}^{\infty} A^{\varepsilon + n/\delta} e^{-\text{const} \cdot (\beta_2 - \beta^{(2)})A} \tag{3.31}$$

Hence the critical exponent ρ_n of $R_n(\underline{\beta})$ is given by

$$\rho_n = 1 + \varepsilon + n/\delta \tag{3.32}$$

The mean radius of gyration $r(\underline{\beta})$ is defined by

$$r(\underline{\beta}) \sim R_{n+1}(\underline{\beta})/R_n(\underline{\beta}) \sim (\beta_2 - \beta^{(2)})^{-1/\delta} \tag{3.33}$$

i.e., its critical exponent is $1/\delta$. It has been argued in refs. 6 and 7 that

$$1/\delta = \nu \tag{3.34}$$

Hence, in the branched polymer regime,

$$\delta = 4 \tag{3.35}$$

by (3.22). By calculating the intersection probability of two independent, noninteracting branched polymers in $\mathbb{Z}^d$ and showing that it tends to zero for $d \geqslant 8$ while it remains positive in $d < 8$, it has been proven in ref. 18 that (3.35) is indeed correct within the branched polymer regime.

If $\underline{\beta}$ approaches $\partial\mathscr{B}$ above the curve $\mathscr{C}$, we expect the ansatz (3.17) to break down, and the exponents ν, γ, δ are not expected to be defined; the quantities $R_n(\underline{\beta})$ are likely to remain finite. The exponent μ may be defined, but the proof of the scaling relation $\nu = \mu/2$ breaks down. In any event, above the curve $\mathscr{C}$, typical surfaces bounded by a loop $\gamma_{L,L}$ must be expected to be smooth (or rough), but strictly two-dimensional objects, as $L \to \infty$, in the sense that $\delta = 2$.

The value of δ in the regime $\underline{\beta} \to \partial\mathscr{B}$, $\beta_1 = 0$, $\beta_0 \to \infty$, is not known, but is likely to stay below 4.

Denoting by $\|S\|$ the total number of *distinct* plaquettes of $\mathbb{Z}^d$ belonging to a surface $S \in \mathscr{E}$, we may define the *recurrence exponent* α by

$$|S| \underset{|S| \to \infty}{\sim} \|S\|^{\alpha}, \qquad S \in \mathscr{E} \tag{3.36}$$

Of course, for self-avoiding random surfaces, $\alpha = 1$, but for planar surfaces, $\alpha > 1$, for $d < 4$, within the branched polymer regime. Some relation between α and the exponent ε and a lower bound on α,

$$\alpha \geqslant \max(\delta/d, 1) \tag{3.37}$$

are derived in ref. 15; see also ref. 7. One expects that

$$\alpha = \delta/d \quad \text{for} \quad d < \delta$$

$$\alpha = 1 \quad \text{for} \quad d > \delta$$

and

$$|S| \underset{|S| \to \infty}{\sim} \|S\| \, (\log \|S\|)^{\kappa} \quad \text{for} \quad d = \delta$$

for some $\kappa > 0$. (Similar results are known to be true for simple random walks in $\mathbb{Z}^d$.)

Finally, we consider the *orientational order* of typical surfaces. Let p and p' be two plaquettes in $P(S)$, $S \in \mathscr{E}$. We define

$$i(p, p') = \begin{cases} 1 & \text{if } p \text{ and } p' \text{ are parallel} \\ & \text{(i.e., they lie in the same plane} \\ & \text{and have the same orientation)} \\ 0 & \text{otherwise} \end{cases}$$

For p and p' in $P(S)$, let $D(p,p')$ be the distance between p and p' in the intrinsic metric of S [i.e., $D(p, p')$ is the minimal number n of plaquettes, $p_1,..., p_n$, in $P(S)$ such that $p_1 = p$, $p_n = p'$, and p_i, p_{i+1} are glued together along a common link, for all $i = 1,..., n-1$]. A measure of orientational order is the expectation

$$E_{\underline{\beta}}(i(p, p') \mid D(p, p') = a) \equiv G_{\underline{\beta}}(\partial p)^{-1} \sum_{\substack{S \in \mathscr{E} \\ \partial S = \partial p}} \sum_{\substack{p' \in P(S) \\ D(p, p') = a}} i(p, p') \exp[-A_{\underline{\beta}}(S)] \qquad (3.38)$$

We are interested in the asymptotic behavior of this quantity as a becomes large when $\underline{\beta}$ reaches $\partial\mathscr{B}$, for it measures orientational order and is suitable

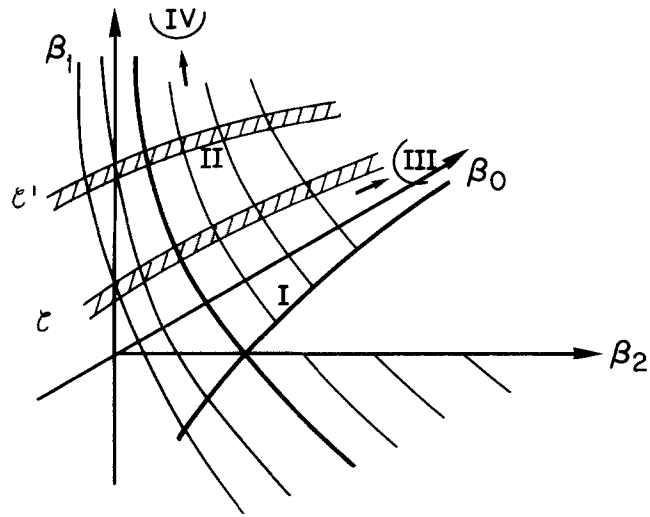


Fig. 1. The part of the critical surface $\partial\mathscr{B}$ in $\beta_i > 0$, $i = 1, 2, 3$. The roman numerals refer to different critical regimes as described in the text. The curve $\mathscr{C}$ separates region I from region II. The bold lines are the intersection of $\partial\mathscr{B}$ with the coordinate planes.

for numerical studies. If (3.38) does not tend to zero as $a \to \infty$, we say there is orientational order.

Our findings (partly heuristic and partly rigorous) are summarized in Fig. 1. We have the following regimes:

I. *Branched polymer regime* (below $\mathscr{C}$): $m(\underline{\beta})$ tends to zero, $\tau(\underline{\beta})$ stays positive, no orientational long-range order.

II. "*Solid-on-solid*" *regime* (above $\mathscr{C}$): $m(\underline{\beta})$ stays positive, orientational order for surface bounded by a planar loop.

III. Regime where $m(\underline{\beta})$ *and* $\tau(\underline{\beta})$ tend to zero (genuine "*surface*" *regime*).

IV. Regime where $\tau(\underline{\beta})$ tends to zero, while $m(\underline{\beta})$ diverges.

$\mathscr{C}'$ indicates a region where a *roughening transition* might be observed as one moves from high to low values of β_1. Note that regimes III and IV are at infinity in coupling constant space.

4. THE RG CONSTRUCTION AND THE EFFECTIVE ACTION

In this section we introduce the notions, notation, and machinery necessary for the decomposition of surfaces into a hierarchy of networks and explain the basic renormalization group step. We specialize to the case of surfaces with one boundary component. The changes that are necessary for the surfaces with two (or more) boundary components are explained in Section 6. At the end of this section we discuss the effective action obtained after carrying out renormalization and show that β_1 always increases as the scale is increased.

Let $\mathscr{E}$ denote the ensemble (vi) of simple planar random surfaces. This is a convenient ensemble for describing our RG method. The extension to the full ensemble of PRS surfaces is in principle straightforward but notationally complicated and is discussed, together with some further generalizations, in Section 8. Throughout Sections 4–6, all surfaces are assumed to belong to $\mathscr{E}$.

Let $S \in \mathscr{E}$ with $\partial S = \mathscr{L}$, where $\mathscr{L}$ is a loop, and denote by E_0 the network in S that contains $\mathscr{L}$. It is possible that $E_0 = \mathscr{L}$.

We shall call E_0 the *base network* of S or the *network at level* 0. Let $F_1, \ldots, F_n$ be the faces of S whose boundaries share links with E_0. These faces are called *faces at level* 0. Let F be one of the faces at level 0, and suppose its boundary consists of k loops $\mathscr{L}_1, \ldots, \mathscr{L}_k$. One of these loops, $\mathscr{L}_1$ say, is the *exterior boundary of* F, meaning that $\mathscr{L}_1$ cannot be removed from ∂F by filling a hole in F. The other boundary components of F (if such exist) $\mathscr{L}_2, \ldots, \mathscr{L}_k$ form the *interior boundary of* F and disappear if the holes they

bound are filled; see Fig. 2. We shall use the notation $\partial_{\text{ext}} F = \mathscr{L}_1$ and $\partial_{\text{int}} F = \mathscr{L}_2 \cup \cdots \cup \mathscr{L}_k$.

Next we show that F shares exactly one of its boundary loops with E_0. Since ∂F intersects E_0 and E_0 is connected, $\partial F \cap E_0$ contains at least one loop. If we cut S along this loop, it becomes disconnected, since it is planar, and therefore $\partial F \cap E_0$ cannot contain another loop, because E_0 is connected and entirely contained in one of the two components that S has been separated into.

A face F is said to be of *type 1* if $\partial_{\text{ext}} F \subseteq E_0$. Otherwise F is of *type 2* (see Fig. 3). Now we are ready to define what we mean by networks at level 1. Let F be a type 1 face at level 0. Let $\partial_{\text{int}} F = \mathscr{L}_1 \cup \cdots \cup \mathscr{L}_n$, where the $\mathscr{L}_i$ are loops. Each of the loops $\mathscr{L}_i$ is contained in some network $E_i \neq E_0$ and $E_i \neq E_j$, if $i \neq j$, since S is planar, by a similar argument as before. The networks $E_1, \ldots, E_n$ are said to be at level 1.

Next, let F be a type 2 face at level 0. Then there is a network $E' \neq E_0$ such that $\partial_{\text{ext}} F \subseteq E'$. We say that the network E' is at level 1, but the other networks that may be attached to the boundary of F (in addition to E_0 and E') are at level 2.

We now give an inductive definition of networks and faces at any level $k \in \mathbb{N}$. Suppose we have defined networks and faces at levels $\leqslant k$ in such a way that the subsurface S_k of S consisting of faces at levels $\leqslant k$ glued together as in S is topologically a sphere with holes and the connected components of ∂S_k different from ∂S are either interior boundary components of faces at level $k-1$ or boundary components (interior or exterior) of faces in S_k at level k. The networks at level $k+1$ are those

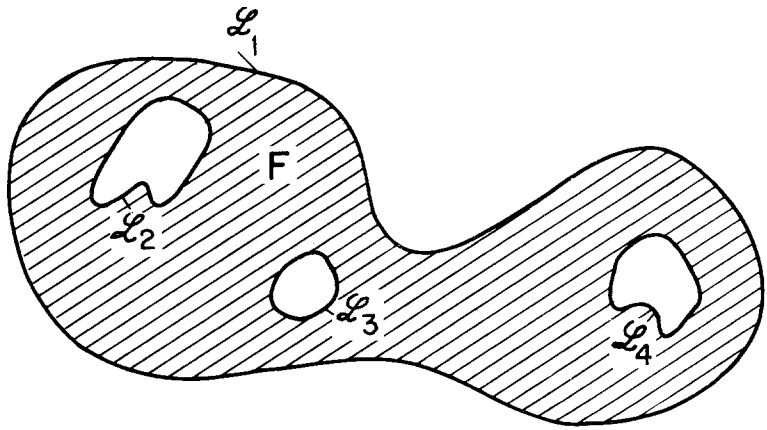
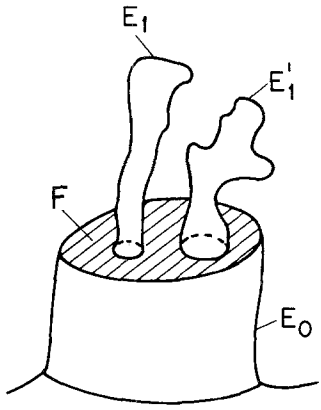


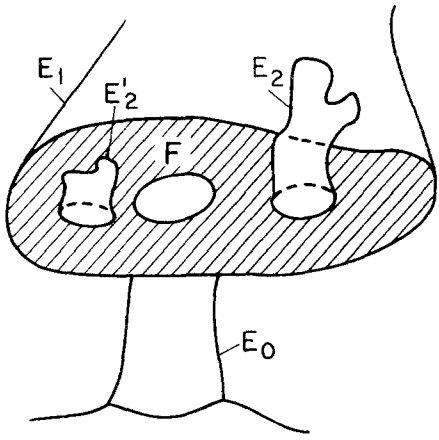
Fig. 2. A face F with exterior boundary $\mathscr{L}_1$ and interior boundary $\mathscr{L}_2 \cup \mathscr{L}_3 \cup \mathscr{L}_4$.

which either contain a boundary component of ∂S_k which is the exterior boundary of a face at level k or an interior boundary component of a face at level k or $k-1$ whose exterior boundary is not in ∂S_k (but in the interior of S_k).

The faces at level $k+1$ are those faces in S whose boundary intersects a level-$(k+1)$ network but not a lower-level network. It is readily seen that the surface S_{k+1}, consisting of all faces at level $\leqslant k+1$, is again a topologi-



(a)



(b)

Fig. 3. (a) A type 1 face F at level 0, with two networks E_1 and E_1' at level 1 attached. (b) A type 2 face F at level 0. The face F intersects a unique network E_1 at level 1 and, in the case depicted here, two networks, E_2 and E_2', at level 2.

cal sphere (with holes) satisfying the properties we assumed for S_k. Note that networks connected to boundary components of ∂S_k that are interior boundary components of faces whose exterior boundary lies in ∂S_k are at level $k+2$ by the above definition.

A face at level k is a type 1 face if its exterior boundary is contained in a level-k network; otherwise the exterior boundary of the face is contained in a level $k+1$ network and the face is said to be of *type 2*. See Fig. 4.

The hierarchical structure of surfaces which we want to exploit should now be manifest. We say that a surface S has rank n, denoted $r(S)=n$, if

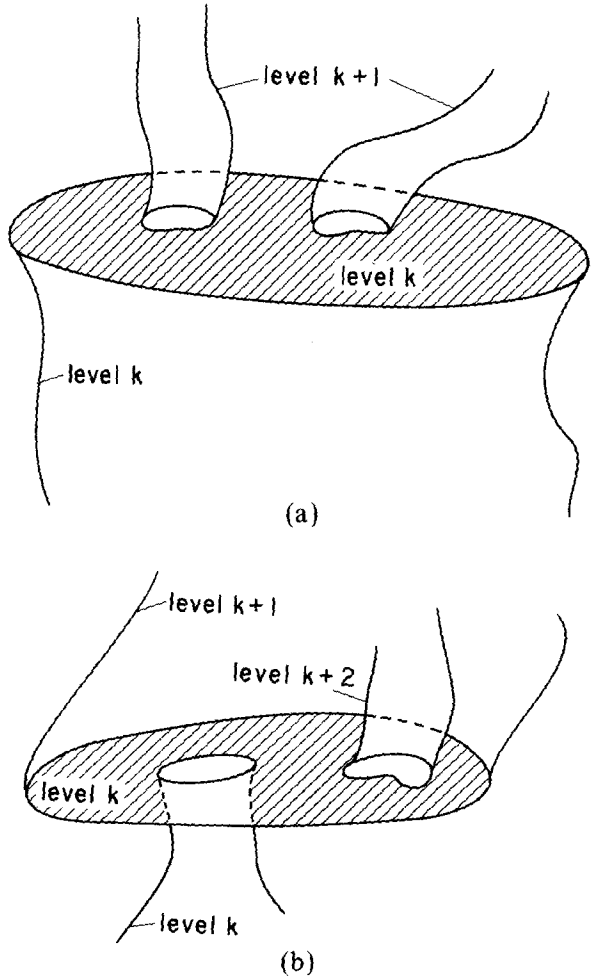


Fig. 4. (a) A type 1 face at level k. (b) A type 2 face at level k.

the highest nonempty level of networks in S is number n. If $r(S)=n$, we define $P_k(S)$ for any $k \leqslant n-1$ to be the subsurfaces of S consisting of all faces at levels $\leqslant k$. This is the surface S_k considered before. If $k \geqslant n$, it is convenient to define $P_k(S)=S$. We say that two surfaces S and S' are *identical up to level k* if and only if $\partial S=\partial S'$ and $P_k(S)=P_k(S')$. In each equivalence class of surfaces that are identical up to level k, we fix, for later use, an ordering of the faces at each level $\leqslant k$.

