



ELSEVIER

Nuclear Physics B 570 [FS] (2000) 590–614

NUCLEAR
PHYSICS B

www.elsevier.nl/locate/npe

Intersecting loop models on \mathbb{Z}^d : rigorous results

L. Chayes^a, Leonid P. Pryadko^b, Kirill Shtengel^c

^a *Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, USA*

^b *School of Natural Sciences, IAS, Princeton, NJ 08540, USA*

^c *Department of Physics, UCLA, Los Angeles, CA 90095-1547, USA*

Received 20 October 1999; accepted 15 November 1999

Abstract

We consider a general class of (intersecting) loop models in d dimensions, including those related to high-temperature expansions of well-known spin models. We find that the loop models exhibit some interesting features – often in the “unphysical” region of parameter space where all connection with the original spin Hamiltonian is apparently lost. For a particular $n = 2$, $d = 2$ model, we establish the existence of a phase transition, possibly associated with divergent loops. However, for $n \gg 1$ and arbitrary d there is no phase transition marked by the appearance of large loops. Furthermore, at least for $d = 2$ (and n large) we find a phase transition characterised by broken translational symmetry. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 05.50.+q; 64.60.-i; 75.10.Hk

Keywords: Loop models; Reflection positivity; Phase transitions

1. Introduction

In recent years there has been much interest in various loop models. Loop models are graphical models defined by drawing closed loops along the bonds of the underlying lattice. The loops may come in n different flavours (colours). No two loops can share a bond, while sharing a vertex is generally allowed. Explicitly, the bond configurations are such that each vertex houses an even number – possibly zero – of bonds of each colour. Each loop configuration is assigned a “weight” that depends on the number of participating vertices of each type. In the cases of interest these weights are actually positive hence, at least in finite volume, they define a *probability measure* on the set of

E-mail address: shtengel@physics.ucla.edu (K. Shtengel).

all loop configurations. Thus, for a finite lattice the loop partition function may be written as

$$Z = \sum_{\mathcal{G}} R^b \nu_1^{m_1} \nu_2^{m_2} \dots \nu_V^{m_V}, \quad (1.1)$$

with the sum running over all allowed loop configurations \mathcal{G} . Here b is the total number of participating bonds, m_i ($i = 1, \dots, V$) is the number of vertices of type i and ν_i is the corresponding vertex factor.¹ This definition is slightly different from the one typically found in literature (cf. Refs. [1,2]) since it also includes the bond fugacity R . Although strictly speaking it is not needed (since the bond fugacity can always be incorporated into the vertex factors), we find it convenient to keep R as a separate parameter. We remark that by relabeling the empty bonds as an additional colour, these models may be formally regarded as “fully packed”.

The reason loop models have been extensively studied is because they appear quite naturally as representations (often approximate) of various statistical-mechanical models. These include, among others, the Ising model (this approach dates back to Kramers and Wannier [3] and was later used to solve the model exactly [4,5]), the Potts model (polygon expansion [6]), $O(n)$ spin models [7–11], 1-D quantum spin models [12], a supersymmetric spin chain [13], the q -colouring problem [14,15] and polymer models [16,17].

Here we consider the loop models explicitly related to the high-temperature expansions of the standard $O(n)$, corner-cubic (AKA diagonal-cubic) and face-cubic spin models. This is, in fact, the same set of models that was treated in Ref. [7]. However, in this paper, we provide a careful treatment of the large n cases – and we treat the standard d -dimensional lattices. As a result, we arrive at quite unexpected results concerning the behaviour of these models in the high fugacity region.

In particular, despite the considerable attention the subject has received, most authors (with certain exceptions, e.g. Refs. [3–5,9,13]) chose to consider models where only loops of *different* colours are allowed to cross each other (if at all). On the other hand, spin systems (in the high-temperature approximation) naturally generate self-intersecting loops. In order to avoid this issue, an exorbitant amount of work has been done on lattices with coordination number $z = 3$ (e.g. the honeycomb lattice), where loop intersections simply cannot occur. Overall this approach appears to be justified since one is usually interested in the critical properties of the underlying spin systems. Indeed, consider the archetypal n -component spin system with $|\mathbf{S}_i| \equiv 1$ and let us write $\exp(\lambda \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j) \sim \prod_{\langle i,j \rangle} (1 + \lambda \mathbf{S}_i \cdot \mathbf{S}_j)$. Although as a spin system the right-hand side makes strict sense only if $|\lambda| \leq 1$ (the “physical regime”), the associated loop model turns out to be well defined for all λ . Since the systems can be identified for $|\lambda| \ll 1$ it can be argued that the critical properties of the spin system and those of the loop model are the same and are independent of the underlying lattice.

¹ Many authors consider an additional factor of the form $F_1^{l_1} F_2^{l_2} \dots F_n^{l_n}$ where F_i is a “loop fugacity” and l_i is the number of loops of the i th colour. Although the objects l_i are unambiguous when self-intersections are forbidden, in the general case they are not easily defined. Nevertheless, the essence of such a term – at least in the case of integer F_i – is captured by the introduction of additional colours.

Notwithstanding, for $n \gg 1$ any phase transition in the actual spin system is not anticipated until temperatures of order $1/n$ (i.e. $\lambda \sim n$), which we note is well outside the physical regime of the loop model. At first glance this appears to be borne out: the natural parameter in the loop model (as well as in the spin system) seems to be λ/n . Thus, the loop model could, in principle, capture the essential features of the spin system up to – and including – the critical point.

We have found such a picture to be overly optimistic. Indeed, depending on the specific details, e.g. the lattice structure, there may be a phase transition in the region $1 \ll |\lambda| \ll n$ (specifically, $\lambda \sim n^{3/4}$), well outside the physical regime but well before the validity of the approximation was supposed to break down. Furthermore, it would seem that both the temperature scale and the nature of the transition (not to mention the existence of the transition) depend on such details. Finally, we shall demonstrate that in contrast to their spin system counterparts, the large- n models have *no* phase transition – for any value of bond fugacity – associated with the formation of large loops (i.e. divergent loop correlations).

The structure of this paper is as follows. Section 2 is dedicated to the description of the spin models and their connection to the loop models. Specific results for those models with the two-dimensional spin variable ($n = 2$) are presented in Section 3. Finally, Section 4 contains the discussion of reflection positivity as well as some results concerning phase transitions in the large n case.

2. n-component models

2.1. O(n) model

Let us start by considering the O(n) model on some finite lattice $\Lambda \subset \mathbb{Z}^d$ defined by the following partition function:

$$Z = \text{Tr} \prod_{\langle i,j \rangle} (1 + \lambda \mathbf{S}_i \cdot \mathbf{S}_j) \quad (2.1)$$

with $\mathbf{S}_i \in \mathbb{R}^n$, $|\mathbf{S}_i| = 1$ and Tr denoting normalised summation (integration) over all possible spin configurations. The corresponding loop model is readily obtained along the lines of a typical “high-temperature” expansion. We write $\mathbf{S}_i \cdot \mathbf{S}_j = S_i^{(1)} S_j^{(1)} + \dots + S_i^{(n)} S_j^{(n)}$ and define n different colours (each associated with a coordinate direction of the O(n)-spins). Expanding the product, we have n choices for each bond plus a possibility of a vacant bond. Thus, various terms are represented by n -coloured bond configurations: $\mathcal{E} = (\mathcal{E}_1, \dots, \mathcal{E}_n)$ with \mathcal{E}_ℓ denoting those bonds where the term $S_i^{(\ell)} S_j^{(\ell)}$ has been selected. Clearly, the various \mathcal{E}_ℓ ’s are pairwise (bond) disjoint. Thus, for each \mathcal{E} we obtain the weight

$$W_{\mathcal{E}} = \text{Tr} \prod_{\langle i,j \rangle \in \mathcal{E}_1} \lambda S_i^{(1)} S_j^{(1)} \dots \prod_{\langle i,j \rangle \in \mathcal{E}_n} \lambda S_i^{(n)} S_j^{(n)}. \quad (2.2)$$

On the basis of elementary symmetry considerations it is clear that $W_{\mathcal{E}} \neq 0$ if and only if each vertex houses an even number (which could be zero) of bonds of each colour. Once this constraint is satisfied, we get an overall factor of $\lambda^{b(\mathcal{E})}$ – with $b(\mathcal{E})$ being the total number of participating bonds – times the product of the *vertex factors* obtained by

performing the appropriate $O(n)$ integrals. The details and results of these calculations are presented in Appendix A. In general, it is seen that the vertex factors depend only on the number of participating colours and the number of bonds of each colour emanating from a given vertex, i.e. not on the particular colours that were involved nor on the directions of these bonds.

In the case of a square lattice ($d = 2$) we have only three main types of (non-empty) vertices: those where two bonds of the same colour join together, those with two pairs of bonds of two different colours and those with four bonds of the same colour. These have weights of $1/n$, $1/n(n + 2)$ and $3/n(n + 2)$ correspondingly. Rescaling the bond fugacity from $R = \lambda$ to $R = \lambda/n$ we arrive at the vertex weights $\nu_1 = 1$, $\nu_2 = n/(n + 2)$ and $\nu_3 = 3n/(n + 2)$.

The factor of 3 relating ν_2 to ν_3 has an interesting interpretation (which, as shown in Appendix A, turns out to be quite general). Indeed, each vertex of the third type may be decomposed into three different vertices as shown in Fig. 1. Each of the new vertices is now assigned equal weight, which is also that of ν_2 . We thus split each \mathcal{G} into 3^{m_3} different graphs – each of equal weight – in which every vertex with four bonds now provides explicit instructions relating outgoing and incoming directions of an individual walk.

