Journal of Statistical Physics, Vol. 106, Nos. 3/4, February 2002 (© 2002)

Surface Tension and the Ornstein–Zernike Behaviour for the 2D Blume–Capel Model

Ostap Hryniv[1,2] **and Roman Kotecký**[3]

Received September 21, 2001

We prove existence of the surface tension in the low temperature 2D Blume–Capel model and verify the Ornstein–Zernike asymptotics of the corresponding finite-volume interface partition function.

KEY WORDS: Lattice systems; interface; Blume–Capel model.

1. SETTING AND RESULTS

Study of Gibbs states describing an interface between coexisting phases of lattice models goes back to the seminal Dobrushin's paper,(1) where he proved existence of translation noninvariant Gibbs state for the 3-dimensional Ising model. His idea of describing the interface in terms of weakly interacting excitations—walls—separated by "flat regions" with minimal energy cost—ceilings—was subsequently extended to a more general class of models(2) whose pure phases are described by the Pirogov–Sinai theory.(3) Even more general situation arises whenever there is a competition between several different types of ceilings.(4) For these models, the existence of an interface Gibbs state, characterized by overwhelming prevalence of a given type of ceiling, depends on the values of parameters like the temperature, external fields, etc. For particular values of these parameters we have coexistence: Gibbs states characterized by different types of ceilings can be constructed as the thermodynamic limit of finite volume states with an

[1] DMA-EPFL, CH-1015 Lausanne, Switzerland.
[2] On leave from: Institute for Applied Problems of Mechanics and Mathematics, 3-b Naukova str., 79601 Lviv, Ukraine.
[3] CTS, Charles University, Jilská 1, Praha 1, Czech Republic; e-mail: kotecky@cucc.ruk.cuni.cz

0022-4715/02/0200-0431/0

appropriately chosen boundary condition. The corresponding phase boundary can be studied using once more the Pirogov–Sinai theory, this time to describe the "gas of walls" representing the interface.

Our aim here is to discuss a similar situation in two-dimensional models. Even though there are no translation noninvariant Gibbs states describing interfaces for a two-dimensional system, as shown first by Gallavotti for the Ising model,[5] the surface tension is still nonvanishing and it is an interesting problem to prove its existence and to show that the asymtotics of the corresponding finite volume interface partition function satisfies what we call, in an analogy with the behaviour of correlations, the Ornstein–Zernike asymptotics. Observe, however, that the study of the surface tension in a 2D system at low temperature can be related to the study of correspondingly inclined interfaces decomposed into regular and irregular "jumps." The former present typical pieces of inclined low-temperature interfaces in two dimensions and play the role of true ceilings in the Dobrushin's picture. The latter are perturbations appearing at positive temperatures and are analogous to the Dobrushin's walls. Roughly in this setting, the Ornstein–Zernike asymptotics has been studied for several models: percolation,[6,7] self-avoiding walks,[8,9] phase boundary of the two-dimensional Ising model.[10]

Occurence of competing ceilings brings a new factor to the problem. Even though we expect that (at least) some of the results could be extended to a rather general situation, we will not aim at full generality and will restrict the formulation of our results to a simple case where, however, competing ceilings play an important role. Namely, we will discuss the direction dependent surface tension and the corresponding asymptotics for the two-dimensional Blume–Capel model. Only some of the abstract principles presented in the next section will be formulated in a general manner.

Considering spin configurations $\sigma \in X \equiv \{-1, 0, 1\}^{\mathbb{Z}^2}$, the Hamiltonian of the Blume–Capel model, in a finite volume $\Lambda \subset \mathbb{Z}^2$ and under fixed boundary conditions $\bar{\sigma} \in \{-1, 0, 1\}^{\mathbb{Z}^2}$, is

$$H_\Lambda(\sigma \mid \bar{\sigma}) = J \sum_{\substack{\langle x, y \rangle \\ x, y \in \Lambda}} (\sigma(x) - \sigma(y))^2 + J \sum_{\substack{\langle x, y \rangle \\ x \in \Lambda,\, y \in \partial\Lambda}} (\sigma(x) - \bar{\sigma}(y))^2 - \lambda \sum_{x \in \Lambda} \sigma(x)^2 - h \sum_{x \in \Lambda} \sigma(x). \tag{1.1}$$