We now introduce a concept that will be important in the subsequent discussion: *labeled networks*. Roughly speaking, a labeled network $\tilde{E}$ is a network E together with instructions for constructing a unique surface S with $E(S)=E$. Consider a network E contained in a surface S. We fix an arbitrary labeling $e_1(j),\ldots, e_n(j)$ of the edges of E that meet at any corner $j \in E$. Note that every edge carries two labels, since it joins two corners. At each corner j we have a reordering $e_{i_1}(j),\ldots, e_{i_n}(j)$ of the edges meeting at j, determined by the condition that $e_{i_k}(j)$ and $e_{i_{k+1}}(j)$ belong to the same face of S, $k=1,\ldots, n$, $e_{i_{n+1}}(j) \equiv e_{i_1}(j)$. The reorderings determined in the manner described above constitute the labeling of E induced by S. It is not hard to see that different surfaces containing E can induce different labelings. A labeled network $\tilde{E}$ is a network E together with a labeling of E induced by some surface S with $E \subseteq E(S)$.

Given a labeled network $\tilde{E}$, we say that S is a *minimal surface spanning* $\tilde{E}$ if $E(S)=E$ and the labeling of E induced by S is the given one.

A network with a *base loop* is an ordered pair $(E, \mathscr{L})$ where E is a network and $\mathscr{L}$ is a planar loop contained in E. A labeled network with base loop is defined similarly. If E is a network at level k in a surface S, then E has a *canonical base loop*, i.e., the loop where E meets a face at level $k'<k$, $k'=k-1$, or $k'=k-2$, depending on the type of the face and which boundary component of the face E meets. If E is the base network of S, $\partial S=\mathscr{L}$, the base loop of E is defined to be $\mathscr{L}$.

Lemma 7. Given a labeled network with base loop, $(\tilde{E}, \mathscr{L})$, there exists a unique minimal surface $S(\tilde{E}, \mathscr{L})$ spanning $\tilde{E}$ with boundary $\partial S(\tilde{E}, \mathscr{L})=\mathscr{L}$.

Proof. The existence and uniqueness of $S(\tilde{E}, \mathscr{L})$ follow from the following prescription for its construction.

Let S be a surface with $E \subset E(S)$. We can assume without loss of generality that $\tilde{E}$ is the labeled base network of S and $\partial S=\mathscr{L}$. Let F be a level 0 face of S. If F is a type 1 face, we replace F by $\bar{F}$, which is the unique (finite) flat surface with boundary $\partial \bar{F}=\partial_{\text{ext}} F$. If F is a type 2 face, then it meets E_0 at one of its interior boundary components $\mathscr{L}'$ and we replace F by the unique (finite) flat surface $\bar{F}$ with $\partial \bar{F}=\mathscr{L}'$; see Fig. 5. ∎

Note that if we take any labeled network $\tilde{E}$ in a surface S together with its canonical base loop $\mathscr{L}_c$, it follows from the lemma that there is a unique minimal surface spanning $\tilde{E}$ and having boundary $\mathscr{L}_c$.

In the sequel we shall exclusively work with labeled networks, so we adopt the convention that the word *network from now on means labeled network.* We also drop the tilde on the labeled network, $\tilde{E}$, from our notation.

If S is a surface with boundary $\partial S = \mathscr{L}$, we denote its base network E_0 with base loop $\mathscr{L}$ by $\mathscr{N}_0(S)$. Let $F_1, F_2, \ldots, F_n$ be the zeroth-level faces of S, ordered by our fixed convention, and let $S(E_0, \mathscr{L})$ be the minimal surface spanning E_0 with boundary $\mathscr{L}$.

Let $\bar{F}_1, \ldots, \bar{F}_n$ denote the faces of $S(E_0, \mathscr{L})$. It follows from the proof of Lemma 7 that there is a natural 1–1 correspondence between the faces of $S(E_0, \mathscr{L})$ and the zeroth-level faces of S. We may assume that $\bar{F}_i$ corresponds to F_i. Then, if F_i is a type 1 face, $\bar{F}_i$ is obtained by "filling the interior holes" in F_i, and if F_i is a type 2 face, then $\bar{F}_i$ is a "filling of a hole in F_i."

Assume that the face F_i is a type 1 face. Let $E_i^1, \ldots, E_i^{k(i)}$ be the first-level networks that meet F_i. Let $\mathscr{L}_i^j$ be the canonical base loop of E_i^j. Then $\mathscr{L}_i^j$ is one of the interior boundary components of F_i. Let $F(\mathscr{L}_i^j)$ be the subsurface of $\bar{F}_i$ bounded by $\mathscr{L}_i^j$.

Now let F_i be a type 2 face. Then F_i and $\bar{F}_i$ meet in the single loop $\partial \bar{F}_i$ and there is only one network at level 1 that meets F_i. This network we denote by E_i^1. We denote the canonical base loop of E_i^1 by $\mathscr{L}_i^1$ and observe

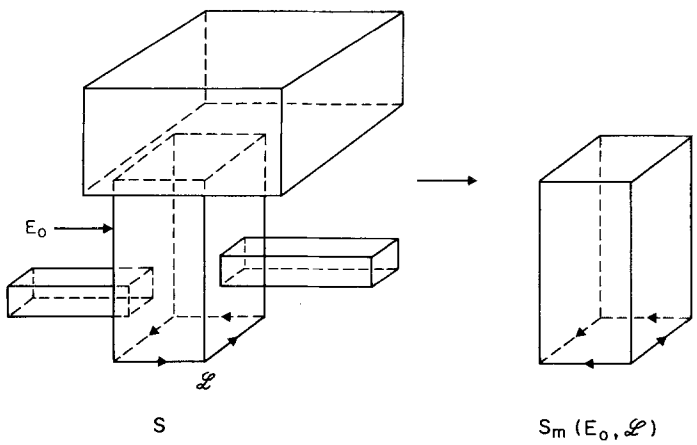


Fig. 5. A surface S with base network E_0 and boundary $\mathscr{L}$. The surface $S_m(E_0, \mathscr{L})$ is obtained by removing all "outgrowths" from S and filling the resulting holes.

that $\partial \bar{F}_i$ lies in the interior of $\mathscr{L}_i^1$. Denote the collection of all the first-level networks, with their canonical base loops, by

$$\mathscr{N}_1(S) = \{(E_i^1, \mathscr{L}_i^1), \ldots, (E_i^{k(i)}, \mathscr{L}_i^{k(i)})\}_{i=1}^n \tag{4.1}$$

where n is the number of level 0 faces in S.

We define the set $\Phi(E_0)$ of excitations of E_0 by

$$\Phi(E_0) = \{\mathscr{N}_1(S) \mid S \in \mathscr{E}, \mathscr{N}_0(S) = E_0, \partial S = \mathscr{L}\} \tag{4.2a}$$

and

$$\Phi(\bar{F}_i) = \{(E_i^1, \mathscr{L}_i^1), \ldots, (E_i^{k(i)}, \mathscr{L}_i^{k(i)})\} \tag{4.2b}$$

Note that $\Phi(E_0)$ contains the empty set, corresponding to $S = S(E_0, \mathscr{L})$.

In the following, we need generalizations of the definitions (4.1), (4.2a), and (4.2b). Given a simple, planar random surface $S \in \mathscr{E}$, we define $\mathscr{N}_k(S)$ to be the collection of all pairs $(E, \mathscr{L})$, where E is a labeled network of S at level k and $\mathscr{L}$ is the base loop of E. Given $(E, \mathscr{L})$ in $\mathscr{N}_k(S)$, we define

$$\Phi(E) = \{\mathscr{N}_1(S') \mid S' \in \mathscr{E}, \mathscr{N}_0(S') = E, \partial S' = \mathscr{L}\} \tag{4.2c}$$

Let F_i be a face of a labeled network E, with $(E, \mathscr{L}) \in \mathscr{N}_k(S)$. The set of excitations $\Phi(\bar{F}_i)$ of F_i is defined as in (4.2b), but, instead of being a face of E_0, F_i is a face of E.

In order to explain the basic ideas of our renormalization group analysis of lattice random surface models, without having to worry about inessential technical difficulties, we now introduce a *modified action* for simple planar random surfaces: Let $S \in \mathscr{E}$, and let $(E, \mathscr{L})$ be an element of $\mathscr{N}_k(S)$. Then $\mathscr{L}$ is a simple loop in a two-dimensional lattice plane. Let $A(\mathscr{L})$ be the number of plaquettes in that plane which are enclosed by $\mathscr{L}$. We set

$$\hat{A}_{\underline{\beta}}(S) = \beta_2 \left(|S| + \sum_k \sum_{(E, \mathscr{L}) \in \mathscr{N}_k(S)} A(\mathscr{L}) \right) + \beta_1 |E(S)| + \beta_0 N(S) \tag{4.3}$$

The techniques developed in the remainder of Section 4 and in Section 5 can be extended to the model studied previously, but the details would be substantially more complicated.

From definitions (4.2a), (4.3) we easily obtain the following identity:

$$G_{\underline{\beta}}(\mathscr{L}) = \sum_{E_0} \left\{ \sum_{\mathscr{F} \in \Phi(E_0)} \sum_{S: \mathscr{N}_1(S) = \mathscr{F}} \exp[-\hat{A}_{\underline{\beta}}(S)] \right\} \tag{4.4}$$

where the E_0 sum runs over all networks with base loop $\mathscr{L}$. This identity allows us in essence, as we shall see, to pass from summation over surfaces

to a summation over networks. Our next step is to represent the weights of the networks as sums over surfaces, which in turn can be expressed as weighted sums over networks.

Let $\mathscr{L}$ be a simple loop lying in a coordinate plane. Let S be a surface with $\partial S = \mathscr{L}$, and, as before, let $F(\mathscr{L})$ denote the unique finite flat surface with boundary $\mathscr{L}$. We shall use the notation $S \perp \mathscr{L}$ to indicate that none of the boundary plaquettes of S are contained in $F(\mathscr{L})$.

Given a surface S with $\partial S = \mathscr{L}$ and a base network E_0, we can regard S as a "collective excitation" of the surface $S(E_0, \mathscr{L})$. The individual "excitations" are represented by the first-level networks, each of which can be attached to a definite face of $S(E_0, \mathscr{L})$. Consider a first-level network with base loop, $(E_i^j, \mathscr{L}_i^j) \in \mathscr{N}_1(S)$. If F_i is a type 1 face, then $\mathscr{L}_i^j \in \bar{F}_i$ and the statistical weight of the sum over all surfaces with base network E_0 having the excitation $(E_j^i, \mathscr{L}_j^i)$ contains the factor

$$\sum_{\substack{S': \mathscr{N}_0(S') = E_i^j \\ \partial S' = \mathscr{L}_i^j, S' \perp \mathscr{L}_i^j}} \exp[-\hat{A}_{\underline{\beta}}(S')] \equiv \eta_{\underline{\beta}}^{(1)}(E_i^j, \mathscr{L}_i^j) \tag{4.5}$$

If F_i were a type 2 face, then $\bar{F}_i \subseteq F(\mathscr{L}_i^j)$ and the statistical weight of surfaces containing $(E_i^j, \mathscr{L}_i^j)$ would include the factor

$$\sum_{\substack{S': \mathscr{N}_0(S') = \partial S = \partial \bar{F}_i \\ \mathscr{N}_1(S) = \{(E_i^j, \mathscr{L}_i^j)\}}} \exp[-\hat{A}_{\underline{\beta}}(S') + \hat{A}(\mathscr{L})] \equiv \eta_{\underline{\beta}}^{(2)}(E_i^j, \mathscr{L}_i^j) \tag{4.6}$$

where $\hat{A}_{\underline{\beta}}(\mathscr{L})$ is the action associated with the links of $\mathscr{L} = \partial F_i$, i.e.,

$$\hat{A}_{\underline{\beta}}(\mathscr{L}) = \beta_1 |\mathscr{L}| + \beta_0 N(\mathscr{L}) \tag{4.7}$$

where $N(\mathscr{L})$ is the number of corners in $\mathscr{L}$. Let us define

$$v_{ij} = \begin{cases} 1 & \text{if } \mathscr{L}_i^j \subseteq \bar{F}_i \\ 2 & \text{otherwise [i.e., if } \bar{F}_i \subseteq F(\mathscr{L}_i^j)] \end{cases} \tag{4.8}$$

Lemma 8. Let

$$\mathscr{F} = \{(E_i^1, \mathscr{L}_i^1), \ldots, (E_i^{k(i)}, \mathscr{L}_i^{k(i)})\}_{i=1}^n$$

be the first-level networks, together with their base loops, of a surface $S \in \mathscr{E}$, with $\partial S = \mathscr{L}$ and $\mathscr{N}_0(S) = E_0$. Then

$$\sum_{\substack{S': \mathscr{N}_0(S') = E_0 \\ \partial S' = \mathscr{L}, \mathscr{N}_1(S') = \mathscr{F}}} \exp[-\hat{A}_{\underline{\beta}}(S')]$$

$$= \exp[-\hat{A}_{\underline{\beta}}(S(E_0, \mathscr{L})] \prod_{i=1}^{n} \left(\prod_{j=1}^{k(i)} \eta_{\underline{\beta}}^{(v_{ij})}(E_i^j, \mathscr{L}_i^j) \right) \tag{4.9}$$

Proof. Since our surfaces are simple and planar, and hence do not satisfy any constraints such as self-avoidance, we can independently sum over pieces of surface bounded by distinct loops $\mathscr{L}_i^j$ and with given base networks E_i^j. It is clear from the definition (4.3) of $\hat{A}_{\underline{\beta}}$ that a factor $\exp[-\hat{A}_{\underline{\beta}}(S(E_0, \mathscr{L}))]$ can be extracted from the sum on the lhs of (4.9). (In this regard, our modified action $\hat{A}_{\underline{\beta}}$ leads to substantial simplifications of our analysis as compared to the original action $A_{\underline{\beta}}$.) Equation (4.9) now follows from the definition of the "activities" $\eta_{\underline{\beta}}^{(\nu_{ij})}(E_i^j, \mathscr{L}_i^j)$; see (4.5), (4.6). ∎

It is convenient to introduce the quantities

$$\mathscr{Z}_{\underline{\beta}}(E_0) = \sum_{\mathscr{F} \in \Phi(E_0)} \prod_{i=1}^{n} \prod_{j=1}^{k(i)} \eta_{\underline{\beta}}^{(\nu_{ij})}(E_i^j, \mathscr{L}_i^j) \tag{4.10}$$

and

$$\mathscr{Z}_{\underline{\beta}}(\bar{F}_i) = \sum_{\mathscr{F}_i \in \Phi(\bar{F}_i)} \prod_{j=1}^{k(i)} \eta_{\underline{\beta}}^{(\nu_{ij})}(E_i^j, \mathscr{L}_i^j) \tag{4.11}$$

so that

$$\mathscr{Z}_{\underline{\beta}}(E_0) = \prod_{i=1}^{n} \mathscr{Z}_{\underline{\beta}}(\bar{F}_i) \tag{4.12}$$

We regard $\mathscr{Z}_{\underline{\beta}}(E_0)$ and $\mathscr{Z}_{\underline{\beta}}(\bar{F}_i)$ as the partitions functions of a gas of excitations living on the minimal surface and the face $\bar{F}_i$, respectively. As Eq. (4.12) shows, different faces are totally decoupled.