Hence in every such graph (now defined with the walking instructions encoded at every vertex) the individual loops are well defined. Furthermore, changing the colour of any loop does not change the weight of the graph. Thus we may write

$$Z = \sum_{\mathcal{G}} \left(\frac{\lambda}{n}\right)^{b(\mathcal{G})} \left(\frac{n}{n + 2}\right)^{m(\mathcal{G})} n^{\ell(\mathcal{G})}, \tag{2.3}$$

where the summation now takes place over all configurations \mathcal{G} of colourless loop graphs in which every vertex housing four bonds is resolved by “walking instructions”, and ℓ is the number of such loops (being now defined completely unambiguously). In addition to the advantages of a manifestly colourless expression, the above permits continuation to non-integer n .

We conclude this subsection with the following series of remarks and observations.

- As shown in Appendix A, such vertex decomposition works, in fact, for an arbitrary lattice in an arbitrary number of spatial dimensions (with the proper weights for vertices housing 6, 8, etc., bonds).
- Notice that only $|\lambda| \leq 1$ region of the parameter space is “physical” (in a sense of the underlying Hamiltonian: $-\beta H = \sum_{\langle i,j \rangle} \ln[1 + \lambda \mathbf{S}_i \cdot \mathbf{S}_j]$), while for $|\lambda| > 1$ one presumes, no spin Hamiltonian can be written at all. The corresponding loop model, however, makes perfect sense in the entire parameter space.

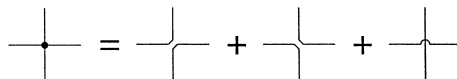


Fig. 1. Decomposition of a type 3 vertex into three new vertices in the two-dimensional $O(n)$ loop model.

- If we consider a 2D XY model ($n = 2$), we notice that the factor $2^{\ell(\mathcal{S}^1)}$ in Eq. (2.3) can be obtained by assigning directions to the colourless loops. The above decomposition of type 3 vertices makes this procedure unambiguous. Having done that, we can turn this model into a random surface model by assigning heights to the plaquettes in such a way that a plaquette to the right of a directed bond is always one step higher than the plaquette to the left. Not surprisingly, this random surface model turns out to be identical to the one obtained by the standard means of Fourier-transforming the original weights in Eq. (2.1) (cf. Refs. [18,19]).
- Finally, it is worth mentioning that in the fully packed limit $R \rightarrow \infty$, the $n = 2$ loop model on the square lattice (with the colour degrees of freedom being replaced by assigning directions to the loops) turns out to be nothing but the square ice model (i.e. the six-vertex model with all six weights being equal – see Ref. [6] for a definition of this model). The mapping between the vertices of these models is shown in Fig. 2. We remark that the perspective of the ice model (and, for that matter, other six-vertex models) as a two colour loop model provides additional flexibility in the analysis of these systems. These issues will be pursued in a future publication.

2.2. Corner-cubic model

We now consider the following “discretised” modification of the above $O(n)$ model (given by Eq. (2.1)):

$$Z = \text{Tr} \prod_{\langle i,j \rangle} \left[1 + \frac{\lambda}{n} (\sigma_i^{(1)} \sigma_j^{(1)} + \sigma_i^{(2)} \sigma_j^{(2)} \dots + \sigma_i^{(n)} \sigma_j^{(n)}) \right] \tag{2.4}$$

with $\sigma_i^{(k)} = \pm 1$. For small values of λ this model may be viewed as a high-temperature limit of a corner-cubic model. Indeed, it describes an interaction of the type in Eq. (2.1) where spins \mathbf{S}_i are allowed to point at the corners of an n -dimensional hypercube (with the origin being placed at the centre of the cube).

Mapping it onto an n -colour loop model is almost identical to the $O(n)$ case, with the only difference being the vertex factor: $(\sigma_i^{(1)})^{2k_1} \dots (\sigma_i^{(n)})^{2k_n} = 1$. We can choose to associate the weight of $R = \lambda/n$ with each bond, thus making all vertex weights ν_i to be equal to unity. In other words, the resulting loops in this model do not interact with

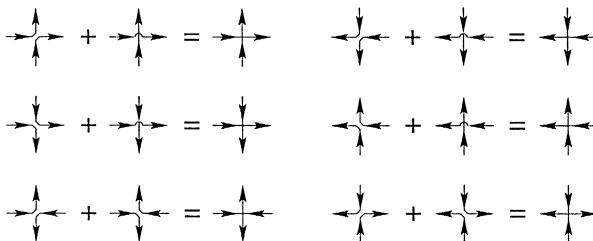


Fig. 2. Mapping of a fully packed $O(2)$ model onto the square ice model.

each other via vertices (there is still a hard core bond repulsion, however). The partition function is then simply

$$Z = \sum_{\mathcal{G}} \left(\frac{\lambda}{n} \right)^{b(\mathcal{G})}. \tag{2.5}$$

2.3. Face-cubic model

Finally, let us examine a different model with cubic symmetry given by the following partition function:

$$Z = \text{Tr} \prod_{\langle i,j \rangle} \left[1 + \lambda (u_i^{(1)} u_j^{(1)} + u_i^{(2)} u_j^{(2)} \dots + u_i^{(n)} u_j^{(n)}) \right]. \tag{2.6}$$

Here $u_i^{(k)} = 0, \pm 1$, and for a given site i exactly one of $u_i^{(k)}$ ($k = 1, 2, \dots, n$) has a non-zero value. In fact, one may think of u 's as components of an n -dimensional unit vector that is only allowed to point along the coordinate axes (or from the centre to the faces of an n -dimensional hypercube – thus the name face-cubic). While the corner-cubic model described earlier had 2^n degrees of freedom per site (the number of corners of a hypercube), the present model has only $2n$ such degrees of freedom (the number of faces).

Once again, the corresponding loop model is obtained by performing multiplication in Eq. (2.6) and then summing the resulting terms over all possible values of u 's. But since for each site i only one of the spin components $u_i^{(k)} \neq 0$ at a time, we notice that no terms that mix different k 's are allowed. In terms of resulting loops this means hard-core repulsion of different colours: only loops of the same colour can share a vertex. The vertex factors are now: zero for any vertex with multiple colours and $1/n$ for vertices with two or more bonds of the same colour; the bond fugacity is given by $R = \lambda$.

3. Results for the $n = 2$ case

3.1. The $n = 2$ models with cubic symmetry, Ashkin–Teller and random surface models

In this section we shall restrict our attention to the models with cubic symmetries. Firstly, let us slightly change the notations for convenience: let $\sigma_i^{(1)} = \sigma_i$ and $\sigma_i^{(2)} = \tau_i$ for the corner-cubic, while $u_i^{(1)} = u_i$ and $u_i^{(2)} = v_i$ for the face-cubic model. The corresponding partition functions are then written as

$$Z_{\text{CC}} = \text{Tr} \prod_{\langle i,j \rangle} \left[1 + \frac{\lambda}{2} (\sigma_i \sigma_j + \tau_i \tau_j) \right], \tag{3.1}$$

$$Z_{\text{FC}} = \text{Tr} \prod_{\langle i,j \rangle} \left[1 + \lambda (u_i u_j + v_i v_j) \right]. \tag{3.2}$$

While we have seen that the loop models generated by these partition functions are very different, the spin models themselves turn out to be identical. Indeed, Eq. (3.2) is

obtained from Eq. (3.1) by the following transformation: $u_i = (\sigma_i + \tau_i)/2$, $v_i = (\sigma_i - \tau_i)/2$. This is equivalent to a 45° rotation in the spin space (along with a $\sqrt{2}/2$ rescaling) and is very specific to the $n = 2$ case. In turn, both models are equivalent to the Ashkin–Teller model [20] with a particular choice of parameters that will be detailed in Section 3.3.

The two-colour loop models generated by Eq. (3.1) and Eq. (3.2) are given by the following sets of parameters in Eq. (1.1): $R = \lambda/2$, with all vertices having weight one, and $R = \lambda$, with all multi-colour vertices given weight zero and all other non-empty vertices given weight one-half respectively.

Turning our attention to the particular case of two spatial dimensions, we remark that in the former model one can sum over all possible colourings to obtain the following result for the partition function:

$$Z = \sum_{\mathcal{G}'} \left(\frac{\lambda}{2}\right)^{b(\mathcal{G}')} 2^{f(\mathcal{G}')} \tag{3.3}$$

with $b(\mathcal{G}')$ being the total number of occupied (colourless) bonds and $f(\mathcal{G}')$ being the total number of *faces* in the clusters they form. The number of faces is the minimum number of bonds that one must remove in order for the remaining clusters to be tree-like. For example, the cluster in Fig. 3 has five *faces*, while it can at most consist of four *loops*. Curiously enough, this result appears to have no simple generalisation for $n > 2$.