Here, the first two sums are over pairs of nearest neighbour sites, $\partial\Lambda$ denotes the outer boundary of the set Λ, $\partial\Lambda = \{x \in \mathbb{Z}^2 \setminus \Lambda : \exists y \in \Lambda, |y - x| = 1\}$; the real parameters λ and h are called external fields, and $J > 0$ is the coupling constant.

The phase diagram,[11] for fixed temperature $T = 1/\beta$ and $J > 0$, features the triple point $(\lambda, h) = (\lambda_0(T, J), 0)$, $\lambda_0(T, J) \downarrow 0$ as $T \downarrow 0$, at which all three phases (predominantly plus, zero, or minus) coexist, and from which three lines of coexistence of two phases emanate. In the following, we will consider

$$h = 0 \qquad \text{and} \qquad \lambda > \lambda_0(T, J), \tag{1.2}$$

i.e., the case of coexistence of the plus and minus phases.

Our aim is to study the asymptotic behaviour of the partition function describing the interface between plus and minus phases inclined by an angle θ. Namely, let σ^a, $a = 1, 2$, denote the basic column configurations:

$$\sigma^a \equiv (\sigma^a_t)_{t=-\infty}^{+\infty}, \tag{1.3}$$

where

$$\sigma^1_t \equiv \begin{cases} +1, & t > 0, \\ -1, & t \leqslant 0, \end{cases} \qquad \sigma^2_t \equiv \begin{cases} +1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0, \end{cases} \tag{1.4}$$

and ω_k be the (vertical) shift, $(\omega_k(\sigma))_t = \sigma_{t-k}$. Now, for any θ, $|\theta| < \frac{\pi}{2}$, we define the θ-inclined basic configurations of the type a, $a = 1, 2$, according to the following formulas:

$$\sigma^{\theta, a} = (\sigma^{\theta, a}(x))_{x \in \mathbb{Z}^2}, \qquad \text{where} \quad (\sigma^{\theta, a}(k, l))_{l=-\infty}^{+\infty} \equiv \omega_{[k \tan \theta]}(\sigma^a). \tag{1.5}$$

Finally, we introduce the mixed boundary conditions $\sigma^{\theta, a, b} = \sigma^{\theta, a, b}(x)$, $x \in \mathbb{Z}^2$, via

$$\sigma^{\theta, a, b}(x) \equiv \begin{cases} \sigma^{\theta, a}(x), & x_1 < 0, \\ \sigma^{\theta, b}(x), & x_1 \geqslant 0. \end{cases} \tag{1.6}$$

For a box $\Lambda \equiv \{x \in \mathbb{Z}^2;\, |x_1| \leqslant L, |x_2| \leqslant M\}$, we consider the partition functions

$$Z^{\theta, a, b}_{L, M} = \sum_{\sigma_\Lambda} \exp\{-\beta H_{\Lambda_L}(\sigma \mid \sigma^{\theta, a, b})\}, \qquad a, b = 1, 2, \tag{1.7}$$

as well as the partition function $Z^+_{L, M}$ with homogeneous boundary condition $\sigma^+(x) = 1$, $x \in \mathbb{Z}^2$. All these partition functions depend on λ and J (recall that $h = 0$).