The statistical weight of all surfaces with base network E_0 and boundary $\mathscr{L}$ is given by

$$\{\exp[-\hat{A}_{\underline{\beta}}(S(E_0, \mathscr{L}))]\}\, \mathscr{Z}_{\underline{\beta}}(E_0) \tag{4.13}$$

We think of $\hat{A}_{\underline{\beta}}(S(E_0, \mathscr{L}))$ as the "bare action" of the network E_0 and $\mathscr{Z}_{\underline{\beta}}(E_0)$ as a "renormalization" due to the presence of higher-level networks, whose main influence, at least in a first approximation, is to modify the values of the bare coupling constants $\underline{\beta} = (\beta_0, \beta_1, \beta_2)$.

Before discussing how $\mathscr{Z}_{\underline{\beta}}(E_0)$ depends on E_0 and how the coupling constants flow upon renormalization, let us briefly return to the hierarchical structure of surfaces. We could apply Lemma 8 to the surface contributing to $\mathscr{Z}_{\underline{\beta}}(E_0)$, peel off the bare action of the first-level networks, and be left with a renormalization due to second-level networks and higher. This process could continue *ad infinitum.*

Rather than thinking of applying the RG from "bottom up" as described above, it is better to consider surfaces of a fixed rank ρ. Consider a network at level ρ. It cannot carry any excitations, so its action is the bare action. Hence, the activities of networks at level $\rho - 1$ are explicit and we can in principle compute all activities. Once we are down at level 0 we can let ρ tend to infinity. This method will of course not allow us to calculate anything explicitly, but enables us to prove bounds, in particular Theorem 1 and, with some assumptions, Quasitheorem 6.

Thus, in the following we consider surfaces of maximal rank $\rho < \infty$. We propose to prove some statements about the effective action of surfaces at some fixed level $\rho < \infty$ which are uniform in the "cutoff" ρ. That permits us to eventually let ρ tend to ∞ without losing control.

Let F be some face at level n. Then the partition function of F is given by

$$\mathscr{Z}_{\beta}(F; n) = \sum_{\mathscr{F} \in \Phi(F)} \prod_{(E, \mathscr{L}) \in \mathscr{F}} \eta^{(v(E))}(E; n+1) \tag{4.14}$$

where $\eta^{(v(E))}(E; n+1)$ are effective activities of labeled networks at level $n+1$, with base loops $\mathscr{L} \subset F$ for $v(E) = 1$ and $F \subseteq F(\mathscr{L})$ for $v(E) = 2$ [in which case the product on the rhs of (4.14) contains a single factor $\eta^{(2)}(E; n+1)$].

The induction hypothesis of our renormalization group analysis will be the following bound, proven in Section 5:

$$\sum_{E\text{: base loop of } E = \mathscr{L}} |\eta^{(v)}(E; n+1)| \leqslant e^{-\kappa_{n+1}|\mathscr{L}|} \tag{4.15}$$

$v = 1, 2$, where $|\mathscr{L}|$ is the length of $\mathscr{L}$ and $\kappa_{n+1} > \log 3$, for all n. Clearly,

$$\sum_{\mathscr{L}\text{: dist}(\mathscr{L}, 0) \leqslant |\mathscr{L}|} e^{-\kappa_{n+1}|\mathscr{L}|} \equiv \eta(\kappa_{n+1}) < \infty \tag{4.16}$$

with $\eta(\kappa) \to 0$ as $\kappa \to \infty$. This is a well-known bound on sums of two-dimensional self-avoiding walks. From (4.14)–(4.16) we conclude that

$$\mathscr{Z}_{\beta}(F; n) \leqslant [1 + \eta(\kappa_{n+1})]^{2|F|} \tag{4.17}$$

Thus, assuming (4.15), $\mathscr{Z}_{\beta}(F; n)$ is finite, and its logarithm is bounded by $O(|F|)$, where $|F|$ is the area of F.

If κ_{n+1} is large enough, then the sum on the rhs of (4.14) can be carried out with the help of a convergent high-temperatuire expansion. Let us describe the results obtained from such an expansion: Let F be a rectangular region with sides of length L and M, contained in $\mathbb{Z}^2$. We set

$$p_2 = \lim_{L,M\to\infty} \frac{1}{LM} \log \mathscr{Z}_{\underline{\beta}}(F; n) \tag{4.18}$$

$$2p_1 = \lim_{L,M\to\infty} \frac{1}{L+M} [\log \mathscr{Z}_{\underline{\beta}}(F; n) - p_2 L \cdot M] \tag{4.19}$$

$$4p_0 = \lim_{L,M\to\infty} [\log \mathscr{Z}_{\underline{\beta}}(F; n) - p_2 L \cdot M - p_1(L+M)] \tag{4.20}$$

It is a well-known consequence of the absolute convergence of the high-temperature expansion that the limits in (4.18)–(4.20) exist; see refs. 19–21 and references given there.

Let $|\partial F|$ denote the length of ∂F and $|C(F)|$ the number of corners of ∂F; the set of corners $C(F)$ of ∂F is the set of sites in ∂F where two edges in ∂F meet at an angle different from π.

Lemma 9. Let $\mathscr{Z}_{\underline{\beta}}(F; n)$ be given by (4.14), and suppose that the constant κ_{n+1} in (4.15) is sufficiently large. Then

$$\log \mathscr{Z}_{\underline{\beta}}(F; n) = p_2 |F| + p_1 |\partial F| + p_0 |C(F)| + O\{\exp[-c(\kappa_{n+1})\, d(F)]\} \tag{4.21}$$

where p_2, p_1, and p_0 are defined in (4.18), (4.19), and (4.20), $c(\kappa_{n+1})$ is positive for κ_{n+1} sufficiently large and tends to $+\infty$ as $\kappa_{n+1} \to \infty$, and $d(F)$ is the smallest distance between edges in ∂F which do *not* share a common corner of ∂F.

Lemma 9 is a fairly standard consequence of absolute convergence of the high-temperature expansion. Here we only sketch the idea of the proof; the technical details can be extracted from refs. 20 and 21. Let F be a finite region in $\mathbb{Z}^2$. Let $\mathscr{P}$ be a countable set of "polymers" (finite subsets of $\mathbb{Z}^2$ with connected boundary). Let z be a complex-valued function on $\mathscr{P}$, called the "activity function," with the property that

$$|z(p)| < e^{-c|\partial p|} \qquad \text{for} \quad p \in \mathscr{P} \tag{4.22}$$

with $c > \log 3$. We define a partition function

$$\mathscr{Z}(F; \mathscr{P}, z) = \sum_n \sum_{\substack{p_1,\ldots,p_n \in \mathscr{P} \\ p_i \subset F, \forall i \\ p_i \cap p_j = \varnothing, \forall i \neq j}} \prod_{j=1}^{n} z(p_j) \tag{4.23}$$

By (4.22), the sum on the rhs of (4.23) can be carried out with the help of a convergent "high-temperature" expansion, for c large enough. To

describe the result of that expansion, we require the following notions: We define a multiplicity function X on $\mathscr{P}$ to be a function on $\mathscr{P}$ with values in $\{0, 1, 2, 3, \ldots\}$. A multiplicity function X is called connected if, for every pair (p_1, p_2), with $p_1, p_2 \in \mathscr{P}$, such that $X(p_1) > 0$ and $X(p_2) > 0$, there is a sequence of polymers $p^1, \ldots, p^N$ such that $p^1 = p_1$, $p^N = p_2$, $X(p^i) > 0$, for $i = 1, \ldots, N$, and $p^i \cap p^{i+1} \neq \emptyset$, for all $i = 1, \ldots, N-1$. If X is a multiplicity function, we define $\operatorname{supp} X$ to be the union over all $p \in \mathscr{P}$ for which $X(p) > 0$.

The high-temperature expansion for $\mathscr{Z}(F; \mathscr{P}, z)$ yields the following representation for $\log \mathscr{Z}(F; \mathscr{P}, z)$: There is a function φ^T defined on the set of all multiplicity functions X on $\mathscr{P}$ such that $\varphi^T(X) = 0$ un less X is connected, and

$$\log \mathscr{Z}(F; \mathscr{P}, z) = \sum_{\substack{X \\ \operatorname{supp} X \subseteq F}} \varphi^T(X) \prod_{p \in \mathscr{P}} z(p)^{X(p)} \tag{4.24}$$

We now assume that z is translation invariant, i.e.,

$$z(p) = z(p_a) \tag{4.25}$$

where p_a is the translate of a polymer $p \in \mathscr{P}$ by a lattice vector $a \in \mathbb{Z}^2$. We also assume, to simplify matters, that z is invariant under rotations of the lattice.

We then define

$$p_2 = \sum_{X:\ \operatorname{supp} X \ni 0} \varphi^T(X) \prod_{p \in \mathscr{P}} z(p)^{X(p)} \tag{4.26}$$

Let L_+ and L_- be the half-lattices $\{(x, y) \in \mathbb{Z}^2;\ x \gtrless 0\}$. We set

$$p_1 = \sum_{\substack{X:\ \operatorname{supp} X \ni 0, \\ \operatorname{supp} X \cap L_+ \neq \emptyset \\ \operatorname{supp} X \cap L_- \neq \emptyset}} \varphi^T(X) \prod_{p \in \mathscr{P}} z(p)^{X(p)} \tag{4.27}$$

The quantity p_2 is interpreted as bulk pressure and p_1 is interpreted as edge pressure. Similarly, a quantity p_0, interpreted as "corner pressure," can be defined. The key fact used in the proof of Lemma 9 is that if the constant c in (4.22) is large enough, the expansions for $\log \mathscr{Z}(F; \mathscr{P}, z)$, p_2, p_1, and p_0 [see (4.24), (4.26), and (4.27)], are all absolutely convergent. Representation (4.21) follows from these convergence properties by using the inclusion–exclusion principle. The "error term" $O\{\exp[-c(\kappa_{n+1})\, d(F)]\}$ on the rhs of (4.21) refers to the contribution of all multiplicity functions X with the property that $\operatorname{supp} X$ contains in its interior two sites belonging to two different edges of ∂F which do *not* have

a common corner. Our estimate, $O\{\exp[-c(\kappa_{n+1})\, d(F)]\}$, on the size of that error term follows from (4.15) and (4.22), respectively.

For further details we refer the reader to the literature.[20]

Lemma 10. If $\eta^{(\nu)}(E; n+1) \geqslant 0$, for all networks E, then

$$p_1 \leqslant 0$$

Proof. Consider a square face F with sides of length NL for some constant L and $N = 2, 3, 4, \ldots$. The face F can be subdivided into square faces $F_1, F_2, \ldots, F_{N^2}$, with sides of length L. In definition (4.14) of $\mathscr{Z}_{\underline{\beta}}(F; n)$ it is understood that, for every $(E, \mathscr{L}) \in \mathscr{F}$,

$$\mathscr{L} \cap \partial F = \varnothing \tag{4.28}$$

We suppose, to begin with, that $\nu(E) = 1$ for all $(E, \mathscr{L}) \in \mathscr{F}$; this defines a modified partition function $\mathscr{Z}'_{\underline{\beta}}(F; n)$. By (4.28),

$$\mathscr{Z}'_{\underline{\beta}}(F, n) \geqslant \prod_{j=1}^{N^2} \mathscr{Z}'_{\underline{\beta}}(F_j; n) \tag{4.29}$$

From this inequality we conclude that

$$\frac{1}{N^2 L^2} \log \mathscr{Z}'_{\underline{\beta}}(F; n) \geqslant \frac{1}{L^2} \log \mathscr{Z}'_{\underline{\beta}}(F_j; n) \tag{4.30}$$

By translation invariance, the rhs of (4.30) is independent of j. We conclude that

$$\frac{1}{(2^N L)^2} \log \mathscr{Z}'_{\underline{\beta}}(F^{(N)}; n) \text{ is monotone increasing in } N \tag{4.31}$$

Here $F^{(N)}$ is a square face with sides of length $2^N L$. Next, we note that

$$\sum_{(E, \mathscr{L})} \eta^{(2)}(E; n+1) \leqslant e^{-CL} \tag{4.32}$$

if the sum extends over all excitations of type 2 of a square face with sides of length L. On the rhs of (4.32), C is a positive constant. The bound (4.32) follows from (4.15). Hence

$$\begin{aligned} \lim_{L \to \infty} \frac{1}{L^2} \log \mathscr{Z}_{\underline{\beta}}(F_j; n) &= \lim_{L \to \infty} \frac{1}{L^2} \log \mathscr{Z}'_{\underline{\beta}}(F_j; n) \\ \lim_{L \to \infty} \frac{1}{N^2 L^2} \log \mathscr{Z}_{\underline{\beta}}(F; n) &= \lim_{L \to \infty} \frac{1}{N^2 L^2} \log \mathscr{Z}'_{\underline{\beta}}(F; n) \end{aligned} \tag{4.33}$$

Comparing (4.30), (4.31), and (4.33) with (4.21), we conclude that

$$p_1 \leqslant 0 \quad \blacksquare$$

Next, let E be a labeled network at level n. Let $S(E)$ be the minimal surface spanned by E, and let $\partial S(E) \equiv \mathscr{L}$ be the base loop of E. Then, by Lemma 9,

$$\begin{aligned}\eta^{(v)}(E, \eta) = \exp\{ -[\beta_2 |S(E)| + \beta_1 |E| + \beta_0 N(E)]\} \\ \times \exp(p_2(n+1) |S(E)| + p_1(n+1) |E| \\ + p_0(n+1) N(E) + O\{\exp[-c(\kappa_{n+1}) d(E)]\})\end{aligned} \tag{4.34}$$

where $p_0(n+1)$, $p_1(n+1)$, and $p_2(n+1)$ are the pressures introduced in (4.18)–(4.20) at level $n+1$, and $d(E)$ is the smallest distance between two edges in E which do not share a common corner.