It appears that the two-dimensional loop model derived from the corner-cubic model can not be mapped directly onto a random surface model. However, the other loop representation (the one obtained via expansion of the face-cubic model) does correspond to a random surface model. Indeed, consider the following ‘‘recipe’’: take a loop configuration generated by a particular term in the expansion of Eq. (3.2). Let red be the colour of loops originating from u ’s, while blue corresponds to v ’s. Take all plaquettes at the outermost region to be at height zero (these plaquettes are said to form a substrate). On this substrate we have clusters of loops. The outermost boundaries of these clusters are themselves closed loops. The plaquettes immediately adjacent to these boundaries are assigned the height of $+1$ if the loop forming a boundary is red, or -1 if it is blue. These plaquettes, along with any other plaquette accessible from them without crossing coloured bonds are said to form a plateau and thus have the same

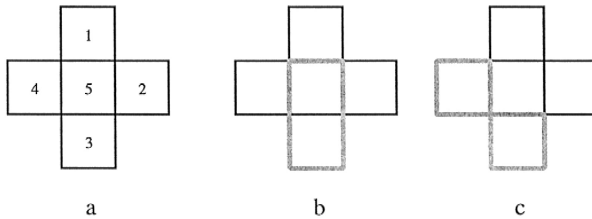


Fig. 3. A fragment of a possible two-dimensional loop configuration for an intersecting loop model. Three different possible colourings (out of the total of 32 in the case of $n = 2$) for a loop model of *corner-cubic* type are represented here as (a), (b) and (c). For the *face-cubic* type, only monotonous clusters like (a) remain allowed.

height.² Inside such a plateau region there may be other loop clusters, that may or may not touch the boundary of a plateau (only corners are allowed to touch, since no bond sharing between the loops is possible). Every such cluster is now treated in the same way: its boundary defines the “secondary” plateau with the height being that of the “primary” plateau ± 1 depending on the colour of the boundary. This procedure is repeated until all plaquettes are assigned their heights. As an example, consider the cluster in Fig. 3a, which may now only consist of the bonds of a single colour. If this were a red cluster, the heights would be $+1$ for the plaquettes 1–4 and $+2$ for the plaquette 5.

In fact, this description is essentially identical to that given in Ref. [21] in the context of wetting transition with the only difference that we allow for two-coloured clusters instead of single-coloured, and therefore the heights in our case may be both positive and negative.³

The important feature of this random surface model is that it must have a phase transition whenever the underlying Ashkin–Teller model undergoes a transition.

3.2. Random cluster representation

Let us derive yet another graphical representation for the $n = 2$ (Ashkin–Teller) model considered in the previous section, this time it will be a *random cluster* representation closely resembling the FK representation for the Potts model. We start from Eq. (2.4) (with $n = 2$), and with the help of the identity $\sigma_i \sigma_j = 2 \delta_{\sigma_i \sigma_j} - 1$ rewrite it as follows:

$$Z \propto \sum_{\sigma} \prod_{\langle i,j \rangle} \left[1 + v (\delta_{\sigma_i \sigma_j} + \delta_{\tau_i \tau_j}) \right] \quad (3.4)$$

with $v = \lambda/(1 - \lambda)$. The random cluster representation is generated by evaluating the product over all bonds in Eq. (3.4) and then summing over the possible values of σ 's and τ 's. If we think of the bonds originating from the σ variables as green (g), and the bonds originating from the τ variables as orange (o), then each of the resulting terms in the partition function can be graphically represented as a collection of green and orange clusters as well as empty sites. The clusters of different colours may share sites, but not

² The strict definition is as follows: plaquettes A and B are said to belong to the same plateau if and only if there exists an unbroken path along the bonds of a *dual* lattice that connects the centre of plaquette A to the centre of plaquette B without crossing a single coloured bond of the direct lattice.

³ This random surface model, however, has a few important differences with those of a more conventional type (like the one obtained for the O(2) case). Firstly, due to an irreducible four-leg vertex factor, it cannot be described by a nearest-neighbour Hamiltonian (i.e. a Hamiltonian that depends only on the height difference of neighbouring plaquettes). Secondly, since no directions are assigned to the loops separating the plateaux, there is no way of deciding on the *sign* of their relative height difference without going through the necessary construction steps starting from the *outside*. By contrast, the O(2)-related random surface model can be constructed starting from *any* plaquette – a particular choice simply determines the overall additive constant. In this sense the mapping between the present random surface and the loop model is *non-local*.

bonds. Denoting the configurations of green and orange bonds as ω_g and ω_o respectively, we can then write the partition function as

$$Z \propto \sum_{\omega} v^{b(\omega_g)} v^{b(\omega_o)} 2^{c(\omega_g)} 2^{c(\omega_o)} \tag{3.5}$$

with $b(\omega)$ being the total number of bonds (of a specified colour), and $c(\omega)$ being the number of corresponding connected components. The rule for counting the connected components is as follows: every site that is not a part of a green cluster is considered to be a separate connected component for the purposes of $c(\omega_g)$, even if this site is a part of an orange cluster, and vice versa. In particular, the quantities $v^{b(\omega_g)} 2^{c(\omega_g)}$ and $v^{b(\omega_o)} 2^{c(\omega_o)}$ are to be interpreted exactly as in the usual random cluster models.

3.3. Self-duality and criticality at $\lambda = 1$

The duality relations for such random cluster representation of the standard AT model in two dimensions were established in Refs. [22,23]. Firstly, let us write the generic AT Hamiltonian as

$$-\beta H = \sum_{\langle i,j \rangle} \left[K(\delta_{\sigma_i \sigma_j} + \delta_{\tau_i \tau_j}) + L \delta_{\sigma_i \sigma_j} \delta_{\tau_i \tau_j} \right]. \tag{3.6}$$

The graphical representation for the partition function is then obtained along the lines of the previous section. The only difference is that this time double-coloured (i.e. green and orange at the same time) bonds are also allowed. The graphical weight of a given bond configuration ω is

$$W(\omega) = A^{b(\omega_g \vee \omega_o)} B^{b(\omega_g \wedge \omega_o)} 2^{c(\omega_g)} 2^{c(\omega_o)}, \tag{3.7}$$

where

$$A = e^K - 1, \quad B = \frac{e^{L+2K} - 2e^K + 1}{e^K - 1}. \tag{3.8}$$

Observe that Eq. (3.5) describes a particular case of this model provided that $A = v \equiv \lambda/(1 - \lambda)$ while $B = 0$.

The dual model is obtained by placing orange bonds between the sites of a dual lattice every time when it does not cross a green bond of the original lattice. Correspondingly, the green bonds on the dual lattice are dual to the original orange bonds. The duality relations are given by

$$A^* = 2B^{-1}, \quad B^* = 2A^{-1}, \tag{3.9}$$

And the model becomes self-dual when $AB = 2$. It is therefore suggestive that our model becomes self-dual at $\lambda = 1$ (or $v = \infty$). In order to show that it is indeed *exactly* self-dual, we shall perform the above duality transformation to the orange bonds only ($\omega_o \rightarrow \omega_o^*$), leaving the green bonds intact. This results in having green bonds on both original and the dual lattices. The green bonds on the dual lattice can be then split into those transversal to the original green bonds and those transversal to the previously vacant bonds, or symbolically $\omega_o^* = \Omega_g \vee \Omega_{\emptyset}$. Here ω_o^* is the configuration of (green) bonds *dual* to the orange bonds while Ω_g and Ω_{\emptyset} are the configurations of bonds *transversal*

to the original green and vacant bonds correspondingly. The corresponding weight is now given by

$$W(\omega) = v^{b(\omega_g)} \left(\frac{2}{v}\right)^{b(\Omega_g)+b(\Omega_\emptyset)} 2^{c(\omega_g)} 2^{c(\Omega_g \vee \Omega_\emptyset)}. \tag{3.10}$$

We now observe that $b(\Omega_g) = b(\omega_g)$, and also that $b(\Omega_\emptyset) \rightarrow 0$ as $v \rightarrow \infty$ (the original random cluster model becomes fully packed according to Eq. (3.5)). Then the weight in this limit becomes simply

$$W = 2^{b(\omega_g)} 2^{c(\omega_g)} 2^{c(\Omega_g)}. \tag{3.11}$$

The model described by such weights is manifestly self-dual since $\omega_g^* = \Omega_o \vee \Omega_\emptyset \rightarrow \Omega_o$ and $\Omega_g^* = \omega_o \vee \omega_\emptyset \rightarrow \omega_o$ as $v \rightarrow \infty$, and there is a symmetry between the green and the orange bonds.

It is tempting to speculate that a phase transition occurs exactly at the self-dual point. Although this is plausible, it is not the only possibility. In particular, there may be a phase transition at some $\lambda_1 < 1$. However, we can say the following: If at $\lambda = 1$ there is no magnetisation (i.e. percolation of green or orange bonds) then the theorem proved by two of us [24] applies; $\lambda = 1$ is a critical point in the sense of infinite correlation length and infinite susceptibility. The only other possibility is positive magnetisation at $\lambda = 1$ which implies a magnetic transition – which could be continuous or first-order – at some $\lambda_1 \leq 1$. (In particular, this is shown to happen, with a first-order transition for the large- q versions of these models [25]). Although we find these alternative scenarios unlikely, we have, in any case, established the existence of a transition in this model for some value of λ between zero and one.