Theorem 1. Let $J>0$, $h=0$, and $\varepsilon, \delta>0$. There exists an inverse temperature $\beta_0(\varepsilon, \delta)$ such that for every $\beta>\beta_0$, $\lambda \geqslant \varepsilon$, and any angle θ, $|\theta|<\frac{\pi}{2}-\delta$, the limit

$$\tau(\theta)=-\frac{1}{\beta} \lim_{L \to \infty} \lim_{M \to \infty} \frac{\cos \theta}{L} \log \frac{Z_{L, M}^{\theta, a, b}}{Z_{L, M}^{+}} \tag{1.8}$$

exists and does not depend on $a, b \in \{1,2\}$. Moreover, for any $a, b \in \{1,2\}$, there exists a constant $C_{a, b}=C_{a, b}(\theta)$ such that

$$\frac{Z_{L, M}^{\theta, a, b}}{Z_{L, M}^{+}}=\frac{C_{a, b}}{\sqrt{L}} \exp \left\{-\frac{\beta L \tau(\theta)}{\cos \theta}\right\}(1+o(1)) \tag{1.9}$$

as $L \to \infty$ provided only that $M \geqslant L \tan \frac{\pi-\delta}{2}$.

Remark 1.1. All coefficients $C_{a, b}(\theta)$ and the surface tension $\tau(\theta)$ above depend on β, λ, and J. For an explicit expression of $C_{a, b}(\theta)$ and $\tau(\theta)$ in terms of a cluster expansion see (2.20)–(2.22).

A remarkable feature of the interfaces in the Blume–Capel model is the phenomenon of "prewetting" of the microscopic $\pm$ interface by zero spins. It turns out that while for $\lambda>2J$ the main contribution to the partition function $Z_{L, M}^{\theta, a, b}$ comes from terms with direct contact between $+$ and $-$ spins over the interface, for $\lambda<2J$ a layer of zero spins is included between them. For λ very close to λ_0 ($\lambda-\lambda_0 \sim O(e^{-\beta})$ as $\beta \to \infty$) the layer of zeros actually spreads over several lattice sites, with its thickness growing due to "entropic repulsion" as $\lambda \searrow \lambda_0$ (actually, in ref. 12 this type of wetting was discussed only for $\lambda=\lambda_0$). However, away from the triple point λ_0 (i.e., for $\lambda>\lambda_0+\varepsilon$ with $\varepsilon>0$) and temperatures low enough, the leading contribution features a thin, one lattice site, layer of zero spins for all values of λ up to $\lambda=2J$.

Yet another characteristic, with threshold at $\lambda \approx 4J$, occurs for the angles $\theta \neq 0$. Namely, at zero temperature and $\lambda \in(2J, 4J)$, the leading contributions contain zeros in the corners at which a horizontal piece of interface meets a vertical jump; on the other hand, these "corner zeroes" disappear for $\lambda>4J$ (see Fig. 3).

Notice that for $d=3$ these features mean an existence of a particular "surface phase transition" (for example, there is a line $\lambda_1(\beta)$, $\lim_{\beta \to \infty} \lambda_1(\beta)=2J$, such that for $\lambda=\lambda_1(\beta)$ there exist two distinct Gibbs states with $\pm$ interface, one with the layer of zeros and one without it ref. 4). Even though in the case $d=2$ discussed here, the resulting surface tension smoothly interpolates between different types of behaviour, an additional technical subtleties occur since one is forced to deal with different leading behaviours simultaneously. This is a typical situation appearing in Pirogov–Sinai theory and this was the tool used for a study of horizontal interfaces in ref. 4. A new technique developed in the present paper is introduced having in mind that we are treating here the case $d = 2$.

The main result—Theorem 1 provides a sharp large deviation asymptotics for the height of the interface. When combined with the ideas developed in refs. 10, 13–15) our approach can be used to obtain the Brownian bridge approximation to the distribution of shapes of appropriately rescaled interfaces. The corresponding results will be discussed elsewhere.

The rest of the paper is organized as follows. Section 2 contains main definitions and presents the key ideas of the proof. The description of the zero-temperature behaviour of the interface model at hand is given Section 3.1. The proof of Theorem 1 is split into two parts—for almost horizontal directions (Section 3.2) and almost diagonal ones (Section 3.3). Finally, Appendices A and B collect some analytical facts from the theory of linear polymer models used in the main text.