Hence, the effective action of $S(E)$ at level n is given by

$$\begin{aligned}\hat{A}^{(n)}(S(E)) = \beta_2(n) |S(E)| + \beta_1(n) |E| \\ + \beta_0(n) N(E) + \text{nonlocal terms}\end{aligned} \tag{4.35}$$

where

$$\begin{aligned}\beta_2(n) &= \beta_2 - p_2(n+1) \\ \beta_1(n) &= \beta_1 - p_1(n+1) \\ \beta_0(n) &= \beta_0 - p_0(n+1)\end{aligned} \tag{4.36}$$

The quantities $p_0(n+1)$, $p_1(n+1)$, and $p_2(n+1)$ depend on $\beta_0(n+1)$, $\beta_1(n+1)$, and $\beta_2(n+1)$. One can argue that the "nonlocal terms" on the rhs of (4.35) are irrelevant. Hence, up to irrelevant terms, the effective action has a form that is *independent* of n. Equations (4.36) are the *renormalization group flow equations* for the coefficients of the local terms in the effective action, the area, edge, and corner terms. In the study of these flow equations, the expansions (4.24), (4.26), (4.27),..., are useful (although the requirement of absolute convergence of these expansions introduces limitations). However, Lemma 10 yields a completely nonperturbative result on these flow equations: This lemma says that $p_1(n+1) \leqslant 0$ and hence

$$\beta_1(n+1) \geqslant \beta_1 \tag{4.37}$$

(β_1 is the bare coupling constant on the highest level), for all n. Inequality (4.36) is a nonperturbative proof of *asymptotic freedom* of the extrinsic

curvature term. If one could achieve better control over the nonlocal, "irrelevant" terms in the effective action, then (4.36) would yield a proof of Quasitheorem 6 (Section 3).

5. PROOF OF THEOREM 1

In this section we use the machinery introduced in the previous section to give a proof of Theorem 1. An "ideal gas" upper bound on $\mathscr{Z}_{\underline{\beta}}(E)$ and an inductive argument yield the desired result. First, we state a lemma, whose combinatorial proof is straightforward. The lemma also follows from the more involved combinatorial estimates of surfaces in Section 8.

Lemma 11. For any $\beta_1 > 0$ there exists a $B_0(\beta_1) < \infty$ such that, if $\beta_0 > B_0(\beta_1)$, then the sum

$$\sum_{E: E \ni x} e^{-\beta_1 |E| - \beta_0 N(E)} \equiv q(\beta_0, \beta_1) \tag{5.1}$$

over all networks that contain a given point $x \in \mathbb{Z}^d$ is convergent.

Define

$$G_{\underline{\beta}}^{(\rho)}(\mathscr{L}) = \sum_{\substack{S: \partial S = \mathscr{L} \\ r(S) \leqslant \rho}} \exp[-\hat{A}_{\underline{\beta}}(S)] \tag{5.2}$$

Clearly, $G_{\underline{\beta}}^{(\rho)}(\mathscr{L}) \nearrow G_{\underline{\beta}}(\mathscr{L})$ as $\rho \to \infty$. Similarly, we define $\eta_{\underline{\beta},\rho}^{(i)}(E, \mathscr{L})$ to be the contribution to $\eta_{\underline{\beta}}^{(i)}(E, \mathscr{L})$ coming from surfaces of rank $\leqslant \rho$, i.e.,

$$\eta_{\underline{\beta},\rho}^{(1)}(E, \mathscr{L}) = \sum_{\substack{S: \mathscr{N}_0(S) = E \\ \partial S = \mathscr{L},\, S \perp \mathscr{L} \\ r(S) \leqslant \rho}} \exp[-\hat{A}_{\underline{\beta}}(S)] \tag{5.3}$$

and similarly for $\eta_{\underline{\beta},\rho}^{(2)}(E, \mathscr{L})$.

Our goal is to prove inductive bounds on $\eta_{\underline{\beta},\rho}^{(i)}(E, \mathscr{L})$ which imply a uniform bound on $G_{\underline{\beta}}^{(\rho)}(\mathscr{L})$, and thereby prove the essential part of Theorem 1. Let $\beta_2 > 0$ be given. By Lemma 11 we can choose $\beta_0 + \beta_1$ so large (with $\beta_1 > 0$) that

$$\sum_{\substack{(E, \mathscr{L}) \\ x \in \mathscr{L}}} |\mathscr{L}|\, e^{-\beta_1 |E| - \beta_0 N(E)} < \beta_2 \tag{5.4}$$

where the summation is over all networks with a base loop that contains a given point $x \in \mathbb{Z}^d$.

Define $\mathscr{Z}_{\beta,0}(F) = 1$ for any face F, and, for $\rho \geqslant 1$, let $\mathscr{Z}_{\underline{\beta},\rho}(F)$ be defined by (4.11) with $\eta_{\underline{\beta}}^{(i)}(E, \mathscr{L})$ replaced by $\eta_{\underline{\beta},\rho}^{(i)}(E, \mathscr{L})$. Note that

$$\eta_{\underline{\beta},1}^{(1)}(E, \mathscr{L}) = \exp[-\hat{A}_{\underline{\beta}}(S(E, \mathscr{L}))] \tag{5.5}$$

and if $\eta^{(2)}_{\underline{\beta},1}(E, \mathscr{L})$ represents an excitation of a face F, then by (4.9)

$$\eta^{(2)}_{\underline{\beta},1}(E, \mathscr{L}) = \exp[-\beta_2(|F(\mathscr{L})| - |F|)] \exp[-\hat{A}_{\underline{\beta}}(S(E, \mathscr{L}))] \tag{5.6}$$

Since $F(\mathscr{L}) \geqslant |F|$, we have

$$\eta^{(i)}_{\underline{\beta},1}(E, \mathscr{L}) \leqslant \exp[-\beta_1 |E| - \beta_0 N(E)] \tag{5.7}$$

for $i = 1, 2$.

Now suppose that, for $k = 1, 2, \ldots, \rho$,

$$\eta^{(i)}_{\underline{\beta},k}(E, \mathscr{L}) \leqslant \exp[-\beta_1 |E| - \beta_0 N(E)] \tag{5.8}$$

By Lemma 8 we have

$$\eta^{(1)}_{\underline{\beta},\rho+1}(E, \mathscr{L}) \leqslant \exp[-\hat{A}_{\underline{\beta}}(S(E, \mathscr{L}))] \times \sum_{\mathscr{F} \in \Phi(E)} \prod_{i=1}^{n} \prod_{j=1}^{k(i)} \eta^{(v_{ij})}_{\underline{\beta},\rho}(E^j_i, \mathscr{L}^j_i) \tag{5.9}$$

Dropping the constraint that the loops $\mathscr{L}^i_j$ are not allowed to intersect (if $v_{ij} = 1$), we obtain from (5.8) that

$$\begin{aligned}
&\sum_{\mathscr{F} \in \Phi(E)} \prod_{i=1}^{n} \prod_{j=1}^{k(i)} \eta^{(v_{ij})}_{\underline{\beta},\rho}(E^j_i, \mathscr{L}^j_i) \\
&\quad \leqslant \prod_{i=1}^{n} \exp\left\{ \sum_{\substack{(E', \mathscr{L}') \\ \mathscr{L}' \cap F_i \neq \varnothing}} \exp[-\beta_1 |E'| - \beta_0 N(E')] \right. \\
&\qquad \left. + \sum_{\substack{(E', \mathscr{L}') \\ F(\mathscr{L}') \geqslant F_i}} \exp[-\beta_1 |E'| - \beta_0 N(E')] \right\} \\
&\quad \leqslant \prod_{i=1}^{n} \exp\left\{ |F_i| \sum_{\substack{(E', \mathscr{L}') \\ \mathscr{L}' \ni 0}} |\mathscr{L}'| \exp[-\beta_1 |E'| - \beta_0 N(E')] \right\} \\
&\quad \leqslant \exp[\beta_2 |S(E, \mathscr{L})|]
\end{aligned}$$

if $\beta_0 + \beta_1$ is chosen according to (5.4). Combining this with (5.9), we reproduce the bound (5.8) for $\eta^{(1)}_{\underline{\beta},\rho+1}$, i.e.,

$$\eta^{(1)}_{\underline{\beta},\rho+1} \leqslant \exp[-\beta_1 |E| - \beta_0 N(E)]$$

The same bound for $\eta^{(2)}_{\underline{\beta},\rho+1}$ is proven similarly. Thus, (5.8) is valid for all k, provided $\beta_0 + \beta_1$ is chosen sufficiently large, for a given $\beta_2 > 0$.

It now follows from (4.3), (4.7), and (5.1) that

$$G_{\beta,\rho}(\mathscr{L}) \leqslant \text{const} \qquad \text{for all} \quad \rho \in \mathbb{N}$$

and hence $G_{\beta}(\mathscr{L}) < \infty$ for arbitrary $\beta_2 > 0$, provided $\beta_0 + \beta_1$ is sufficiently large.

6. THE RG METHOD FOR THE TWO-LOOP FUNCTION

In this section we describe how the procedure of Section 4 can be generalized in order to decompose surfaces with two (or more) boundary components into a hierarchy of networks and faces. Then we give sufficient conditions for Quasitheorem 6 to be valid.

Let S be a surface with two boundary components $\mathscr{L}$ and $\mathscr{L}'$. Let E be a network in S, and let $E = \bigcup_{i=1}^{n} \mathscr{L}_i$ be the decomposition of E into loops, i.e., each $\mathscr{L}_i$ is a boundary component of one of the faces in S that meet E. Suppose we cut the surface S along one of the loops $\mathscr{L}_i$. Then, since S has planar topology, it becomes separated into two components, S_1 and S_2. If we assume that S_1 is the component that contains $\mathscr{L}$, there are two possibilities: Either $\partial S_1 = \mathscr{L} \cup \mathscr{L}_i$ and $\partial S_2 = \mathscr{L}_i \cup \mathscr{L}'$ or $\partial S_1 = \mathscr{L} \cup \mathscr{L}_i \cup \mathscr{L}'$ and $\partial S_2 = \mathscr{L}_i$. In the first case we say that $\mathscr{L}_i$ is *relevant*, but in the second case we say that it is *irrelevant* (see Fig. 6). If one of the loops $\mathscr{L}_i$ in the network E is relevant, we say that the network itself is relevant.

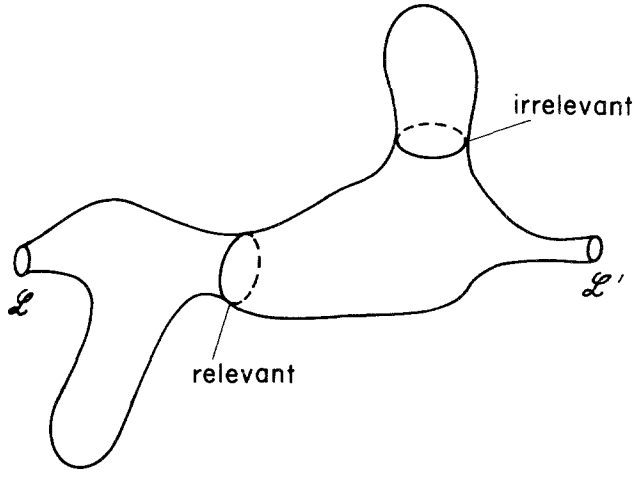


Fig. 6. Relevant and irrelevant loops in a surface with two boundary components $\mathscr{L}$ and $\mathscr{L}'$.

Otherwise, it is irrelevant. Denote by $\mathscr{M}_0(S)$ the collection of all relevant networks in S. As the notation suggests, $\mathscr{M}_0(S)$ will now play a role similar to the one played by the base network for surfaces with one boundary component. It is convenient to adopt the convention that the boundary loops $\mathscr{L}$ and $\mathscr{L}'$ are relevant. Note that relevant loops in S can be characterized by the property that any path in S that connects the two boundary components intersects all relevant loops but need not intersect an irrelevant loop.

Given a surface $S \in \mathscr{E}$ with $\partial S = \mathscr{L} \cup \mathscr{L}'$, it is not hard to see that there exists a unique minimal surface with boundary $\mathscr{L} \cup \mathscr{L}'$ whose edge network is $\mathscr{M}_0(S)$. This can be proven by a simple generalization of the argument leading to Lemma 7. We denote this minimal surface by $S(\mathscr{M}_0(S))$. Now we can write

$$G_{\underline{\beta}}(\mathscr{L}, \mathscr{L}') = \sum_{\mathscr{G}} \sum_{S:\mathscr{M}_0(S)=\mathscr{G}} \exp[-\hat{A}_{\underline{\beta}}(S(\mathscr{G}))]\, \mathscr{Z}_{\underline{\beta}}(\mathscr{G}) \tag{6.1}$$

where the sum over $\mathscr{G}$ runs over all families of relevant networks that can arise in surfaces with boundary $\mathscr{L} \cup \mathscr{L}'$ and $\mathscr{Z}_{\underline{\beta}}(\mathscr{G})$ is the sum over all excitations that can be attached to the faces of $S(\mathscr{G})$. Since

$$\mathscr{Z}_{\underline{\beta}}(\mathscr{G}) = \prod_{i=1}^{n} \mathscr{Z}_{\underline{\beta}}(F_i) \tag{6.2}$$

where $\{F_1, \ldots, F_n\}$ are the faces of $S(\mathscr{G})$, the analysis of Section 4 applies and we have

$$\mathscr{Z}_{\underline{\beta}}(\mathscr{G}) \leqslant \prod_{i=1}^{n} e^{P(F_i)} \tag{6.3}$$

with $P(F) = \log \mathscr{Z}_{\underline{\beta}}(F; 0)$; see (4.14).

Proposition 14. If

$$P(F) \leqslant p_2(\underline{\beta})\, |F| + \bar{p}_1\, |\partial F| + \bar{p}_0\, |C(F)| \tag{6.4}$$

for any face F, where $\bar{p}_0$ and $\bar{p}_1$ are constants, then there exist $\beta_1^* > 0$ and $\varepsilon > 0$ such that

$$m(\underline{\beta}) \geqslant \varepsilon(\beta_1 - \beta_1^*) \tag{6.5}$$

for $\beta_1 > \beta_1^*$.

Proof. Let $\mathscr{L}$ be any loop, and let $\mathscr{L}_x$ be the translate of $\mathscr{L}$ a distance x along a coordinate axis. If $\partial S = \mathscr{L} \cup \mathscr{L}_x$, then clearly $|E(S)| \geqslant 4x$. Hence, the action of any surface contributing to $G_{\underline{\beta}}(\mathscr{L}, \mathscr{L}_x)$ contains a term $4\beta_1 x$ and we shall prove that for β_1 sufficiently large, a fraction of this term

can be removed from the action without affecting the convergence of the sum defining $G_{\underline{\beta}}(\mathscr{L}, \mathscr{L}_x)$, leading to a bound of the form

$$G_{\underline{\beta}}(\mathscr{L}, \mathscr{L}_x) \leqslant \text{const} \cdot e^{-\varepsilon x} \tag{6.6}$$

which implies the desired result.

Let $\mathscr{G}$ be the collection of relevant networks in a surface S with $\partial S = \mathscr{L} \cup \mathscr{L}_x$. We can order the elements of $\mathscr{G}$, $E_1, E_2, \ldots, E_n$, starting at $\mathscr{L}$, such that if we cut S along a relevant loop in E_i, then $E_1, \ldots, E_{i-1}$ are in the same component as $\mathscr{L}$, while $E_{i+1}, \ldots, E_n$ are in the same component as $\mathscr{L}_x$. Let $S(\mathscr{G})$ denote the minimal surface spanning $\mathscr{G}$. Then E_{i+1} is connected to a hole in exactly one of the faces of E_i, since $S(\mathscr{G})$ is planar. Denote the boundary of this hole by γ_i. The face in S containing γ_i cannot carry any type 2 excitations. Note that $|\gamma_i| \leqslant \max\{|E_i|, |E_{i+1}|\} \equiv M_i$. We can think of $S(\mathscr{G})$ as a "tower" of n surfaces glued together along the loops $\gamma_1, \ldots, \gamma_{n-1}$. Given E_i, the number of possible γ_i is bounded by $M_i^2 3^{M_i}$. Hence, by Lemma 11, the sum

$$\sum_{\mathscr{G}} e^{-B_1|\mathscr{G}| - B_0 N(\mathscr{G})} \tag{6.7}$$

where

$$|\mathscr{G}| = \sum_{E_i \in \mathscr{G}} |E_i|, \qquad N(\mathscr{G}) = \sum_{E_i \in \mathscr{G}} N(E_i)$$

is convergent for B_1 sufficiently large.