3.4. Speculative remarks on relation to the critical 4-state Potts model

As mentioned above, in our opinion the most likely scenario is that a phase transition occurs precisely at $\lambda = 1$. The interesting question then is that of the universality class. Without any supporting mathematical statements, we suggest that *at* this point our model behaves similarly to the 4-state Potts ferromagnet at its critical point. In order to substantiate this claim, let us first recall the random cluster representation for a q -state Potts model:

$$Z = \sum_{\omega} K^{b(\omega)} q^{c(\omega)}. \tag{3.12}$$

For $q \leq 4$ this model is universally accepted to have a continuous transition at the self-dual point $K = \sqrt{q}$. Thus for the $q = 4$ model at the self-dual point we have

$$Z = \sum_{\omega} 2^{b(\omega)} 4^{c(\omega)}. \tag{3.13}$$

On the other hand, we can use Eq. (3.11) to rewrite the partition function of our model at $\lambda = 1$ as follows:

$$Z \propto \sum_{\omega} 2^{b(\omega)} 4^{c(\omega)} 2^{c(\Omega) - c(\omega)}. \tag{3.14}$$

The difference between the two models is in the last factor of $2^{c(\Omega) - c(\omega)}$ in Eq. (3.14). It is, however, reasonable to speculate that it can be neglected. Indeed, on average $c(\Omega) = c(\omega)$, and therefore one would expect the typical value of the difference $c(\Omega) - c(\omega)$ to be sublinear in the system size. By contrast, the individual terms $c(\Omega)$ and $c(\omega)$ indeed scale linearly with the size of the system so this correction may be “unimportant”.

This, however, does not mean that the two models approach the self-dual point in a similar fashion. In other words, we expect the exponents associated with the critical point itself (such as η and δ) of the two models to be the same, while this needs not be true for the exponents associated with the *approach* to the critical point (such as α , β and ν). In fact, the $\lambda = 1$ point of our model may well be an edge of a critical Kosterlitz–Thouless phase in which case the approach exponents would take on extreme values (zero or infinity).

4. Reflection positivity and phase transitions in the large n limit

4.1. Reflection positivity

This section concerns the reflection positivity property of the loop models defined by Eq. (1.1) which in turn permits the analysis of their large n limit. Let Λ denote a d -dimensional torus. Here, and for the remainder of this paper, it will be assumed that the linear dimensions of Λ are all the same, and are of the form $L = 2^\ell$. We denote by $N = L^d$ the number of sites in the torus. Let \mathcal{G} denote the set of all possible loop configurations on Λ . Finally let \mathcal{P} denote a hyperplane perpendicular to one of the coordinate axes which cuts through the bonds parallel to this axis dividing the torus into two equal parts. Let $\mathcal{G}_\mathcal{L}$ and $\mathcal{G}_\mathcal{R}$ be the bond configurations on the two sides of the “cut”, with the bonds intersected by \mathcal{P} belonging to *both* sets. Thus $\mathcal{G} = \mathcal{G}_\mathcal{L} \cup \mathcal{G}_\mathcal{R}$, while $\mathcal{G}_\mathcal{P} \equiv \mathcal{G}_\mathcal{L} \cap \mathcal{G}_\mathcal{R}$ contains only the intersected bonds. We now define a map $\vartheta_\mathcal{P}: \mathcal{G}_\mathcal{L} \rightarrow \mathcal{G}_\mathcal{R}$ such that it simply reflects the configuration on the “left” to that on the “right”. Let $f: \mathcal{G}_\mathcal{R} \rightarrow \mathbb{R}$ be a function that depends only on the bond configuration on the right and define $\vartheta_\mathcal{P} f(\mathfrak{g}_\mathcal{L}) \equiv f(\vartheta_\mathcal{P}(\mathfrak{g}_\mathcal{L}))$ for any $\mathfrak{g}_\mathcal{L} \in \mathcal{G}_\mathcal{L}$. A probability measure μ on the set \mathcal{G} is said to be *reflection positive* if for every such \mathcal{P} and any functions f and h as described above $\langle f \vartheta_\mathcal{P} f \rangle_\mu \geq 0$ and $\langle h \vartheta_\mathcal{P} f \rangle_\mu = \langle f \vartheta_\mathcal{P} h \rangle_\mu$.

Theorem 1. The measures μ determined by the weights in Eq. (1.1) are reflection positive on any even d -dimensional torus.

Proof. Let Λ denote one such torus and \mathcal{P} denote one of the above described planes. Let $\mathfrak{g}_\mathcal{P} \in \mathcal{G}_\mathcal{P}$ be a configuration of bonds going through this plane. Assuming that $\mu(\mathfrak{g}_\mathcal{P}) \neq 0$, let us consider the measure $\mu(\cdot | \mathfrak{g}_\mathcal{P})$. Our claim is that this splits into two measures, which we will call $\mu_\mathcal{L}(\cdot | \mathfrak{g}_\mathcal{P})$ and $\mu_\mathcal{R}(\cdot | \mathfrak{g}_\mathcal{P})$, defined on $\mathcal{G}_\mathcal{L}$ and $\mathcal{G}_\mathcal{R}$ which are independent and identical under the reflection $\vartheta_\mathcal{P}$.

Indeed, it is not hard to see that $\mu(\mathfrak{g}_\mathcal{P}) \neq 0$ if and only if $\mathfrak{g}_\mathcal{P}$ has an even number of bonds of each colour. In each half of the torus, the endpoints of these bonds serve as

“source/sinks” for bond configurations. In other words, a configuration in, say, $\mathcal{G}_{\mathcal{S}}$ must contain lines of the appropriate colour that pair up these sources. But, aside from having to satisfy these “boundary conditions”, the weights are the same as given in Eq. (1.1). These two measures defined accordingly on $\mathcal{G}_{\mathcal{S}}$ and $\mathcal{G}_{\mathcal{R}}$ are the above mentioned $\mu_{\mathcal{S}}(\cdot | \mathfrak{g}_{\mathcal{S}})$ and $\mu_{\mathcal{R}}(\cdot | \mathfrak{g}_{\mathcal{S}})$ respectively.

It is clear that if $\mathfrak{g}_{\mathcal{S}} \in \mathcal{G}_{\mathcal{S}}$ then

$$\mu_{\mathcal{S}}(\mathfrak{g}_{\mathcal{S}} | \mathfrak{g}_{\mathcal{S}}) = \mu_{\mathcal{R}}(\vartheta_{\mathcal{S}}(\mathfrak{g}_{\mathcal{S}}) | \mathfrak{g}_{\mathcal{S}}). \tag{4.1}$$

Furthermore, if $\mathfrak{g}_{\mathcal{S}}$ is a configuration and $\mathfrak{g}_{\mathcal{S}}$ is any configuration that agrees with $\mathfrak{g}_{\mathcal{S}}$ and has non-zero weight then for every $\mathfrak{g}_{\mathcal{R}} \in \mathcal{G}_{\mathcal{R}}$ we see that

$$\mu_{\mathcal{R}}(\mathfrak{g}_{\mathcal{R}} | \mathfrak{g}_{\mathcal{S}}) = \mu(\mathfrak{g}_{\mathcal{R}} | \mathfrak{g}_{\mathcal{S}}). \tag{4.2}$$

Thence, for every f that is determined on $\mathcal{G}_{\mathcal{S}}$ we have

$$\langle f \vartheta_{\mathcal{S}} f \rangle_{\mu} = \sum_{\mathfrak{g}_{\mathcal{S}}} \mu(\mathfrak{g}_{\mathcal{S}}) \langle f | \mathfrak{g}_{\mathcal{S}} \rangle_{\mu_{\mathcal{R}}} \langle \vartheta_{\mathcal{S}} f | \mathfrak{g}_{\mathcal{S}} \rangle_{\mu_{\mathcal{S}}} = \sum_{\mathfrak{g}_{\mathcal{S}}} \mu(\mathfrak{g}_{\mathcal{S}}) \langle f | \mathfrak{g}_{\mathcal{S}} \rangle_{\mu_{\mathcal{R}}}^2 \tag{4.3}$$

which cannot be negative. Similarly we get that $\langle h \vartheta_{\mathcal{S}} f \rangle_{\mu} = \langle f \vartheta_{\mathcal{S}} h \rangle_{\mu}$. □

One of the important consequences of reflection positivity is a Cauchy–Schwartz-type inequality:

$$\langle f \vartheta_{\mathcal{S}} h \rangle_{\mu} \leq \sqrt{\langle f \vartheta_{\mathcal{S}} f \rangle_{\mu} \langle h \vartheta_{\mathcal{S}} h \rangle_{\mu}}. \tag{4.4}$$

which in turn leads to the *chessboard estimates* to be described below. (The reader interested in a more detailed description of reflection positivity is referred to the review [26] and the references therein).

4.2. Uniform exponential decay for large n

In this subsection we will consider some n -colour models with $n \gg 1$ and vertex factors that are uniformly bounded above and below independently of n : $0 < c \leq \nu_1, \dots, \nu_m \leq C$. Examples include the $O(n)$ -type models and the corner-cubic models discussed in Section 2. However, the face-cubic model does not fall into this category since all the multi-coloured vertex factors vanish. It is no coincidence that we cannot treat these models since, as is obvious such models have a colour-symmetry broken phase for high enough value of bond fugacity (for brevity we omit a formal proof).

The suppression of long contours will be established by showing that long lines of any particular colour are exponentially rare in the length of the line. To prove this we will need the so called chessboard estimate which in the present context reads as follows:

Proposition 2. For $x \in \Lambda$ let $\omega_1, \dots, \omega_k$ denote indicator functions for bond events that are determined by the bonds emanating from the site x . (The ω_j need not all be distinct.) For any of these ω_j , cover the lattice with (multiple) reflections of the

corresponding event and let Z_j denote the partition function constrained so that at each site, the appropriately reflected event is satisfied. Then

$$\left\langle \prod_j \omega_j(x) \right\rangle_\mu \leq \left(\frac{Z_1}{Z} \right)^{1/N} \dots \left(\frac{Z_k}{Z} \right)^{1/N}. \tag{4.5}$$

Proof. See Subsection 2.4 of Ref. [26].