2. CONTOURS AND INTERFACES

We introduce contours in a standard way, counting the sites occupied by zero spins as a part of contour. Namely, given a configuration $\sigma \in X$, its *contour* is any finite connected component of the *boundary* $B(\sigma)$ defined as the union of all unit bonds of dual lattice separating the sites with unequal values of σ and all plaquettes (closed unit squares) with spin zero at its center. A contour is thus a finite collection of bonds and plaquettes; however, it is often useful to think about it in terms of the corresponding closed connected subset of $\mathbb{R}^2$.

The boundary $B(\sigma)$ of any configuration σ from the set $X_{L,M}^{\theta,a,b} \subset X$ of all configurations coinciding, outside of $\Lambda_{L,M}$, with $\sigma^{\theta,a,b}$ introduced above, has one infinite component. We call it *an interface* and use $I(\sigma)$ to denote it. Whenever I is an interface and $\Lambda_{L,M}$ is fixed, we introduce the set $\Delta(I)$ of all sites attached to I, $\Delta(I) = \{x \in \mathbb{Z}^2;\ \mathsf{dist}(x, I) \leqslant \frac{1}{2}\}$, distance taken in max norm, and use $\Delta_{L,M}(I)$ for its intersection with $\Lambda_{L,M}$, $\Delta_{L,M}(I) = \Delta(I) \cap \Lambda_{L,M}$. Further, we use $\Lambda_{L,M}^{+}(I)$ (resp. $\Lambda_{L,M}^{-}(I)$) to denote the intersection of the infinite component of $\mathbb{Z}^2 \setminus \Delta(I)$, lying above (resp. below) I, with $\Lambda_{L,M}$, and $\mathsf{Int}\, I$ to denote the "interior" of the interface I, $\mathsf{Int}\, I = \Lambda_{L,M} \setminus (\Delta(I) \cup \Lambda_{L,M}^{+}(I) \cup \Lambda_{L,M}^{-}(I))$. An example of I is pictured on Fig. 1 below. Notice that all spins on $\partial\Lambda_{L,M}^{+}(I)$ ($\partial\Lambda_{L,M}^{-}(I)$) are necessarily fixed to $+1$ (-1). Also, every component of $\mathsf{Int}\, I$ has a fixed spin, either $+1$ or -1, on its boundary.

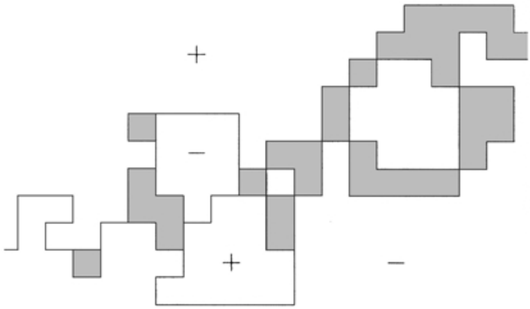

Fig. 1. A phase boundary between + and − phases with two neutral interior components; the regions occupied by zeroes are shadowed.

Finally, we use $E_{L,M}(I)$ to denote the energetical cost of the phase boundary—the coupling J multiplied by the number of bonds of I, in $\Lambda_{L,M}$, separating spins $\pm$ from spin 0 plus $4J$ times the number of bonds between + and − spins plus λ times the number of zero spin plaquettes in I.

Given an interface I, there is a class of configurations σ that have I for its interface, $I(\sigma)=I$. Summing over over these configurations, we get

$$Z^{\theta,a,b}_{L,M} = \sum_{I\in\mathcal{I}^{\theta,a,b}_{L,M}} \exp\{-\beta E_{L,M}(I)\}\, Z_{\mathrm{Int}\, I} Z^{+}_{\Lambda^{+}_{L,M}(I)} Z^{-}_{\Lambda^{-}_{L,M}(I)}. \tag{2.1}$$