For any $\underline{\beta} \in \mathscr{B}$, $\beta_2 - p_2(\underline{\beta}) \geqslant 0$, since otherwise the loop functions diverge, so using the bound (6.4), we obtain

$$\begin{aligned} G_{\underline{\beta}}(\mathscr{L}, \mathscr{L}_x) &\leqslant \sum_{\mathscr{G}} \exp[-(\beta_1 - \tfrac{1}{2}\bar{p}_1)|\mathscr{G}| - (\beta_0 - \tfrac{1}{2}\bar{p}_0) N(\mathscr{G})] \\ &\leqslant \text{const} \cdot \exp(-\varepsilon x) \end{aligned} \tag{6.8}$$

for $\beta_1 > \bar{p}_1/2 + \varepsilon/4$ sufficiently large. ∎

The above proposition is as close as we can get to proving Quasitheorem 6. The problem is that it is beyond the reach of our present understanding to prove bounds on the error term in (4.21) that are uniform in the shape of the region under consideration and depend linearly on the length of the boundary. We are, however, convinced that a bound of the form

$$|P(F) - p_2(\underline{\beta})\,|F|\,| \leqslant \text{const} \cdot |\partial F| \tag{6.9}$$

is valid in general.

It is not hard to see how to generalize the method of this section to decompose surfaces with three or more boundary components into a family of basic networks + excitations. Since we do not study the higher loop functions in this paper, we do not describe this in detail.

7. GENERALIZATIONS

In this section we briefly discuss the complications that arise in the RG procedure described in Sections 4–6 if one drops the condition that surfaces are simple. We do not describe in full detail how to deal with these complications, but discuss the most important problems which are encountered and the main ideas that go into their solution.

(i) *Branch points.* A branch point is a vertex j in a flat surface F such that $\#\partial^* j > 4$ if j is an interior vertex and $\#\partial^* j > 5$ if j is a boundary vertex.

The canonical example of a branch point is a flat surface F with the topology of a disc and a boundary $\partial F = \mathscr{L}$ which is a totally self-overlapping loop and has winding number 2 with respect to all its internal points (see Fig. 7); i.e., F looks like the Riemann surface of the square root function and consists of two sheets that overlap. The location of the branch point inside $\mathscr{L}$ is arbitrary, so there can be $|F|/2$ flat surfaces with boundary $\mathscr{L}$. It follows that a network which contains a self-overlapping loop can in general be spanned by many distinct minimal surfaces. The extra entropy of surfaces due to branch points can be controlled by using the increased number of corners in the networks of surfaces that contain branch points. This is in fact what is done in the combinatoric estimate of Section 8.

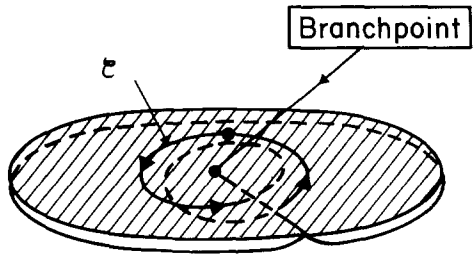


Fig. 7. A branch point of degree two in a surface S imbedded in $\mathbb{R}^3$. A curve $\mathscr{C}$ which winds once around the branch point in S winds twice around it in $\mathbb{R}^3$. Every plaquette in a neighborhood of the branch point is covered twice by the surface. Clearly, there exist branch points of an arbitrarily high degree.

(ii) *Tongues.* Let F be a flat surface with exterior boundary $\mathscr{L}$. A *tongue* in F is a subsurface of F that lies outside $\mathscr{L}$ (see Fig. 8). The problem with tongues is that they can have an arbitrarily large area for surfaces with a fixed exterior boundary. Clearly one should regard tongues as excitations of underlying minimal surfaces with no tongues, but there is no unique way of specifying where the tongues begin and the minimal surfaces end. This problem can be overcome by a cumbersome convention at the price of not having excitations separated from the underlying surface by edge links.

(iii) *Network excitations.* Consider a simple surface S with a network E. Suppose we cut S along one link $l \in E$ that lies in the boundary of two plaquettes p_1 and p_2. Take a large, flat disc D perpendicular to p_1 and p_2, cut D along one link that does not meet the boundary of D, and glue the cut D to the cut S as indicated in Fig. 9. We can imagine that D is the bottom of some cylinder T, so the surface S' obtained by gluing T to S has the same boundary as S and should be regarded as an excitation of S. In fact, the cut D is a type 2 face of S', spanning a hole of area 0. However, T is not in a natural way associated to any particular face of S, but rather to the network E. The entropy of excitations of this type is of course easily controlled by the action of the edge links and corners of E. See Fig. 9.

Note also that given a minimal surface S, we can disconnect the faces of S by excitations of the type we just described; see Fig. 10. Hence, the faces of a surface can coalesce when we shave off excitations. The entropy of excitations of this type is easily controlled by the activities of networks but makes the renormalization process more complicated.

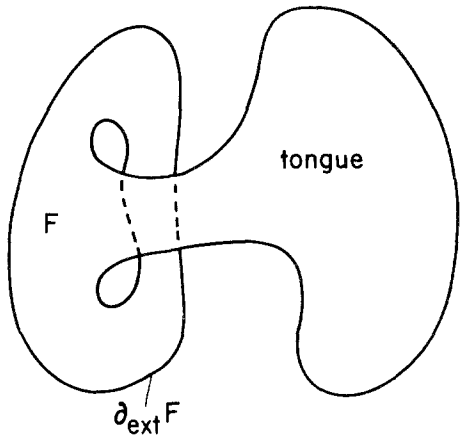


Fig. 8. A flat surface F with a tongue.

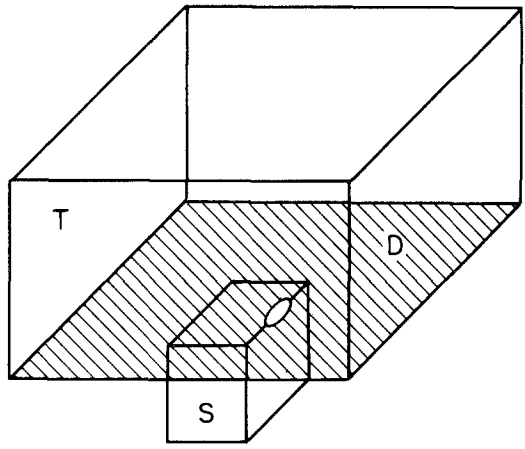


Fig. 9. The box T with bottom D is an excitation of S not associated with a particular face of S.

The three types of surface phenomena described above represent the most important obstacles encountered in extending the RG method to the full ensemble of PRS. These obstacles are technical rather than conceptual. If, however, we want to extend the RG method to surfaces with nonplanar topology, a host of conceptual problems arises, since the decomposition of the networks into a hierarchy breaks down. There is no simple generalization of our method of separating "excitations" from "minimal surfaces" in this case. We can of course use the techniques of this paper to deal with *planar* SA surfaces and prove upper bounds, which allow us to conclude that planar SA surfaces have qualitatively the phase diagram depicted in Fig. 1 and the mass is strictly positive for β_1 sufficiently large.

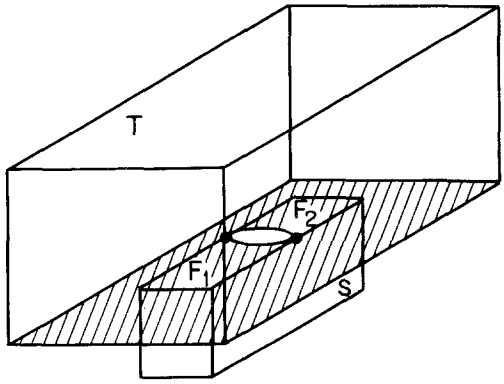


Fig. 10. The excitation T splits one of the face of S into two parts, F_1 and F_2.

8. COMBINATORIAL ESTIMATES

In this section we prove the bound (3.5), and we shall see that Lemma 11 easily follows from the proof. Let $n_{\mathscr{L}}(A, E, N)$ be defined by (3.4), i.e., $n_{\mathscr{L}}(A, E, N)$ equals the number of connected planar surfaces in $\mathbb{Z}^d$ with boundary $\mathscr{L}$, area A, E edge links, and at most N corners, counted with multiplicity. We prove the following.

Proposition 15. For any $\varepsilon > 0$ there exists a convex region $\mathscr{R}_\varepsilon \subseteq \mathbb{R}^2$ containing the half-lines

$$I_0 = \{(\beta^0, 0) \mid B^0(\varepsilon) < \beta^0 < \infty\}$$

and

$$I_1 = \{(0, \beta^1) \mid B^1(\varepsilon) < \beta^1 < \infty\}$$

for some finite $B^0(\varepsilon)$ and $B^1(\varepsilon)$ such that, for $(\beta^0, \beta^1) \in \mathscr{R}_\varepsilon$, we have

$$n_{\mathscr{L}}(A, E, N) \leqslant \text{const} \cdot e^{\varepsilon A + \beta^1 E + \beta^0 N} \tag{8.1}$$

Before proving the proposition, we make a few remarks. First, it is obvious that the subset $\mathscr{R}_\varepsilon$ of $\mathbb{R}^2$ consisting of those $(\beta^0, \beta^1) \in \mathbb{R}^2$ for which (8.1) holds is convex.

Second, if we define

$$n^\delta_{\mathscr{L}}(A) = \#\{S \in \mathscr{E} \mid \partial S = \mathscr{L}, |S| = A, N(S) \leqslant \delta A\} \tag{8.2}$$

and

$$m^\delta_{\mathscr{L}}(A) = \#\{S \in \mathscr{E} \mid \partial S = \mathscr{L}, |S| = A, E(S) \leqslant \delta A\} \tag{8.3}$$

then we claim that it suffices tro prove the bounds

$$b^\delta_{\mathscr{L}}(A) \leqslant \text{const} \cdot e^{f(\delta)A} \tag{8.4}$$

$$m^\delta_{\mathscr{L}}(A) \leqslant \text{const} \cdot e^{g(\delta)A} \tag{8.5}$$

where the functions f and g tend to zero as $\delta \downarrow 0$. Here, and elsewhere in this section, $\mathscr{E}$ denotes the ensemble (v) of planar random surfaces.

In order to prove the claim, note, first, that

$$|E(S)| \leqslant 2A + |\mathscr{L}|/2$$

for any $S \in \mathscr{E}$. Replacing δ by 3 in (8.3), we obtain

$$n_{\mathscr{L}}(A) = \#\{S \in \mathscr{E} \mid \partial S = \mathscr{L}, |S| = A\} \leqslant \text{const} \cdot e^{c_0 A} \tag{8.6}$$

for some finite constant c_0. A bound of the form (8.6) was first proven in ref. 16. Now, for a given $\varepsilon > 0$, we choose $\delta > 0$ such that $g(\delta) < \varepsilon$ and conclude from (8.5)–(8.6) that

$$\begin{aligned} n_{\mathscr{L}}(A, E, N) &\leqslant \#\{S \in \mathscr{E} \mid \partial S = \mathscr{L}, |S| = A, E(S) \leqslant E\} \\ &\leqslant \begin{cases} \text{const} \cdot e^{\varepsilon A} & \text{if } E/A \leqslant \delta \\ \text{const} \cdot e^{c_0 A} & \text{if } E/A > \delta \end{cases} \\ &\leqslant \text{const} \cdot e^{\varepsilon A + c_0 \delta^{-1} E} \end{aligned} \tag{8.7}$$

The bound (8.7) implies that the half-line I_1 in Proposition 15 is contained in $\mathscr{R}_\varepsilon$ if we choose $B^1(\varepsilon) = c_0 \delta^{-1}$. Similarly, by combining (8.4) and (8.6), we conclude that $I_0 \subseteq \mathscr{R}_\varepsilon$ for a suitable choice of $B^0(\varepsilon)$.

Clearly, $|E(S)| \geqslant N(S)$ for any surface S, so

$$m^{\delta}_{\mathscr{L}}(A) \leqslant n^{\delta}_{\mathscr{L}}(A) \tag{8.8}$$

Hence, it suffices to establish the bound (8.4) in order to prove the proposition.

The proof of (8.4) occupies the remainder of this section. It is convenient to introduce an ensemble of surfaces which satisfy the following three conditions:

(i) If v and v' are nearest neighbor edge vertices, i.e., v and v' both belong to an edge link and there is a link joining v and v', then the link joining v and v' is an edge link.

(ii) If v and v' are diagonally opposite edge vertices in the boundary of a plaquette in a surface, then the boundary of the plaquette contains exactly two edge links and these connect v and v'.

(iii) If v and v' are nearest neighbor vertices in a surface, then they are not both corners.

Note that it follows from (ii) that two opposite links in the boundary of a plaquette cannot both be edge links.

It may be seen as follows that we can assume that our surfaces satisfy the above constraints: Let $S \in \mathscr{E}$, with $\partial S = \mathscr{L}$, and $|S| = A$. Divide every plaquette in S into four equal smaller plaquettes, as indicated in Fig. 11. In

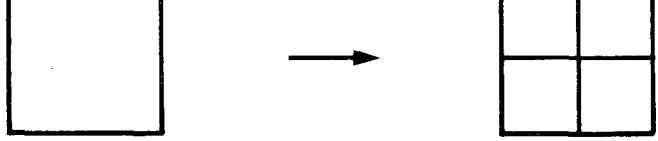


Fig. 11. Subdividing a plaquette into four smaller ones.

this fashion we obtain a surface with vertices in $\frac{1}{2}\mathbb{Z}^d$. Scaling all distances in $\frac{1}{2}\mathbb{Z}^d$ by a factor of 2, we obtain a surface $S' \in \mathscr{E}$, with $\partial S' = \mathscr{L}'$, $|S'| = 4A$, and $N(S') = N(S)$. Furthermore, it is clear that S' satisfies the constraints (i)–(iii). Thus, defining

$$n'^{\delta}_{\mathscr{L}} = \#\{S \in \mathscr{E} \mid \partial S = \mathscr{L},\ |S| = A,\ N(S) \leqslant \delta A,\ S \text{ fulfills (i)–(iii)}\} \tag{8.9}$$

we have shown that

$$n^{\delta}_{\mathscr{L}}(A) \leqslant n'^{\delta/4}_{\mathscr{L}'}(4A) \tag{8.10}$$

which in turn proves that it suffices to verify the bound (8.4) for $n'^{\delta}_{\mathscr{L}}(A)$ instead of $n^{\delta}_{\mathscr{L}}(A)$.

We therefore conclude that Proposition 15 has been established once we have proven the following lemma.

Lemma 16. There is a constant C and a positive function f, defined for $0 \leqslant \delta \leqslant 1$, such that

$$n'^{\delta}_{\mathscr{L}}(A) \leqslant Ce^{f(\delta)A} \tag{8.11}$$

and

$$f(\delta) \sim \delta |\ln \delta| \qquad \text{as} \quad \delta \to 0$$

Proof. For the sake of clarity, we first present the proof for a restricted class of surfaces whose faces do not overlap, i.e., in addition to the constraints (i)–(iii), we shall assume the following.