Our principal result of this subsection:

Theorem 3. Consider an n -colour loop model as described by Eq. (1.1) on the torus Λ (which is taken to be “sufficiently large”) and suppose that the vertex factors are bounded below by $c > 0$ and above by $C < \infty$ uniformly in n . For sites x and y in Λ , let $\mathcal{L}_{x,y}$ denote the probability that these sites belong to the same loop. Then, provided n is sufficiently large, there is a $\xi_n > 0$ such that for all values of R ,

$$\mathcal{L}_{x,y} \leq K e^{-|x-y|/\xi_n}, \tag{4.6}$$

where $|x - y|$ denotes the minimum length of a walk between x and y and K is a constant.

Remark. For conceptual clarity, we will start with a proof of the case $d = 2$; all of the essential ideas are contained in this case. The problems in $d > 2$ involve some minor technicalities and the general proof can be omitted on a preliminary reading.

Proof of Theorem 3 ($d = 2$). Let us focus on a particular colour – red – and show the statement is true for red loops; this only amounts to a factor of n in the prefactor. We define the “red event” as an event where at least two red bonds are connected to the site in question. It is clear that there are two main types of red events: those where the red bonds attached to a given site line up along the straight line and those where they form the right angle. These two types are shown in Figs. 4Ia and 4IIa respectively. We will denote by $\omega_I^{(\alpha)}$, $\alpha = 1, 2$ the events of the first type and $\omega_{II}^{(\beta)}$, $\beta = 1, \dots, 4$ those of

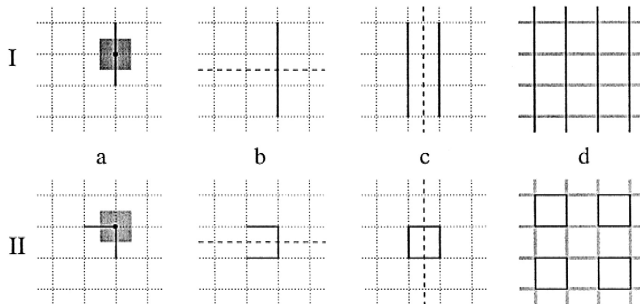


Fig. 4. Chess-board estimate on the “red” events of types I and II. The events are shaded in (a). Parts (b) and (c) represent the results of the first two reflections with respect to the dashed lines. The resulting tilings of the entire plane (torus) are shown in part (d).

the second type. Obviously there are only two distinct constrained partition functions which we respectively denote by Z_I and Z_{II} . Fig. 4 shows these single events (a), the results of the first two reflections with respect to the dashed lines (b), (c) and finally the configurations obtained by applying each reflection $\ell d = \log_2 N$ times in order to completely tile the surface of the torus. The grey lines correspond to yet unidentified bonds – these are the degrees of freedom left after the process of tiling has been completed.

Let us perform the estimate on Z_{II}/Z first. We claim that if r_{II} is any legitimate configuration that contributes to Z_{II} each of the red squares – of which there are $N/4$ – can be independently replaced by vacant bonds or a square loop of any other colour or left as red. Of course this may cost us an exchange of the “best” for the “worst” vertex factor but even so, the result is

$$\frac{Z_{II}}{Z} \leq \left(\frac{R^4}{1 + nR^4} \right)^{N/4} \left(\frac{C}{c} \right)^N \leq \left(\frac{1}{n^{1/4}} \frac{C}{c} \right)^N. \tag{4.7}$$

Let us now turn attention to the Z_I estimate. We start with a factor of R^N for the red bonds already in place – as well as another worst case scenario of C^N . As for the lines that are orthogonal to (horizontal in Fig. 4Id) once started in any colour they must continue until they wrap the torus, a total length of $L = \sqrt{N}$. There are n possible choices of colour for each line as well as the possibility of no bonds at all. Since there are a total of L lines altogether, this gives

$$Z_I \leq C^N R^N (1 + nR^L)^L. \tag{4.8}$$

To obtain our estimate on Z , we simply pick the even (or odd) sublattice of dual sites and surround each site with one of n coloured elementary loops or with a “loop” of vacant bonds. Folding in the worst case scenario for the vertex factors this gives

$$Z \geq c^N (1 + nR^4)^{N/2}. \tag{4.9}$$

The ratio may be expressed as a product of two terms namely $[R/(1 + nR^4)^{1/4}]^N$ and $(1 + nR^L)^L/(1 + nR^4)^{N/4}$ – times an additional $(C/c)^N$. Clearly the first term is bounded by $n^{-N/4}$. As for the second ratio, if $R < 1$ we may neglect nR^L for $N \gg 1$ and the ratio is bounded by one. On the other hand, if $R > 1$, we may neglect the 1 and we get, modulo a factor of n^L , another $n^{-N/4}$; we will settle for the bound of 1.

Thus we have

$$\lim_{N \rightarrow \infty} \left(\frac{Z_I}{Z} \right)^{1/N} \leq \left(\frac{1}{n^{1/4}} \frac{C}{c} \right) \tag{4.10}$$

with the same upper bound for $(Z_{II}/Z)^{1/N}$ valid for all N . We thus denote the mutual upper bound by $\epsilon_n \sim n^{-1/4}$. The desired result now follows from a standard Peierls argument: If x and y are part of the same loop, some subset of this loop must be a self-avoiding walk of length at least $2|x - y|$. We enumerate all such walks and use the chess board estimate on each particular walk. Then if $\epsilon_n \lambda_2 < 1$ where λ_2 is the two-dimensional connectivity constant we write

$$\epsilon_n \lambda_2 = \exp\{-1/2 \xi_n\} \tag{4.11}$$

(the factor of 2 because we must go there and back) and the stated result follows. \square

Proof of Theorem 3 ($d > 2$). The preliminary steps are the same as the two-dimensional case: There are again only two types of ω 's (but with more indices) and two constrained partition functions which we again denote by Z_I and Z_{II} . The pattern for Z_{II} is the two-dimensional pattern in Fig. 4IIId reflected in all directions orthogonal to the plane visualised. Noting that each reflection doubles the number of red squares, we see that in the Z_{II} patterns there are a total of $(2^d)^{d-2} \times \frac{1}{4}L^2 = \frac{1}{4}N$ squares altogether. Repeating the argument leading to Eq. (4.7) we end up with exactly the same bound. What is a little harder is the estimates on Z_I and the partition function itself.

We first claim that Z can be estimated by

$$Z \geq c^N (1 + nR^4)^{\frac{d}{4}N}. \tag{4.12}$$

To achieve this, we assert that the following holds: There is a set of plaquettes of the lattice with the property that each bond of the lattice belongs to exactly one plaquette. Once this claim is established it is clear that Eq. (4.12) holds; indeed there are just $[1/4] \times Nd$ plaquettes in question, we consider those configurations in which each of them is independently left vacant or traversed with an elementary loop in one of the n possible colours. Let us turn to a proof of the above assertion.

Let $\hat{e}_1, \dots, \hat{e}_d$ denote the elementary unit vectors. We adapt the following notation for plaquettes: If, starting at $\mathbf{x} \in \mathbb{Z}^d$ we first move in the \hat{e}_j direction then in the \hat{e}_k direction and then complete the circuit we will denote this plaquette by $[\hat{e}_j \diamond \hat{e}_k]_{\mathbf{x}}$. In general we can have $[\pm \hat{e}_j \diamond \pm \hat{e}_k]_{\mathbf{x}}$ and it is noted that $[\hat{e}_j \diamond \hat{e}_k]_{\mathbf{x}} = [\hat{e}_k \diamond \hat{e}_j]_{\mathbf{x}}$. Starting with the origin, consider the following list of instructions for plaquettes:

$$[\hat{e}_1 \diamond - \hat{e}_2]_0, [\hat{e}_2 \diamond - \hat{e}_3]_0, \dots, [\hat{e}_d \diamond - \hat{e}_1]_0. \tag{4.13}$$

So far so good – each bond emanating from the origin belongs to exactly one plaquette. If $\mathbf{x} = (x_1, \dots, x_d)$ let us define $\pi_j(\mathbf{x}) = (-1)^{x_j}$ to be the parity of the j th coordinate. Then at the site \mathbf{x} , select the following: $[\pi_1(\mathbf{x})\hat{e}_1 \diamond - \pi_2(\mathbf{x})\hat{e}_2]_{\mathbf{x}}, \dots, [\pi_d(\mathbf{x})\hat{e}_d \diamond - \pi_1(\mathbf{x})\hat{e}_1]_{\mathbf{x}}$. Taking the union of all these lists, it is clear that indeed each bond belongs to *at least* one plaquette, the only question is whether there has been an over-counting.

To settle this issue we establish the following assertion: Let \mathbf{x} denote a site. Then a plaquette is specified by the instructions at \mathbf{x} if and only if the same plaquette is specified by the neighbours of \mathbf{x} on that plaquette.

Indeed, consider one such plaquette namely $[\pi_j(\mathbf{x})\hat{e}_j \diamond - \pi_{j+1}(\mathbf{x})\hat{e}_{j+1}]_{\mathbf{x}}$. (If necessary we use the convention $\hat{e}_{d+1} = \hat{e}_1$.) One of the relevant neighbours is $\mathbf{x}' = \mathbf{x} + \pi_j(\mathbf{x})\hat{e}_j$. But at \mathbf{x}' , we have instructions for the plaquette $[\pi_j(\mathbf{x}')\hat{e}_j \diamond - \pi_{j+1}(\mathbf{x}')\hat{e}_{j+1}]_{\mathbf{x}'}$. Since $\pi_j(\mathbf{x}') = -\pi_j(\mathbf{x})$ and $\pi_{j+1}(\mathbf{x}') = \pi_{j+1}(\mathbf{x})$ this is seen to be the same plaquette. The other neighbour, at $\tilde{\mathbf{x}} = \mathbf{x} - \pi_{j+1}(\mathbf{x})\hat{e}_{j+1}$, follows from the same argument. Reversing the rôles of \mathbf{x} and \mathbf{x}' (as well as \mathbf{x} and $\tilde{\mathbf{x}}$) establishes the ‘only if’ part of the assertion.