(iv) In each face of the surface no two plaquettes overlap (i.e., are copies of the same plaquette in $\mathbb{Z}^d$).

Subsequently, we explain what modifications are needed in order to deal with the general case.

We first describe the general line of argument. To each surface S satisfying (i)–(iv) with $\partial S = \mathscr{L}$, $|S| = A$, and $N(S) \leqslant \delta A$, we associate an ordered sequence $q_1, \ldots, q_A$ of (abstract) plaquettes some of whose $4A$ links are marked by an arrow and some of whose vertices are distinguished, by being colored red, say. We specify a process which allows us to determine the markings of the plaquette sequence that corresponds to a given surface and, at the same time, we bound the number of surfaces that can give rise to each sequence.

The surface S can be constructed by successively embedding the abstract plaquettes $q_1, \ldots, q_A$ into $\mathbb{Z}^d$ and gluing them together in a prescribed way, explained below. In this process, a marked link is either mapped into a link incident with a corner, such that the arrow points toward the corner, or it is the first edge link in some connected component of the edge network that we encounter in the construction of the surface.

Similarly, a colored vertex is the first vertex encountered in some connected component of the edge network.

It follows that the number of decorated links is $\leqslant 5\delta A$, and the number of colored vertices is $\leqslant \delta A$. Thus, the total number of sequences $q_1,..., q_A$ is at most

$$\sum_{i=j}^{5\delta A} \sum_{j=1}^{\delta A} \binom{4A}{i}\binom{4A}{j} 2^i \leqslant e^{C_1 \delta |\ln \delta| A} \tag{8.12}$$

for some constant C_1 (provided $\delta \leqslant 1/2$, say).

We shall also prove that at most $\exp(C_2 \delta |\ln \delta| A)$, $C_2 = \text{const}$, different surfaces, $S \in \mathscr{E}$, with $\partial S = \mathscr{L}$ and $|S| = A$, give rise to the same sequence $q_1,..., q_A$. In fact, the reconstruction of a surface S from a sequence $q_1,..., q_A$ requires at most $4A$ steps, as we explain below, and each step is uniquely determined, except when a plaquette is glued onto an edge link emerging from a corner, in which case there are at most $2d-3$ possibilities. The number of steps in which this ambiguity arises is bounded by $2N(S)$, and this proves the desired bound (8.11).

We now describe in detail the procedure by which a given surface S, with $|S| = A$, is constructed by successively embedding plaquettes $q_1,..., q_A$ into $\mathbb{Z}^d$ and gluing together overlapping links. At the same time, we show how the corners and links of the plaquettes $q_1,..., q_A$ are marked so that S is also uniquely determined by the marked q_i. The basic idea in this inductive construction of S is that one face is constructed at a time and once a boundary component of a face is encountered, all plaquettes containing links from that boundary component are added to the surface. In the construction of each face, all steps (except the first one) will be uniquely determined by the markings of the q_i. In each step we glue together two overlapping links or add one plaquette and glue it to the surface along one link.

Let $S \in \mathscr{E}$ satisfy conditions (i)–(iv), $\partial S = \mathscr{L}$, $|S| = A$. Below, we shall order the plaquettes in S, $p_1,..., p_A$. We begin marking the plaquettes by coloring all boundary links green. If p_i contains a corner and l is a link in ∂p_i containing this corner, then we label l by an arrow pointing toward the corner. Note that this marking is well defined, since, by (iii), l can meet at most one corner.

After performing these markings, we proceed to construct a sequence $S_1,..., S_{k(S)}$, $1 \leqslant k(S) \leqslant 4A$, of connected, nondecreasing subsurfaces of S such that S_1 consists of one plaquette and $S_{k(S)} = S$. With each surface S_n in this sequence we must associate some data: Denote the boundary components of S_n by $\mathscr{L}_n^0, \mathscr{L}_n^1,..., \mathscr{L}_n^{l(n)}$, $l(n) \geqslant 0$.

The number of vertices in $\mathscr{L}_n^i$, $i = 0,..., l(n)$, is denoted by $k(n, i)$ and the vertices are ordered consistently with a chosen orientation of S. Thus,

to specify the ordering it is enough to specify the first vertex in each component. We denote the vertices in $\mathscr{L}_n^i$ by $0_1^{n,i},\ldots, 0_{k(n,i)}^{n,i}$ in increasing order. Similarly, the links in $\mathscr{L}_n^i$ are denoted by $l_1^{n,i},\ldots, l_{k(n,i)}^{n,i}$, where $l_j^{n,i}$ connects $0_j^{n,i}$ and $0_{j+1}^{n,i}$ (with $0_{k(n,i)+1}^{n,i}=0_1^{n,i}$). The construction of the surface S_n is such that the boundary loops $\mathscr{L}_n^1,\ldots, \mathscr{L}_n^{l(n)}$ consists entirely of edge links. There are numbers $s(n,i)\in\{0,1,\ldots, k(n,i)-1\}$ for $i=1,\ldots, l(n)$ such that the links $l_1^{n,i},\ldots, l_{s(n,i)}^{n,i}$ are either boundary links (i.e., colored green) or have to be glued to other links present in $\mathscr{L}_n^i$, but $l_j^{n,i}$ is not to be glued to $l_{j+1}^{n,i}$ for $j=1,\ldots, s(n,i)$. On the other hand, $\mathscr{L}_n^0$ is either empty or there exist numbers $r_1(n)$, $r_2(n)\in\{1,\ldots, k(n,i)\}$, with $r_1(n)\leqslant r_2(n)$, such that the vertices $0_{r_1(n)}^{n,0},\ldots, 0_{r_2(n)}^{n,0}$ are not edge vertices, while the other vertices in $\mathscr{L}_n^0$ (i.e., $0_1^{n,0},\ldots, 0_{r_1(n)-1}^{n,0}$ and $0_{r_2(n)+1}^{n,0},\ldots, 0_{k(n,i)}^{n,0}$) are edge vertices. Furthermore, there is a nonnegative integer $s(n)\leqslant r_1(n)$ such that the links $l_1^{n,0},\ldots, l_{s(n)}^{n,0}$ are either boundary links or have to be glued to other links in $\mathscr{L}_n^0$, but $l_j^{n,0}$ is not to be glued to $l_{j+1}^{n,0}$ for $j=1,\ldots, s(n)$. Note that by assumption (i) the links $l_{r_1(n)-1}^{n,0},\ldots, l_{r_2(n)}^{n,0}$ (with $l_0^{n,0}=l_{k(n,0)}^{n,0}$) are not edge links while the other links in $\mathscr{L}_n^0$ are edge links. In the following the notion of a subsurface S_n of S also includes an ordering of the boundary loops, as above, together with an assignment of numbers $s(n,i)$, $i=1,\ldots, l(n)$, and $s(n)$, $r_1(n)$, $r_2(n)$.

We now explain how the subsurfaces S_n are constructed inductively. At the same time, we define an ordering of the plaquettes in S. First, we define S_1: Choose an orientation of S and fix some oriented link $l\in\mathscr{L}$. By definition, S_1 equals the oriented plaquette p_1 in S that contains l in its boundary. Furthermore, the ordering of the vertices in ∂S_1 is fixed by letting the endpoint of l be the first vertex $0_1^{1,0}$. Clearly, $l(1)=0$ and $k(1,0)=4$. The last link $l_4^{1,0}=l$ is an edge link, while the second link is not an edge link, by assumption (ii), and hence the vertices $0_2^{1,0}$ and $0_3^{1,0}$ are not both edge vertices by assumption (i). Moreover, $0_2^{1,0}$ (resp. $0_3^{1,0}$) is an edge vertex exactly if $l_1^{1,0}$ (resp. $l_3^{1,0}$) is an edge link by assumption (i).

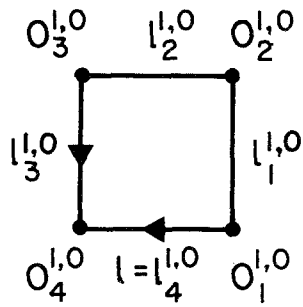


Fig. 12. An illustration of S_1 consisting of one plaquette. $0_4^{1,0}$ is a corner and l is a boundary link.

Thus, by changing the orientation of S, if necessary, we may assume that $0_2^{1,0}$ is not an edge vertex. Clearly, $l_3^{1,0}$ is an edge link exactly if $0_4^{1,0}$ is a corner and hence $l_3^{1,0}$ marked by an arrow. Thus we set $s(1)=0$, $r_1(1)=2$, and we set $r_2(1)=3$ if $l_3^{1,0}$ is not marked by an arrow whereas we set $r_2(1)=2$ if $l_3^{1,0}$ is marked by an arrow. This finishes the construction of S_1 (see Fig. 12).

Assume now that S_n has been constructed. There are three cases to consider:

1. $\mathscr{L}_n^0$ is nonempty and $(r_1(n), r_2(n)) \neq (1, k(n, 0))$, i.e., there is at least one edge vertex in $\mathscr{L}_n^0$.
2. $\mathscr{L}_n^0$ is nonempty and $(r_1(n), r_2(n)) = (1, k(n, 0))$, i.e., there are no edge vertices in $\mathscr{L}_n^0$.
3. $\mathscr{L}_n^0$ is empty, i.e., all links in ∂S_n are edge links.

We construct the subsurfaces S_{n+1} as follows.

Case 1. Consider the link $l_{r_1(n)-1}^{n,0}$ emerging from $0_{r_1(n)-1}^{n,0}$ in positive direction along $\mathscr{L}_n^0$. Since this is not an edge link, while its starting point is an edge vertex, by the construction of S_n, the only link in ∂S_n to which $l_{r_1(n)-1}^{n,0}$ can possibly be glued is $l_{r_2(n)}^{n,0}$, since this is the only other link in ∂S_n whose boundary contains exactly one edge vertex. Since the plaquettes in S_n containing $l_{r_1(n)-1}^{n,0}$ and $l_{r_2(n)}^{n,0}$, respectively, clearly belong to the same face in S, it follows from (iv) that $l_{r_1(n)-1}^{n,0}$ and $l_{r_2(n)}^{n,0}$ have to be glued together exactly if they are overlapping with opposite orientation. Thus we must distinguish between two subcases (see Fig. 13):

1a. $l_{r_1(n)-1}^{n,0}$ and $l_{r_2(n)}^{n,0}$ do not overlap with opposite orientation.

1b. $l_{r_1(n)-1}^{n,0}$ and $l_{r_2(n)}^{n,0}$ overlap with opposite orientation.

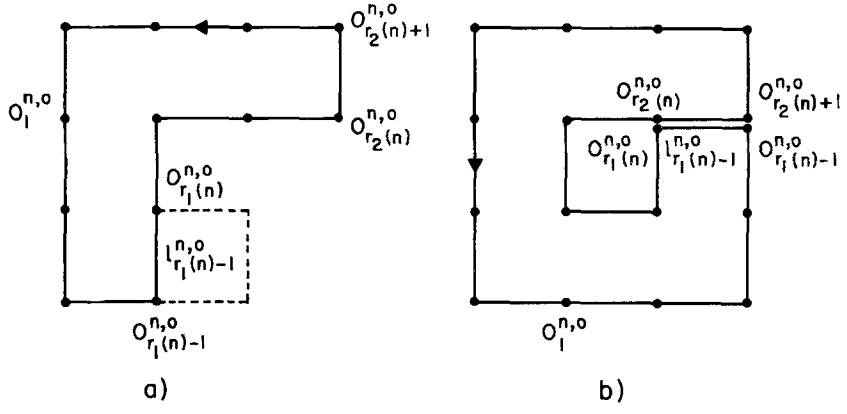


Fig. 13. (a) Case 1a; (b) case 1b.

Case 1a. In this case a new plaquette has to be glued onto S_n along $l^{n,0}_{r_1(n)-1}$. If the last plaquette added to S_n was number i, the new plaquette is number $i+1$. Since $l^{n,0}_{r_1(n)-1}$ is not an edge link, this plaquette has to lie in the same plane and has the same orientation as the plaquette in S_n onto which it is glued, i.e., the plaquette in S_n containing the link $l^{n,0}_{r_1(n)-1}$. This defines S_{n+1}. Obviously, the number of boundary components in S_{n+1} is the same as in S_n and, as we shall see, the component $\mathscr{L}^0_{n+1}$ containing the points in $\mathscr{L}^0_n$ contains both edge vertices and non-edge vertices. Thus, we have $l(n+1)=l(n)$, and we set $\mathscr{L}^i_{n+1}=\mathscr{L}^i_n$ and $s(n+1,i)=s(n,i)$, $i=1,\ldots,l(n+1)$, keeping the ordering of the vertices. The ordering of the vertices in $\mathscr{L}^0_{n+1}$ is chosen such that $0^{n,0}_1$ is the first vertex in $\mathscr{L}^0_{n+1}$.

We next determine $r_1(n+1)$, $r_2(n+1)$, $s(n+1)$ using only the information encoded in the markings on the links. Since clearly $0^{n,0}_i=0^{n+1,0}_i$ for $i=1,\ldots,r_1(n)-1$, we set $s(n+1)=s(n)$. The new vertices in $\mathscr{L}^0_{n+1}$ are $0^{n+1,0}_{r_1(n)}$ and $0^{n+1}_{r_1(n)+1}$ and the new links are $l^{n+1,0}_{r_1(n)-1}$, $l^{n+1,0}_{r_1(n)}$, $l^{n+1,0}_{r_1(n)+1}$. Of these, $l^{n+1,0}_{r_1(n)+1}$ is not an edge link, since that would contradict assumption (ii). It also follows from (ii) that $0^{n+1,0}_{r_1(n)+1}$ is an edge vertex if and only if both $l^{n+1,0}_{r_1(n)-1}$ and $l^{n+1,0}_{r_1(n)}$ are edge links. In that case, $0^{n+1,0}_{r_1(n)}$ is a corner, so that $l^{n+1,0}_{r_1(n)-1}$ and $l^{n+1,0}_{r_1(n)}$ are marked by an arrow, and we set $r_1(n+1)=r_1(n)+2$ and $r_2(n+1)=r_2(n)+2$. If, on the other hand, $0^{n+1,0}_{r_1(n)+1}$ is not an edge vertex, then $l^{n+1,0}_{r_1(n)}$ is not an edge link, and we have the two possibilities that $0^{n+1,0}_{r_1(n)}$ is an edge vertex or it is not. If $l^{n+1,0}_{r_1(n)-2}$ is an edge link, i.e., if $r_2(n)-r_1(n)<k(n,0)-2$, we can decide which is the case by looking only at the markings of the links in ∂S_{n+1}, as follows: If $l^{n+1,0}_{r_1(n)-2}=l^{n,0}_{r_2(n)-2}$ has no arrow or has an arrow pointing away from $0^{n+1,0}_{r_1(n)-1}$, then $0^{n+1,0}_{r_1(n)-1}$ is not a corner, and we may conclude that $l^{n+1,0}_{r_1(n)-1}$ is an edge link, because otherwise $0^{n+1,0}_{r_1(n)-1}$ is a branch point, which is excluded by assumption (iv). If, on the other hand, $l^{n+1,0}_{r_1(n)-2}$ has an arrow pointing toward $0^{n+1,0}_{r_1(n)-1}$, then $0^{n+1,0}_{r_1(n)-1}$ is a corner, and $l^{n+1,0}_{r_1(n)-1}$ is an edge link exactly if it is marked by an arrow. Thus, we set $r_1(n+1)=r_1(n)+1$ and $r_2(n+1)=r_2(n)+2$ if $l^{n+1,0}_{r_1(n)-2}$ is not marked by an arrow or if it is marked by an arrow pointing away from $0^{n+1,0}_{r_1(n)-1}$ or if $l^{n+1,0}_{r_1(n)-1}$ and $l^{n+1,0}_{r_1(n)-2}$ are marked by arrows pointing toward $0^{n+1,0}_{r_1(n)-1}$. Otherwise, we set $r_1(n+1)=r_1(n)$ and $r_2(n+1)=r_2(n)+2$.

Finally, if $l^{n+1,0}_{r_1(n)-2}$ is not an edge link, then there is only one edge vertex in $\mathscr{L}^0_n$, namely $0^{n,0}_{r_1(n)-1}$. This is clearly the first vertex we encounter in some connected component of the edge network of S, and we color it red. If $l^{n+1,0}_{r_1(n)-1}$ is an edge link, it is the first link encountered in some connected component of the edge network of S, and we color it (and the corresponding link in $q_{i(n+1)}$) red. Otherwise, S_{n+1} is defined as above, according to whether $l^{n+1,0}_{r_12(n)-1}$ is an edge link or not.

This finishes the construction of S_{n+1} if $l^{n,0}_{r_1(n)-1}$ and $l^{n,0}_{r_2(n)}$ do not overlap with opposite orientation.

Case 1b. As mentioned above, $l^{n,0}_{r_1(n)-1}$ has to be glued to $l^{n,0}_{r_2(n)}$, by assumption (iv). After doing this, we obtain a surface S_{n+1} with $l(n+1) \equiv l(n)+1$ boundary components, since $\mathscr{L}^0_n$ splits into two pieces $\mathscr{L}^0_{n+1}$ and $\mathscr{L}^{l(n)+1}_{n+1}$ whose vertices are $0^{n,0}_{r(n)}$ (which is identified with $0^{n,0}_{r_2(n)}$), $0^{n,0}_{r_1(n)+1}$,..., $0^{n,0}_{r_2(n)-1}$, and $0^{n,0}_1$,..., $0^{n,0}_{r_1(n)-1}$ (which is identified with $0^{n,0}_{r_2(n)+1}$), $0^{n,0}_{r_2(n)+2}$,..., $0^{n,0}_{k(n,0)}$, respectively. In particular, $\mathscr{L}^0_{n+1}$ contains no edge vertices, while $\mathscr{L}^{l(n)+1}_{n+1}$ consists entirely of edge vertices. Letting $0^{n,0}_{r_1(n)}$ be the first vertex in $\mathscr{L}^0_{n+1}$ and $0^{n,0}_1$ [or $0^{n,0}_{k(n,0)}$ if $r_1(n)=1$] be the first vertex in $\mathscr{L}^{l(n)+1}_{n+1}$, we may set $r_1(n+1)=1$, $r_2(n+1)=k(n+1,0)=r_2(n)-r_1(n)$, $s(n+1)=0$, and $s(n+1, l(n+1))=s(n)$. Moreover, we let $\mathscr{L}^i_{n+1}=\mathscr{L}^i_n$, for $i=1,\ldots, l(n)$, and keep the ordering of the vertices. We have disregarded the case $r_1(n)=r_2(n)$ or $r_2(n)-r_1(n)=k(n,0)-1$. In the first case, $\mathscr{L}^0_n$ is empty, and in the second case $\mathscr{L}^{l(n)+1}_{n+1}$ is empty, and we should put $l(n+1)=l(n)$.

This finishes the construction of S_{n+1} in case 1b.

Case 2. In this case there is no edge vertex in $\mathscr{L}^0_n$. Again we distinguish between two subcases:

2a. Each link in $\mathscr{L}^0_n$ overlaps another link in $\mathscr{L}^0_n$ with opposite orientation.

2b. There is a link in $\mathscr{L}^0_n$ which does not overlap any other link in $\mathscr{L}^0_n$ with opposite orientation.

Case 2a. From assumption (iv) and the fact that the plaquettes in S_n containing a link or vertex in $\mathscr{L}^0_n$ belong to the same face in S, it follows that each link in $\mathscr{L}^0_n$ overlaps a unique link in $\mathscr{L}^0_n$ with opposite orientation and that these links have to be glued together. Hence, there is a unique way of gluing the links in $\mathscr{L}^0_n$ pairwise together. After doing this we obtain the surface S_{n+1} with boundary components $\mathscr{L}^1_{n+1}=\mathscr{L}^1_n$,..., $\mathscr{L}^{l(n)}_{n+1}=\mathscr{L}^{l(n)}_n$. Thus, $l(n+1)=l(n)$, and $\mathscr{L}^0_{n+1}$ is empty. Furthermore, we keep the ordering of the vertices in $\mathscr{L}^1_{n+1}$,..., $\mathscr{L}^{l(n+1)}_{n+1}$ and set $s(n+1, i)=s(n, i)$, $i=1,\ldots, l(n+1)$. This finishes the construction of S_{n+1} in case 2a.

Case 2b. Consider the first link l in $\mathscr{L}^0_n$ which does not overlap any other link in $\mathscr{L}^0_n$ with opposite orientation. Then a new plaquette has to be glued onto l as in case 1a. Thus, $l(n+1)=l(n)$ and the loops $\mathscr{L}^0_{n+1}$, $\mathscr{L}^1_{n+1}$,..., $\mathscr{L}^{l(n+1)}_{n+1}$ are, together with the ordering of vertices in $\mathscr{L}^1_{n+1}$,..., $\mathscr{L}^{l(n+1)}_{n+1}$ and $s(n+1, i)$, $i=1,\ldots, l(n+1)$, defined as in case 1a. Clearly, the vertices in $\mathscr{L}^0_{n+1}$ are those of $\mathscr{L}^0_n$, together with two new nearest neighbor vertices, P and Q, in the new plaquette. If neither P nor Q is an edge vertex, then there are no edge vertices in $\mathscr{L}^0_{n+1}$, and we define $0^{n,0}_1$ to be the first vertex in $\mathscr{L}^0_{n+1}$ and set $r_1(n+1)=1$, $r_2(n+1)=k(n+1,0)=k(n,0)+2$, and $s(n+1)=0$. If exactly one of the new vertices P and Q is

an edge vertex, then it is the first edge vertex in some connected component of the edge network of S, and we color it red and define it to be the first vertex in $\mathscr{L}^0_{n+1}$. Furthermore, we set $r_1(n+1)=2$, $r_2(n+1)=k(n+1,0)=k(n,0)+2$, and $s(n+1)=0$.

Finally, if both P and Q are edge vertices, then, by assumption (i), the link connecting P and Q is an edge link and, clearly, it it is first edge link encountered in some connected component of the edge network of S. We then color this link red and let its endpoint be the first vertex in $\mathscr{L}^0_{n+1}$. Moreover, we set $r_1(n+1)=2$, $r_2(n+1)=k(n+1,0)-1=k(n,0)+1$, and $s(n+1)=0$. This finishes the construction of S_{n+1} in case 2b.

Case 3. If all links in ∂S_n are colored green, it is clear that $S_n=S$, and the construction is finished. Assume that this is not the case, and let i_0 be the smallest index i, $1\leqslant i\leqslant l(n)$, such that $\mathscr{L}^i_n$ contains at least one link which is not colored green. We then let l_0 denote the link $l^{n,i_0}_{s(n,i_0)+1}$ if $s(n,i_0)>0$. If $s(n,i_0)=0$, we let $l_0=l^{n,i_0}_{j_0}$, where $0^{n,i_0}_{j_0}$ is the first corner in $\mathscr{L}^{i_0}_n$ if $\mathscr{L}^{i_0}_n$ contains a corner, whereas if $\mathscr{L}^{i_0}_n$ contains no corner, we let $l_0=l^{n,i_0}_{j_0}$, where j_0 is the smallest index j, $1\leqslant j\leqslant k(n,i_0)$, such that l^{n,i_0}_j and l^{n,i_0}_{j+1} are overlapping with opposite orientation. The existence of such a j_0 in this case is easily verified.

We now once more thave to distinguish between two subcases:

3a. l_0 is not incident with a corner, i.e., l_0 carries no arrow.

3b. l_0 is incident with a corner.

Case 3a. Assume first that $0<s(n,i_0)<k(n,i_0)-1$. Then the link $l^{n,i_0}_{s(n,i_0)}$ is either green or has to be glued to another link, $l'\neq l_0$, in $\mathscr{L}^{i_0}_n$ by definition of $s(n,i_0)$. Suppose, in the latter case, that $l'=l^{n,i_0}_{k_0}$. Then, since l_0 is not incident with a corner, it follows that l_0 is colored green if $l^{n,i_0}_{s(n,i_0)}$ is; otherwise, it has to be glued to the link $l^{n,i_0}_{k_0-1}$ (see Fig. 14).

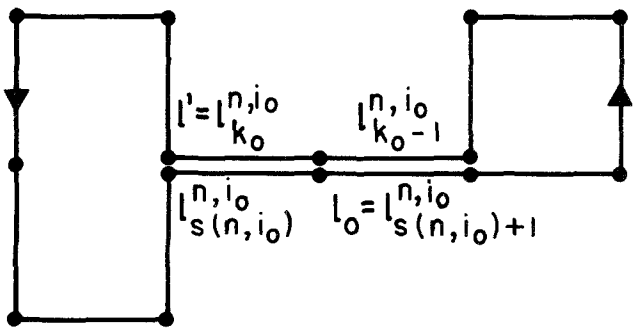


Fig. 14. Occurrence of case 3a.

If l_0 is colored green or if $k_0 - 1 \neq s(n, i_0) + 2$, mod $k(n, i_0)$, (i.e., if l_0 and $l^{n,i_0}_{k_0-1}$ are not neighboring links in $\mathscr{L}^{i_0}_n$), we define S_{n+1} to be equal to S_n, except that $s(n+1, i_0)$ is defined to be $s(n, i_0)+1$, which is less than $k(n, i_0) = k(n+1, i_0)$. If $k_0 - 1 = s(n, i_0) + 2$, we let S_{n+1} be the surface obtained by gluing l_0 to $l^{n,i_0}_{k_0-1}$. Clearly, S_{n+1} has $l(n) = l(n+1)$ boundary components $\mathscr{L}^{i_0}_{n+1}$ and $\mathscr{L}^{i}_{n+1} = \mathscr{L}^{i}_n$, for $i \neq i_0$ and $1 \leqslant i \leqslant l(n+1)$. The vertices in $\mathscr{L}^{i_0}_{n+1}$ are those of $\mathscr{L}^{i_0}_n$, except $0^{n,i_0}_{k_0-1}$, and $0^{n,i_0}_{s(n,i_0)+1}$ has been identified with $0^{n,i_0}_{k_0}$. We let $0^{n,i_0}_1$ be the first vertex in $\mathscr{L}^{i_0}_{n+1}$ and set $s(n+1, i_0) = s(n, i_0) - 1$. The ordering of the other boundary components is kept, and, of course, $s(n+1, i) = s(n, i)$, for $i \neq i_0$ and $1 \leqslant i \leqslant l(n)$.

If $s(n, i_0) = 0$, then it follows from assumption (iv) that $l_0 = l^{n,i_0}_{j_0}$ has to be glued to $l^{n,i_0}_{j_0+1}$, and we define S_{n+1} as described above for the case $k_0 - 1 = s(n, i_0) + 2$, the only difference being that $s(n+1, i_0)$ is defined to be zero. If $k(n, i_0) = 2$, then $\mathscr{L}^{i_0}_{n+1}$ is empty, in which case we have to relabel the boundary components in the obvious way.

Finally, if $s(n, i_0) = k(n, i_0) - 1$, we claim that $l^{n,i_0}_{s(n,i_0)+1} = l^{n,i_0}_{k(n,i_0)}$ has to be glued to l^{n,i_0}_1. To see this, we first note that $l^{n,i_0}_{k(n,i_0)}$ is either colored green or has to be glued to another link in $\mathscr{L}^{i_0}_n$. If it was green or had to be glued to a link different from l^{n,i_0}_1, we would conclude that any link in $\mathscr{L}^{i_0}_n$ was either colored green or had to be glued to a nonneighboring link in $\mathscr{L}^{i_0}_{n_0}$. But this contradicts the fact that S is planar with one boundary component, as an easy application of Euler's formula (2.6) shows. Thus, $l^{n,i_0}_{k(n,i_0)}$ has to be glued to l^{n,i_0}_1.

Gluing these two links together, we obtain a subsurface $S^{(1)}_n$ of S with boundary components $\mathscr{L}^1_n, \ldots, \mathscr{L}^{i_0-1}_n, \mathscr{L}^{i_0+1}_n, \ldots, \mathscr{L}^{l(n)}_n$, and $\mathscr{L}^{(1)}_n$, where the links in $\mathscr{L}^{(1)}_n$ are those of $\mathscr{L}^{i_0}_n$, except $l^{n_0}_1$ and $l^{n,i_0}_{k(n,i_0)}$. By the argument just presented, we conclude that either $\mathscr{L}^{(1)}_n$ consists only of green links, in which case $\mathscr{L}^{(1)}_n = \mathscr{L}$, or l^{n,i_0}_2 has to be glued to $l^{n,i_0}_{k(n,i_0)-1}$. Gluing these links together, we obtain the surface $S^{(2)}_n$. We may continue in this manner until we obtain a subsurface S_{n+1} with boundary components $\mathscr{L}^1_{n+1} = \mathscr{L}^1_n, \ldots, \mathscr{L}^{i_0-1}_{n+1} = \mathscr{L}^{i_0-1}_n$, $\mathscr{L}^{i_0+1}_{n+1} = \mathscr{L}^{i_0+1}_n, \ldots, \mathscr{L}^{l(n+1)}_{n+1} = \mathscr{L}^{l(n)}_n$, and $\mathscr{L}^{i_0}_{n+1} = \mathscr{L}$. We choose an arbitrary ordering of $\mathscr{L}$, set $s(n+1, i_0) = 0$, keep the ordering of the other loops, and set $s(n+1, i) = s(n, i)$, for $i \neq i_0$, and $1 \leqslant i \leqslant l(n+1) = l(n)$.

This completes the construction of S_{n+1} in case 3a.

Case 3b. If the link l_0 is green, we define S_{n+1} to be equal to S_n except that $s(n+1, i_0) = s(n, i_0) + 1$.

Assume now that l_0 is not green. Then there are two possibilities: Either a new plaquette has to be glued onto l_0, and this plaquette has to be orthogonal or parallel with opposite orientation to the plaquette in S_n containing l_0, or it has to be glued to another link in $\mathscr{L}^{i_0}_n$. Note that l_0

cannot be glued to a link in another boundary component of S_n than $\mathscr{L}_n^{i_0}$, since that would give rise to a handle in S, as is easily seen by using Euler's formula (2.6).