It is thus evident that no bond can belong to two plaquettes: If \mathbf{x} and \mathbf{x}' are neighbours and there are instructions coming from \mathbf{x} to make some particular plaquette that includes the bond $\langle \mathbf{x}, \mathbf{x}' \rangle$ then:

- (a) These are the only instructions coming from \mathbf{x} pertaining to this bond.
- (b) The list at \mathbf{x}' (and the other corners of this plaquette) also include this plaquette.
- (c) If \mathbf{y} is another neighbour of \mathbf{x} which is *not* in this plaquette there cannot be instructions at \mathbf{y} to include the plaquette containing \mathbf{x} , \mathbf{x}' and \mathbf{y} .

Items (a) and (b) are immediate. To see item (c), note that by the assertion, such instructions would also have to appear on the list at \mathbf{x} . But, by item (a), they do not. Eq. (4.12) is now established.

Finally, let us estimate Z_I from above. First it is noted that the Z_I patterns consist of parallel lines (pointing in the direction of the original red bonds) that pass through every site. Hence the only loops we can draw are confined to the various $(d - 1)$ -dimensional hyperplanes orthogonal to these lines. Modulo vertex factors – which we estimate by C^N – each hyperplane yields the $(d - 1)$ -dimensional partition function. We thus have

$$Z_{II} \leq R^N C^N (\Xi)^L, \tag{4.14}$$

where Ξ is the $(d - 1)$ -dimensional partition function with all vertex factors equal to unity. (This observation is not of crucial importance – the key ingredient is the number of available bonds – but it serves to compartmentalise the argument.) Let us tend to the estimate of Ξ .

We denote by Ξ_M the contribution to Ξ that comes about when there are exactly M bonds in the configuration. Given the placement of the bonds, let us count (estimate) the number of ways that they can be organised into self-returning walks and then independently colour each walk. Throwing in a combinatoric factor for the placement of the M bonds and the fact that there cannot conceivably be more than $M/4$ loops to colour we arrive at

$$\Xi_M \leq \binom{(d - 1)L^{d-1}}{M} R^M n^{M/4} \times Y(M), \tag{4.15}$$

where $Y(M)$ is defined as follows: For M bonds, consider their placement on the lattice chosen so as to maximise the number of ways in which the resulting (colourless) graph can be decomposed into distinct loops. (Such decompositions are done by “resolving” all intersecting vertices into connected pairs of bonds, similar to Fig. 1.) Then the quantity $Y(M)$ denotes this maximum possible number of decompositions.

Let β (which equals $2(d - 1) - 1 = 2d - 3$) denote the greatest possible number of local options for an on-going self-avoiding walk. We claim that

$$Y(M) \leq (2\beta)^M. \tag{4.16}$$

Indeed, take this optimal placement of M bonds and, starting at some predetermined bond (and moving in some predetermined direction) draw a loop of length P . There are no more than β^P possibilities for this loop. Then what is left over cannot yield a total better than $Y(M - P)$. This must be done for every possible value of P and summed. Defining $Y(0) = 1$, we arrive at the recursive inequality

$$Y(M) \leq \sum_{0 < P \leq M} \beta^P Y(M - P) \tag{4.17}$$

and the bound in Eq. (4.16) can be established inductively. Putting together Eqs. (4.15) and (4.16) we find

$$\Xi \leq \sum_M \binom{(d - 1)L^{d-1}}{M} R^M n^{M/4} (2\beta)^M = (1 + 2\beta R n^{1/4})^{(d-1)L^{d-1}}. \tag{4.18}$$

And thus

$$Z_l \leq C^N R^N \left[(1 + 2\delta R n^{1/4})^{(d-1)L^{d-1}} \right]^L = \left[CR(1 + 2\delta R n^{1/4})^{(d-1)} \right]^N. \quad (4.19)$$

An estimate for the straight segments is nearly complete. We write

$$\left[\frac{Z_l}{Z} \right]^{1/N} \leq \frac{c}{C} \frac{R(1 + 2\delta R n^{1/4})^{(d-1)}}{(1 + nR^4)^{d/4}}. \quad (4.20)$$

Let us split the power of the denominator: $d/4 = 1/4 + (d-1)/4$; the first term will handle the R in the numerator and this ratio is less than $n^{-1/4}$. As for what is left over, it is easy to see that

$$\frac{1 + 2\delta R n^{1/4}}{(1 + nR^4)^{1/4}} = \frac{1}{(1 + nR^4)^{1/4}} + 2\delta \frac{R n^{1/4}}{(1 + nR^4)^{1/4}} \leq 1 + 2\delta. \quad (4.21)$$

Thus we have

$$\left[\frac{Z_l}{Z} \right]^{1/N} \leq \frac{C}{c} (1 + 2\delta)^{d-1} \frac{1}{n^{1/4}}, \quad (4.22)$$

so again all of the ω 's have estimates with the scaling of $n^{-1/4}$. The remainder of the argument follows the same course as the two-dimensional argument with modifications where appropriate. \square

4.3. Translational symmetry breaking and related phase transitions

So far we have shown that for a sufficiently large n , there are no phase transitions associated with the divergent loop correlation length. However, this does not rule out the possibility of phase transitions of a totally different nature – in fact we shall see that such transitions do indeed take place. The type of these transitions happens to be sensitive to the lattice structure and the vertex factors that correspond to loop intersections. This leads to a variety of critical phenomena which has not been observed in the case of a honeycomb lattice [7].

We shall consider the case of all vertex factors being positive and uniformly bounded. Let the number of colours n be very large and consider the high fugacity limit of $R \rightarrow \infty$. The model becomes fully packed in this limit, i.e. every bond is occupied. But since n is large, according to the estimates of the previous subsection the partition function is dominated by configurations in which there are as many loops as possible. In two dimensions, it is clear that these “maximum entropy states” break the translational symmetry: the loops of different colours may run around either odd or even plaquettes of the lattice. (We remark that in $d > 2$, the situation is considerably more complicated and is currently under study.)

On the other hand, if the bond fugacity is low, the system must be translationally invariant. Therefore one would anticipate that at some large but finite value of R a translational symmetry-breaking transition takes place. Due to a doubly positionally degenerate nature of the resulting high-density state, we find it natural to expect that

such transition is of the Ising type. Indeed, it appears to be very similar to the transition in the Ising antiferromagnet.

Let us now turn to the rigorous arguments supporting this qualitative picture.

Theorem 4. Consider an n -colour loop model on \mathbb{Z}^2 with vertex factors bounded below by $c > 0$ and above by $C < \infty$ uniformly in n . Then, if n is sufficiently large there is a phase transition characterised by the breaking of translational symmetry. In particular, there exist ϵ (sufficiently small) and Δ (sufficiently large) such that for $R^4 n < \epsilon$ all states that emerge as limits of torus states are translation invariant. On the other hand, if $R^4 n > \Delta$, there are (at least) two states with broken translational symmetry.

Proof. Let us start with the formal proof of translation-invariance at low fugacity. Any given site belongs to zero, two or four bonds. Let \odot , \ominus and \oplus denote the corresponding events and Z_\odot , Z_\ominus and Z_\oplus the constrained partition functions in which every site is of the stated type. We will estimate $\text{Prob}(\ominus) \leq (Z_\ominus/Z_\odot)^{1/n}$ and $\text{Prob}(\oplus) \leq (Z_\oplus/Z_\odot)^{1/n}$; obviously $Z_\odot = 1$. In calculating Z_\oplus we notice that in all configurations, each bond must be coloured and there are at most $N/2$ separate loops. This gives

$$Z_\oplus \leq [R^2 n^{1/2}]^N Y(2N), \tag{4.23}$$

where $Y(M)$ is the same quantity that appears in Eq. (4.15). Similarly, in Z_\ominus half the bonds are used (resulting in a factor of R^N) and there are no more than $N/4$ loops. We get

$$Z_\ominus \leq [Rn^{1/4}]^N Y(N) \tag{4.24}$$

– essentially the square root of the above.

In order to show that there is translation invariance it is sufficient to establish the following: Let A and B denote local events, let T_x be the translation operator to the point $x \in \Lambda$ and let \hat{e} denote any unit vector. Then we must show that for all x with $|x|$ large the probabilities of $\mu_\Lambda(A \cap T_x B)$ and $\mu_\Lambda(A \cap T_{x+\hat{e}} B)$ are essentially the same. We note that either the supports of A and $T_x B$ are attached by a $*$ -connected path of sites of type Z_\ominus and Z_\oplus , or they are separated by a connected circuit of sites of type Z_\odot . (Two sites are considered $*$ -connected if they share the same plaquette.) However, in case the support of A is surrounded by a connected circuit of empty sites, it is clear that for any \hat{e} , A and $T_{\hat{e}} A$ have the same probability. Thus we arrive at

$$\left| \mu_\Lambda(A \cap T_x B) - \mu_\Lambda(A \cap T_{x+\hat{e}} B) \right| \leq \mu_\Lambda(C_{A,T_x B}), \tag{4.25}$$

where $C_{A,T_x B}$ is the event that there is a $*$ -connected path of non- \odot sites between the support of A and that of $T_x B$. Now, provided $Rn^{1/4}$ is small, $\mu_\Lambda(C_{A,T_x B}) \leq e^{-\kappa|x|}$ with $e^{-\kappa} \sim \alpha Rn^{1/4}$ for some constant α . Thence translation invariance (among the states that emerge from the torus) is established.