If a new plaquette is glued onto l_0 we obtain a surface S_{n+1} with $l(n)$ boundary components. The vertices in one of these are those of $\mathscr{L}_n^{i_0}$, together with two new neighbor vertices in the new plaquette. We call this component $\mathscr{L}_{n+1}^0$ and let $0_1^{n,i_0}$ be the first vertex in it. The other boundary components are $\mathscr{L}_n^1,\ldots,\ \mathscr{L}_n^{i_0-1},\ \mathscr{L}_n^{i_0+1},\ldots,\ \mathscr{L}_n^{l(n)}$. Then the new vertices are $0_{s(n,i_0)+2}^{n+1,0}$ and $0_{s(n,i_0)+3}^{n+1,0}$. It follows from assumptions (i) and (ii) that not both of these vertices are edge vertices and that $0_{s(n,i_0)+2}^{n+1,0}$ (resp. $0_{s(n,i_0)+3}^{n+1,0}$) is an edge vertex exactly if $l_{s(n,i_0)+1}^{n+1,0}$ (resp. $l_{s(n,i_0)+3}^{n+1,0}$) is an edge link. Clearly, this is equivalent to stating that $0_{s(n,i_0)+1}^{n+1,0}$ (resp. $0_{s(n,i_0)+4}^{n+1,0}=0_{s(n,i_0)+2}^{n,i_0}$) is a corner, or that $l_{s(n,i_0)+1}^{n+1,0}$ (resp. $l_{s(n,i_0)+3}^{n+1,0}$) carries an arrow. If $0_{s(n,i_0)+2}^{n+1,0}$ (resp. $0_{s(n,i_0)+3}^{n+1,0}$) is an edge vertex, we set $r_1(n+1)=r_2(n+1)=s(n,i_0)+3$ [resp. $=s(n,i_0)+2$], whereas, if neither of the new vertices is an edge vertex, we set $r_1(n+1)=s(n,i_0)+2$ and $r_2(n+1)=s(n,i_0)+3$. In both cases we set $s(n+1)=s(n,i_0)$. Moreover, we set $\mathscr{L}_{n+1}^i=\mathscr{L}_{n+1}^i$ and $s(n+1,i)=s(n,i)$, for $i=1,\ldots,i_0-1$, whereas $\mathscr{L}_{n+1}^i=\mathscr{L}_n^{i+1}$ and $s(n+1,i)=s(n,i-1)$, for $i=i_0,\ldots,\ l(n+1)=l(n)-1$.

Finally, if l_0 is to be glued to another link in $\mathscr{L}_n^{i_0}$, we define S_{n+1} as described above in case 3a, depending on whether it is glued to the subsequent link or not.

This completes the constructions of S_{n+1}.

Let us note that the number of links decorated by an arrow is $2N(S)$, the number of red links is bounded by the number of connected components in the edge network of S, which is clearly bounded by $N(S)/2$, and the same bound holds for the number of colored vertices. The number of green links is of course $|\mathscr{L}|$. Thus, since $N(S)\leqslant\delta A$, it follows, as mentioned earlier, that the total number of ways to color and mark the links in $q_1,\ldots,q_A$ is bounded by $e^{c|\delta\ln\delta|A}$ as $A\to\infty$ for some constant c. On the other hand, as we have been, the procedure described above to reconstruct S from $q_1,\ldots,q_A$ us unambiguous, except at those steps where case 3b above arises, and in which case there are *a priori* $2d-1$ possibilities, namely either to glue onto l_0 one of the $2(d-1)-1$ possible plaquettes, or to glue l_0 onto the subsequent link in $\mathscr{L}_n^{i_0}$, or to glue l_0 onto a link in $\mathscr{L}_n^{i_0}$ which is not a neighboring link in $\mathscr{L}_n^{i_0}$. Furthermore, the total number of steps required to reconstruct S is easily seen to be bounded by $4A$, and the number of times case 3b may arise is clearly bounded by $2N(S)$. From this the bound (8.11) follows.

It remains to describe the modifications needed if assumption (iv) is dropped. We used this assumption in cases 1a, 1b, and 2a. In case 1b we used it to conclude that $l_{r_1(n)-1}^{n,0}$ and $l_{r_2(n)}^{n,0}$ have to be glued together if they

overlap with opposite orientation. It is not hard to check that the number of times this possibility arises is bounded by $2N(S)$, and each time we have two possibilities: to glue $l^{n,0}_{r_1(n)-1}$ to $l^{n,0}_{r_2(n)}$, or to glue on a new plaquette in the same plane and with same orientation as the plaquette in S_n containing $l^{n,0}_{r_1(n)-1}$. Hence it is clear that this ambiguity only modifies f by a constant multiple.

In case 1a we used assumption (iv) to exclude branch points, i.e., vertices shared by more than four plaquettes in the same face. If, for any vertex i, we define

$$Q_i = \begin{cases} \sigma_i - 4 & \text{if } i \text{ is an interior vertex} \\ \sigma_i - 3 & \text{if } i \text{ is a boundary vertex} \end{cases}$$

where σ_i is the order of i, then it follows from Euler's formula (2.6) that, for a planar surface S with b boundary components,

$$\sum_{i \in S} Q_i = 4(b-2) \tag{8.13}$$

Applying this formula to each face in S, it is easy to verify that the number of plaquettes in S containing a branch point is bounded by $40N(S)$. Hence, we can deal with the presence of branch points by coloring them blue, say, and adopting the convention that, once a branch point is encountered, all the plaquettes containing this vertex are glued on at once.

Finally, we used assumption (iv) in case 2a to conclude that if $\mathscr{L}^0_n$ consists only of non-edge links, with each link overlapping some other link in $\mathscr{L}^0_n$ with opposite orientation, then the links in $\mathscr{L}^0_n$ have to be glued pairwise together and this gluing is unique. If we allow overlapping plaquettes in a face, this does not necessarily hold anymore. In particular, a given link in $\mathscr{L}^0_n$ may overlap several other links in $\mathscr{L}^0_n$ with opposite orientation. This problem may be overcome by using the fact that if $\mathscr{L}_0$ is an oriented loop lying in some 2-plane in $\mathbb{Z}^d$, then it is the boundary of at most one flat, bounded, oriented surface without branch points. In order to apply this result to $\mathscr{L}^0_n$, we must ensure that $\mathscr{L}^0_n$ is the boundary of a piece of a face containing no branch points. This can be ensured as follows: In case the face in S containing $\mathscr{L}^0_n$ has no boundary components other than $\mathscr{L}^1_n, \ldots, \mathscr{L}^{l(n)}_n$ and no branch points other than those contained in S_n, there is a unique way of finishing the construction of this face by gluing the unique bounded flat surface with boundary $\mathscr{L}^0_n$ onto S_n, as mentioned above. If, on the contrary, the face containing $\mathscr{L}^0_n$ contains more boundary components than $\mathscr{L}^1_n, \ldots, \mathscr{L}^{l(n)}_n$, or it contains other branch points than those contained in S_n, there exists a (nonunique) straight strip of plaquettes which connects $\mathscr{L}^0_n$ to either a new boundary component or to a new branch point. We then choose to color the link in $\mathscr{L}^0_n$ to which this strip is glued yellow, say. Thus, once case 2a is encountered and $\mathscr{L}^0_n$ contains yellow links, we obtain S_{n+1} by successively gluing plaquettes in a strip

starting at the first of those yellow links, until a blue vertex or a red vertex or link appears. We then choose this colored (red or blue) vertex or the endpoint of the red link to be the first one in the boundary component $\mathscr{L}^0_{n+1}$ containing it and keep the ordering of $\mathscr{L}^1_n \equiv \mathscr{L}^1_{n+1}, \ldots, \mathscr{L}^{l(n)}_n \equiv \mathscr{L}^{l(n+1)}_{n+1}$. We are then no longer in case 2a, and the construction may proceed. This modification of the construction only requires coloring yellow a number of links which is bounded by the sum of the number of branch points and the number of boundary components in the faces of S. Hence, it is bounded by $4N(S)$, which is easily seen, and it follows that the desired bound (8.11) can still be maintained.

This completes the proof of Lemma 16. ▌

Remark. For the bound (8.11) to hold it is clearly not necessary to keep $\mathscr{L}$ fixed, but sufficient to require that $\mathscr{L}$ contain some fixed link or point in $\mathbb{Z}^d$, as the proof shows.

Let now $(E, \mathscr{L})$ be a connected edge network with base loop $\mathscr{L}$, and let $R(E, \mathscr{L})$ be a surface with boundary $\mathscr{L}$, whose edge network is E. Clearly, $(E, \mathscr{L})$ is determined by the part of $R(E, \mathscr{L})$ consisting of plaquettes containing edge vertices. It is easy to verify that the number of such plaquettes is $\leqslant C_1 |E|$ for some constant C_1 independent of E. Note also that the only stages in the construction of a given surface described above at which a plaquette without edge vertices is added is when case 2a or case 2b arises. Thus, in order to estimate

$$n^\delta(K) = \#\{(E, \mathscr{L}) \mid x \in \mathscr{L},\ |E| = K,\ N(E) \leqslant \delta |E|\}$$

we may start with a sequence of $\leqslant C_1 K$ abstract plaquettes, mark them as previously, and apply the same construction, except that each time case 2 arises, i.e., $\mathscr{L}^0_n$ contains no edge vertices, we leave $\mathscr{L}^0_n$ out of further consideration, or, alternatively, we just glue an arbitrary flat piece of surface (possibly containing branch points) with boundary $\mathscr{L}^0_n$ to S_n. In this way we obtain

$$n^\delta(K) \leqslant \text{const} \cdot e^{f(\delta)K} \tag{8.14}$$

where $f(\delta)$ is as in Lemma 16.

Given $\beta_1 > 0$, choose δ_0 small enough such that $f(\delta_0) < \beta_1$. Then, for fixed $x \in \mathbb{Z}^d$,

$$\begin{aligned}
&\sum_{\substack{(E,\mathscr{L}) \\ \mathscr{L} \ni x}} e^{-\beta_1 |E| - \beta_0 N(E)} \\
&\quad \leqslant \sum_{\substack{(E,\mathscr{L}),\ \mathscr{L} \ni x \\ N(E) \leqslant \delta_0 |E|}} e^{-\beta_1 |E|} + \sum_{\substack{(E,\mathscr{L}),\ \mathscr{L} \ni x \\ N(E) > \delta_0 |E|}} e^{-(\beta_1 + \delta_0 \beta_0)|E|} \\
&\quad \leqslant \text{cst} \cdot \sum_{K=1}^{\infty} e^{-[\beta_1 - f(\delta_0)]K} + \text{cst} \cdot \sum_{K=1}^{\infty} e^{-[\beta_1 + \delta_0 \beta_0 - f(1)]K}
\end{aligned}$$

where we have used (8.14) with $\delta = \delta_0$, and that

$$\#\{(E, \mathscr{L}) \mid x \in \mathscr{L}, |E| = K\} = n^1(K) \leqslant \text{cst} \cdot e^{f(1)K}$$

since $N(E) \leqslant |E|$. Thus, it follows that

$$\sum_{(E,\mathscr{L}),\mathscr{L} \ni x} e^{-\beta_1 |E| - \beta_0 N(E)} < \infty$$

provided $\beta_1 + \beta_0 \delta_0 > f(1)$, which proves Lemma 11.

ACKNOWLEDGMENT

The work of T.J. was supported in part by the Icelandic Science Foundation.

REFERENCES

1. D. R. Nelson, T. Piran, and S. Weinberg, eds. *Statistical Mechanics of Membranes and Surfaces* (World Scientific, Singapore, to be published).
2. A. M. Polyakov, *Nucl. Phys. B* **268**:406 (1986); D. C. Brydges, B. Durhuus, J. Fröhlich, and C. Giffen, *Nucl. Phys. B* **275**[FS17]:459 (1986), and references therein; F. David and E. Guitter, *Europhys. Lett.* **3**:1169 (1987); P. Olesen and S.-K. Yang, *Nucl. Phys. B* **283**:73 (1987); E. Braten, R. D. Pisarksi, and S.-M. Tse, *Phys. Rev. Lett.* **58**:93 (1987); A. M. Polyakov, Gauge fields and strings, (Harwood, Chur, London 1987).
3. J. Ambjørn, B. Durhuus, and T. Jónsson, *Phys. Rev. Lett.* **58**:2619 (1987).
4. J. Ambjørn, B. Durhuus, J. Fröhlich, and T. Jónsson, *Nucl. Phys. B* **290**[FS20]:480 (1987).
5. B. Durhuus, J. Fröhlich, and T. Jónsson, *Nucl. Phys. B* **257**[FS14]:779 (1985).
6. B. Durhuus, J. Fröhlich, and T. Jónsson, *Nucl. Phys. B* **240**[FS12]:453 (1984).
7. B. Durhuus, J. Fröhlich, and T. Jónsson, *Phys. Lett.* **137B**:93 (1984).
8. B. Durhuus and T. Jónsson, *Phys. Lett.* **108B**:385 (1986).
9. T. Jónsson, *Commun. Math. Phys.* **106**:679 (1986).
10. W. Helfrich, *J. Phys.* (Paris) **46**:1263 (1985); A. M. Polyakov, *Nucl. Phys. B* **268**:406 (1986); L. Peliti and S. Leibler, *Phys. Rev. Lett.* **54**:1690 (1985); D. Förster, *Phys. Lett.* **114A**:115 (1986); H. Kleinert, *Phys. Lett.* **174B**:335 (1986); F. David, *Europhys. Lett.* **2**:577 (1986); R. D. Pisarski, *Phys. Rev. Lett.* **58**:1300 (1987); F. Alonso and D. Espriu, *Nucl. Phys. B* **283**:393 (1987).
11. J. Fröhlich and T. Spencer, *Commun. Math. Phys.* **81**:527 (1981).
12. M. Wortis, R. Pandit, and M. Schick, in *Melting, Localization and Chaos*, R. K. Kalia and P. Vashishta, eds. (Elsevier, Houston, Texas, 1982); K. Binder, in *Phase Transitions and Critical Phenomena*, Vol. 8, C. Domb and J. L. Lebowitz, eds. (Academic Press, 1983); M. E. Fisher, *J. Stat. Phys.* **34**:667 (1984); P.-G. De Gennes, *Rev. Mod. Phys.* **57**:827 (1985); D. B. Abraham, in *Phase Transitions and Critical Phenomena*, Vol. 10, C. Domb and J. L. Lebowitz, eds.
13. J. Bricmont, A. El Mellouki, and J. Fröhlich, *J. Stat. Phys.* **42**:743 (1987); J. Fröhlich and C.- E. Pfister, *Europhys. Lett.* **3**:845 (1987); *Commun. Math. Phys.* **109**:493, **112**:51 (1987).

14. H. Kawai and Y. Okamoto, *Phys. Lett.* **130B**:415 (1983).
15. J. Fröhlich, in *Applications of Field Theory to Statistical Mechanics*, L. Garido, ed. (Springer-Verlag, Berlin, 1985).
16. B. Durhuus, J. Fröhlich, and T. Jónsson, *Nucl. Phys. B* **225**:185 (1983).
17. B. Baumann and B. Berg, *Phys. Lett.* **164B**:131 (1985); B. Baumann, *Nucl. Phys. B* **285**[FS19]:391 (1987).
18. J. Fröhlich, in *Proceedings of the 1984 Les Houches Summer School*, K. Osterwalder and R. Stora, eds. (Plenum Press, New York, 1986).
19. D. Ruelle, *Statistical Mechanics, Rigorous Results* (Benjamin, New York, 1969).
20. E. Seiler, *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics* (Springer-Verlag, Berlin, 1982).
21. C. Cammarota, *Commun. Math. Phys.* **85**:517 (1982); R. Kotecky and D. Preiss, *Commun. Math. Phys.* **103**:491 (1986).
22. M. Aizenman, J. T. Chayes, L. Chayes, J. Fröhlich, and L. Russo, *Commun. Math. Phys.* **92**:19 (1983).