Let us now turn attention to the opposite limit, namely $Rn^{1/4} \gg 1$. We claim that with high probability the plaquettes on the lattice are of one of the two types: they are either surrounded by four bonds of the same colour, or by bonds of four different colours.

We will again define site events representative of the two purported ground states. If a site is on the even sublattice, we say it is of the A-type if it has four bonds of two different colours, one colour occupying the positive \hat{e}_1 and \hat{e}_2 directions and the other, the negative \hat{e}_1 and \hat{e}_2 directions. The odd site of the A-type is defined as the mirror reflection of an even A-type site. A site is said to be of B-type if the rôles of the even and the odd sublattices are reversed – see Fig. 5. Any site which is not of one of these two types will be called “bad”. There are three choices for such bad sites, namely housing no bonds, only two bonds and a “+” configuration in which the \hat{e}_1 and $-\hat{e}_1$ directions have the same colour and similarly for the $\pm\hat{e}_2$ directions. The first two are the old \odot and \ominus from the preceding portion of this proof, while the constrained partition function for the latter event, denoted (for the lack of better notations) by $Z_{\#}$ can be bounded by

$$Z_{\#} \leq R^{2N} n^{2\sqrt{N}}. \tag{4.26}$$

On the other hand, we can always estimate Z from below by $[Rn^{1/4}]^2 N$ as discussed previously. Thus we have (as $|\Lambda| \rightarrow \infty$)

$$\mu_{\Lambda}(\odot) \leq R^{-2} n^{-1/2}, \tag{4.27a}$$

$$\mu_{\Lambda}(\ominus) \leq n^{-1/2}, \tag{4.27b}$$

$$\mu_{\Lambda}(\#) \leq n^{-1/4}. \tag{4.27c}$$

Thus almost all sites are either A-type or B-type. However, there is no constraint that A and B sites cannot be neighbours. But, as we show below, this possibility is also suppressed for $n \gg 1$. Firstly, let us notice that the reflections through bonds map the even into the odd sites (and vice versa). By definition, the A-even and A-odd events are mirror images, and similarly for B. Thus, under multiple reflections, A-sites map to A-sites and B’s to B’s. Hence, if (A-B) is the event that the origin is of the A-type and its right nearest neighbour is of the B-type, we have

$$\mu_{\Lambda}(A-B) \leq \left(\frac{Z_{A-B}}{Z} \right)^{2/N}, \tag{4.28}$$

where Z_{A-B} is the partition function constrained according to the pattern shown in Fig. 6. Now, the interior of the A or B columns still consists of the four-bond loops – just as a pure A or pure B-type tiling would. However, as is easily seen in Fig. 6, a boundary

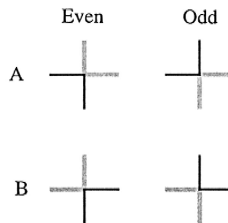


Fig. 5. Even and odd sites of A- and B-types.

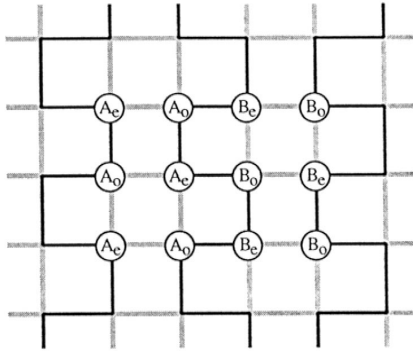


Fig. 6. Tiling of the torus obtained by multiple reflections of the (A-B) event. The sites marked as A_e are the even sites of the A-type, etc. The zigzag lines that form loops wrapping the torus and separating the columns of A- and B-sites are shown black.

between the columns is formed by a zigzag line of a single colour running around the torus. Thus we arrive at

$$Z_{A-B} = R^{2N} n^{\frac{1}{4}N} n^{\frac{1}{2}\sqrt{N}} \tag{4.29}$$

and hence, in the large N limit, $\mu_A(A-B) \leq n^{-1/2}$.

The rest of the proof follows easily: A connected cluster of A-type sites has, on its boundary a “bad” site or an A-B pair; similarly for the clusters of B-sites. Thus, if $Rn^{1/4}$ and n are large enough, the possibility that any given site is isolated from the origin is small. Thence, given that the origin is of the A-type – which has probability very near one half – the population of B-type sites is small and vice versa. Evidently, there are two states, one with an abundance of A-sites and the other with an abundance of B’s. It is not hard to see that in these two states, translation symmetry is broken – both are dominated by the appropriate staggered patterns. □

5. Final remarks and conclusions

So far we have succeeded in proving the following general statements:

- A generic multi-coloured loop model with *all* vertex factors ν_1, \dots, ν_m uniformly bounded from above and below does not have a phase transition corresponding to the divergence of the loop size in *any* dimension *provided* that the number of colours n is sufficiently large.
- In two dimensions these models undergo a different phase transition, presumably of the Ising type, that is associated with breaking the *translational* symmetry⁴. While the examples of the models that have been considered in this paper all

⁴ At present it is not clear whether this statement holds in $d > 2$. While the “ground states” (the states at $R = \infty$) of the system are *not* translationally invariant, their degeneracy appears to grow very fast with d , and it is not obvious that the resulting “entropy” will not destroy the transition at any finite value of R .

have their vertex factors independent of the particular arrangements of different colours entering the vertex, all these results remain valid even if this is not the case as long as all different vertex factors are still reflection-symmetric and bounded from above and below by some positive numbers.

A clear example of a model that does not follow this rule is the loop model derived from the face-cubic spin model in Section 2.3. Here the bonds of different colours are simply not allowed to share a vertex, making the corresponding vertex factor vanish. It is clear that as the bond fugacity R increases, the system will have a phase transition (possibly of the first-order) associated with breaking the *colour* symmetry. Indeed, the only way to “pack” more bonds into the system is to force all of them to be of the same colour. This transition is very similar to the Widom–Rowlinson transition.

Another interesting observation can be made about the two-dimensional loop model with all two- and four-leg vertices having the same weight (this is the model originating from the corner-cubic spin model – see Section 2.2). Comparing the states of the system at $R = \infty$ when all bonds are occupied (the fully packed limit), and at $R = 1$ when no additional weight is associated with placing extra loops, we conclude that the $(n + 1)$ -coloured model at $R = \infty$ is identical to the n -coloured model at $R = 1$. Indeed, one should simply consider the vacant bonds at $R = 1$ as being coloured grey to turn the n -coloured loop system into a fully packed $(n + 1)$ -coloured system.

Now start with the large enough n at $R = \infty$. The result of Section 4.3 guarantees the existence of a broken translational symmetry in this case. By the argument presented here this means the existence of a broken symmetry in $(n - 1)$ -coloured, $R = 1$ case. Now let us continuously increase the value of the bond fugacity in this $(n - 1)$ -coloured model. As R grows, there are only two possible scenarios: the broken translational symmetry is either lost via a phase transition or is retained all the way to $R = \infty$. The former case corresponds to the *intermediate* symmetry-broken phase surrounded by the phase transitions at $R \leq 1$ and $R \geq 1$. If the latter scenario is realised, then go to $R = \infty$ and repeat the process of mapping it onto a lower n model. Notice that this latter scenario can not continue all the way to $n = 1$ because the $n = 1$, $R = 1$ case is nothing but a loop representation for the Ising magnet at $T = 0$ which does not have broken translational symmetry.⁵ Therefore one *must* find an intermediate phase at some not very large value of n (although we cannot rule out a possibility of it being just the point $R = 1$).

From the above we also learn that the $n = 2$ (Ashkin–Teller-like) model is *not* critical in its fully packed limit (since it is just the $T = 0$ Ising model). On the other hand, the $O(2)$ loop model discussed in Section 2.1 *is* critical in this limit (it maps onto the square ice model). But the only difference between the two is the factor of 3 in the same-colour four-leg vertex factor.

On the basis of this observation (along with the first two remarks of this section) we reiterate that the vertex factors and lattice details are often important for determining the phase diagram of a particular model.

⁵ Indeed, these loops appearing in the $T = 0$ “high temperature” expansion are also the domain walls of the Ising model on the dual lattice at $T \rightarrow \infty$. These dual spins are assigned their values independently with probability $1/2$, and the contours separating regions of opposite type are manifestly translation invariant.

Acknowledgements

The authors would like to thank J. Kondev, R. Kotecky and T. Spencer for the interesting discussions. L.C. and K.S. were supported in part by NSA Grant No. MDA904-98-1-0518 and NFS Grant No. 99-71016; L.P.P. was supported in part by DOE Grant No. DE-FG02-90ER40542.

Appendix A. Vertex factor for the $O(n)$ vertices and loop decomposition

In this appendix we present the calculation of the generic vertex factor appearing in the loop expansion of the $O(n)$ -type model featured in Subsection 2.1. As it turns out, this calculation is independent of the lattice details (such as the dimensionality or coordination number) which is reflected in our notation. Thus let Ω_n denote the n -dimensional solid angle and let $\mathbf{S} = (S_1, \dots, S_n)$ denote the components of a n -dimensional unit vector. Since we only consider the spin at one particular site, the actual site index is conveniently dropped for now. The vertex factor we need to calculate is

$$\nu_{m_1, \dots, m_n} = \frac{1}{|\Omega_n|} \int_{\Omega_n} S_1^{2m_1} \dots S_n^{2m_n} d\Omega_n, \quad (\text{A.1})$$

where $|\Omega_n| = 2\pi^{n/2} [\Gamma(n/2)]^{-1}$ is the total solid angle and $d\Omega_n/|\Omega_n|$ is the Haar measure. The object ν_{m_1, \dots, m_n} is, of course, the ‘‘vertex factor’’ originating from $2m_1$ terms from the first component, \dots $2m_n$ terms from the n th component associated with these numbers of these types of bonds entering into a vertex.

Proposition 5. Let $\phi(2m)$ – which is equal to $(2m - 1)!!$ – denote the number of different ways of dividing $2m$ objects into m distinct pairs. Then

$$\nu_{m_1, \dots, m_n} = D_n(M) \prod_{i: m_i \neq 0} \phi(2m_i), \quad (\text{A.2})$$

where $M = \sum_{i=1}^n m_i$ and

$$D_n(M) = 2^{-M} \Gamma(n/2) / \Gamma(M + n/2) = (n - 2)!! / (n + 2M - 2)!! . \quad (\text{A.3})$$

Remark. The calculation of the value for ν_{m_1, \dots, m_n} has appeared before, cf. Ref. [27] and references therein. However, the interpretation that emerges from this formula has, to our knowledge, not been discussed previously (cf. the remark following the proof).

Proof. The proof boils down to the calculation of the above integral. We write

$$d\Omega_n = \sin^{n-2} \theta_n d\theta_n d\Omega_{n-1}, \quad (\text{A.4})$$

where $\theta_n = \mathbf{S} \cdot \hat{\mathbf{e}}_n$. Also,

$$\mathbf{S} = (\mathbf{T} \sin \theta_n, \cos \theta_n), \quad (\text{A.5})$$

where \mathbf{T} is an $(n - 1)$ -dimensional unit vector. Finally, we note the fact that

$$|\Omega_n| = |\Omega_{n-1}| \frac{\sqrt{\pi} \Gamma\left(\frac{n}{2} - 1\right)}{\Gamma(n/2)}. \tag{A.6}$$

Starting with the definition (A.1) with the help of Eqs. (A.4)–(A.6), the following recursion relation is obtained:

$$\begin{aligned} \nu_{m_1, \dots, m_n} &= \frac{1}{|\Omega_n|} \int_{\Omega_n} S_1^{2m_1} \dots S_n^{2m_n} d\Omega_n \\ &= \frac{1}{|\Omega_{n-1}|} \int_{\Omega_{n-1}} T_1^{2m_1} \dots T_{n-1}^{2m_{n-1}} d\Omega_{n-1} \frac{\Gamma(n/2)}{\sqrt{\pi} \Gamma\left(\frac{n}{2} - 1\right)} \\ &\quad \times \int (\sin \theta_n)^2 \sum_{i=1}^{n-1} m_i + n - 2 (\cos \theta_n)^{2m_n} d\theta_n \\ &= \nu_{m_1, \dots, m_{n-1}} \frac{\Gamma(m_n + \frac{1}{2}) G_{n-1}}{\sqrt{\pi} G_n}, \end{aligned} \tag{A.7}$$

where

$$G_k = \frac{\Gamma\left(2 \sum_{i=1}^k m_i + \frac{k}{2}\right)}{\Gamma(k/2)}. \tag{A.8}$$

Using the recursion relation (A.7) to reduce the dimensionality of the spin space to 1 (and the fact that $\nu_{m_1} = 1$), we arrive at the desired expression:

$$\begin{aligned} \nu_{m_1, \dots, m_n} &= \frac{\Gamma(n/2)}{\pi^{n/2} \Gamma(M + n/2)} \prod_{i=1}^n \Gamma(m_i + \frac{1}{2}) \\ &= \frac{\Gamma(n/2)}{2^M \Gamma(M + n/2)} \prod_{i: m_i \neq 0} (2m_i - 1)!!. \end{aligned} \tag{A.9}$$

□

Remark. As a consequence of the expression (A.2) we may decompose the coloured bond configurations to obtain a colourless loop model (which generalises the result of Eq. (2.3) obtained for the square lattice). In particular, at each vertex we have a factor which depends only on the total number of bonds entering the vertex (and the dimensionality of the spin n) times the number of ways that the bonds of the various colours can be paired. By choosing a particular pairing scheme at each vertex, the bond configuration breaks, unambiguously, into a collection of self-returning walks (loops). Configurations with the same set of loops – that is to say the same set of bonds and the same pairing schemes (walking instructions) at each vertex – but which differ in the colours of the loops are seen to have identical weights. Thus we consider configurations

\mathcal{Z} of loops (bonds + walking instructions) and the partition function becomes the analog of the expression (2.3):

$$Z = \sum_{\mathcal{Z}} \left(\frac{\lambda}{n} \right)^{b(\mathcal{Z})} v_1^{m_1} \dots v_k^{m_k} n^{\ell(\mathcal{Z})} \quad (\text{A.10})$$

with v_p being the factor for a vertex with $2p$ bonds given by

$$v_p = v_p(n) = \frac{n^p \Gamma(n/2)}{2^p \Gamma(p + n/2)} = \frac{n^p}{n(n+2) \dots (n+2p-2)}. \quad (\text{A.11})$$

We observe that the right-hand side of Eq. (A.10) provides a well-defined model for all n and can be continued – essentially with no ambiguity – to non-integer values. As usual, we may *define* correlation functions via loop probabilities or be content with the numerators containing strings divided by denominators without.

In the previous discussions of these issues (see Ref. [27] and references therein) the goal has been to derive the self-avoiding walk as the $n \rightarrow 0$ limit of the $O(n)$ models. Thus the vertex weights themselves get continued without reference to intermediate models. (And indeed, the self-avoiding walk does emerge as $n \rightarrow 0$.) Although we have only treated the case where the original spin models were defined by $Z = \text{Tr} \prod_{\langle i,j \rangle} (1 + \lambda \mathbf{S}_i \cdot \mathbf{S}_j)$ as opposed to $Z = \text{Tr} \prod_{\langle i,j \rangle} \exp(\lambda \mathbf{S}_i \cdot \mathbf{S}_j)$ – we claim that the latter case also leads to graphical models that are well defined for non-integer n . (The derivation is somewhat intricate and will eventually appear in a future publication.) However, we remark that in either case, the $n \rightarrow 0$ limit is, in this context, quite simple to understand.

References

- [1] J. Kondev, Int. J. Mod. Phys. B 11 (1997) 153, cond-mat/9607181.
- [2] S.O. Warnaar, B. Nienhuis, J. Phys. A: Math. Gen. 26 (1993) 2301.
- [3] H.A. Kramers, G.H. Wannier, Phys. Rev. 60 (1941) 252.
- [4] M. Kac, J.C. Ward, Phys. Rev. 88 (1952) 1332.
- [5] N.V. Vdovichenko, Sov. Phys. JETP 20 (1965) 477.
- [6] R.J. Baxter, Exactly Solved Models in Statistical Mechanics (Academic Press, New York, 1982).
- [7] E. Domany et al., Nucl. Phys. 190 (1981) 279.
- [8] B. Nienhuis, Phys. Rev. Lett. 49 (1982) 1062.
- [9] B. Nienhuis, Int. J. Mod. Phys. B 4 (1990) 929.
- [10] H.W.J. Blöte, B. Nienhuis, J. Phys. A: Math. Gen. 22 (1989) 1415.
- [11] M.T. Batchelor, B. Nienhuis, S.O. Warnaar, Phys. Rev. Lett. 62 (1989) 2425.
- [12] M. Aizenman, B. Nachtergaele, Commun. Math. Phys. 164 (1994) 17.
- [13] M.J. Martins, B. Nienhuis, R. Rietman, Phys. Rev. Lett. 81 (1997) 504, cond-mat/9709051.
- [14] R.J. Baxter, J. Phys. A: Math. Gen. 19 (1986) 2821.
- [15] J. Kondev, C.L. Henley, Phys. Rev. B 52 (1995) 6628.
- [16] J.L. Jacobsen, J. Kondev, Nuclear Physics B 532 (1998) 635, cond-mat/9804048.
- [17] J.L. Jacobsen, J. Kondev, J. Stat. Phys. 96 (1999) 21, cond-mat/9811085.
- [18] J.V. Jose et al., Phys. Rev. B 16 (1977) 1217.
- [19] H.J.F. Knops, Phys. Rev. Lett. 39 (1977) 766.
- [20] J. Ashkin, E. Teller, Phys. Rev. 64 (1943) 178.
- [21] D.B. Abraham, C.M. Newman, Phys. Rev. Lett. 61 (1988) 1969.

- [22] L. Chayes, J. Machta, *Physica A* 239 (1997) 542.
- [23] C.E. Pfister, Y. Velenik, *J. Stat. Phys.* 88 (1997) 1295, cond-mat/9704017.
- [24] L. Chayes, K. Shtengel, *Commun. Math. Phys.* 204 (1999) 353, cond-mat/9811203.
- [25] T. Baker, L. Chayes, *J. Stat. Phys.* 93 (1998) 1.
- [26] S.B. Shlosman, *Russ. Math. Surv.* 41:3 (1986) 83.
- [27] N. Madras, G. Slade, *The Self-Avoiding Walk* (Birkhäuser, Boston, 1993).