Here, the sum is over the set $\mathcal{I}^{\theta,a,b}_{L,M}$ of all interfaces consistent with the boundary condition $\sigma^{\theta,a,b}$ and $Z_{\mathrm{Int}\, I}$ is the product of the partition functions $Z^{\sigma_i}_{V_i}$ over all components of the interior of I, $V_i\subset \mathrm{Int}\, I$. Usually the b.c. σ_i for a component V_i are predetermined by the contour I itself; however, if a component V_i is surrounded by zeroes, then we call it "neutral" and take $Z_{V_i}=Z^{+}_{V_i}+Z^{-}_{V_i}$. Now, we will use the fact that, assuming $\lambda\geqslant\lambda_0+\varepsilon$, both phases + and − are stable (cf. ref. 11). We can thus employ cluster expansion and, moreover, referring to the spin flip symmetry, infer that it yields an identical result for plus and minus phase,

$$\log Z^{+}_{V} = \log Z^{-}_{V} = \sum_{C\subset V} \Phi^{T}(C). \tag{2.2}$$

Here the terms $\Phi^{T}(C)$ correspond to connected subsets C of $\mathbb{Z}^2$—clusters, or rather sums over all clusters with the same support—and are quickly decaying with the size of C. Namely, using $|C|$ to denote the number of sites in C and applying standard cluster expansion estimates,[10, 16] one can show that there exists $\kappa_0(J, \lambda)$ such that

$$\sum_{C \ni 0} |\Phi^T(C)|\, e^{\kappa_0 \beta |C|} \leqslant 1$$

for sufficiently large β; in particular,

$$|\Phi^T(C)| \leqslant e^{-\kappa_0 \beta |C|}. \tag{2.3}$$

As a result, in the same way as in ref. 10, we get

$$\frac{Z^{\theta, a, b}_{L, M}}{Z^{+}_{L, M}} = \sum_{I \in \mathscr{I}^{\theta, a, b}_{L, M}} \exp\{-\beta E_{L, M}(I)\} \exp\left\{-\sum_{C \cap \Delta_{L, M}(I) \neq \varnothing} \Phi^T(C)\right\}. \tag{2.4}$$

Mayer expanding the second term, we get

$$\frac{Z^{\theta, a, b}_{L, M}}{Z^{+}_{L, M}} = \sum_{I \in \mathscr{I}^{\theta, a, b}_{L, M}} \exp\{-\beta E_{L, M}(I)\} \sum_{\substack{\mathscr{C} = \{C\} \\ C \cap \Delta_{L, M}(I) \neq \varnothing}} \prod_{C \in \mathscr{C}} (e^{-\Phi^T(C)} - 1).$$

One can now control the limit $\lim_{M \to \infty}$ in a standard way and get, with the obvious notation,

$$\mathscr{Z}^{\theta, a, b}_{L} = \lim_{M \to \infty} \frac{Z^{\theta, a, b}_{L, M}}{Z^{+}_{L, M}} = \sum_{I \in \mathscr{I}^{\theta, a, b}_{L}} \exp\{-\beta E_{L}(I)\} \sum_{\substack{\mathscr{C} = \{C\} \\ C \cap \Delta_{L}(I) \neq \varnothing}} \prod_{C \in \mathscr{C}} (e^{-\Phi^T(C)} - 1).$$

To develop the polymer representation of the partition function $\mathscr{Z}^{\theta, a, b}_{L}$ we use the natural decomposition into irreducible pieces (cf. refs. 1, 5, 17–19, and ref. 10, Section 4.4). Namely, let us use π to denote the projection $\pi: \mathbb{R}^2 \to \mathbb{R}$, $\pi(x_1, x_2) = x_1$, and consider any pair $(I, \mathscr{C})$, where $I \in \mathscr{I}^{\theta, a, b}_{L}$ is an interface and $\mathscr{C}$ is a collection of clusters connected to it. We say that $(I, \mathscr{C})$ is *regular* (*of type a*) *in the column* x_1 if the following two conditions are met:

(1) $\pi^{-1}(x_1) \cap \mathscr{C} = \varnothing$,

(2) $\pi^{-1}(x_1) \cap I = (\pi^{-1}(x_1) \cap I(\omega_k(\sigma^a)))$ for some integer k,

i.e., the vertical line $\pi^{-1}(x_1)$ does not intersect any cluster $C \in \mathscr{C}$ and the intersection of $\pi^{-1}(x_1)$ and the contour I coincides (up to vertical translation) with the intersection of $\pi^{-1}(x_1)$ and $I(\sigma^a)$, the contour of the basic column configuration σ^a. Given any pair $(I(\sigma), \mathscr{C})$, we cut it with any vertical line $\pi^{-1}(x_1)$, where $(I(\sigma), \mathscr{C})$ is regular (of any type) and obtain a

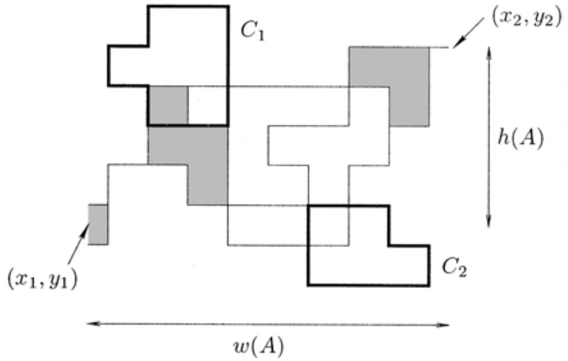

Fig. 2. An aggregate decorated with two clusters C_1 and C_2.

collection $\mathscr{A}(\sigma, \mathscr{C})$ of irreducible components (*polymers, aggregates*) A labeled according to the type of regularity at the cutting lines. Note that this cutting procedure creates a one-to-one correspondence between pairs $(I, \mathscr{C})$ and sequences of labeled aggregates $(I_A, \mathscr{C}_A)$ such that each aggregate has two labels (on the left and on the right) and the corresponding labels of touching aggregates match (like in dominoes game).

For an aggregate $A = (I_A, \mathscr{C}_A)$ we introduce, in a natural way, the *initial* and the *ending* points (the points (x_1, y_1) and (x_2, y_2) on Fig. 2) as well as the following characteristics:[4] the *width* $w(A) = x_2 - x_1$, the *height* $h(A) = y_2 - y_1$, and the *activity*

$$\mathfrak{z}(A) = \exp\{-\beta E(I_A)\} \prod_{C \in \mathscr{C}_A} (e^{-\Phi^T(C)} - 1) \tag{2.5}$$

As a result, we get

$$\mathscr{Z}_L^{\theta, a, b} = \sum_{\mathscr{A}_L^{a, b}(k_\theta)} \prod_{A \in \mathscr{A}_L^{a, b}(k_\theta)} \mathfrak{z}(A) \tag{2.6}$$

where the sum runs over all collections of aggregates corresponding to interfaces from $\mathscr{I}_{L, M}^{\theta, a, b}$. Having rewritten the interface partition function in terms of a gas of aggregates, we can employ the methods of cluster expansions. The only obstacle, however, lies in the fact that aggregates have two different types of endpoints, $a = 1, 2$, and these have to match for the neighbouring aggregates. An analog of this problem for a two-dimensional

[4] Which are shift-invariant that allows us not to distinguish between different aggregates that are shift-congruent.

interface in the case of a three-dimensional model is solved with the help of Pirogov–Sinai theory applied to the gas of aggregates.[4]

Here, the gas of aggregates is one-dimensional (by the projection onto the horizontal axis). This fact suggests that one should study them as a linear polymer system employing the renewal approach similar to the one used in ref. 9. However, the endpoints of our polymers are labeled, the latter being subject to the matching rule (like in the dominoes game). This naturally generates matrix formalism above the classical renewal theory.

On the way to such reformulation, we first introduce the "grand-canonical partition sum"

$$\mathscr{Z}_L^{a,b}(u) = \sum_k e^{k\beta u} Z_L^{\theta_k, a, b}, \qquad \mathscr{Z}_0^{a,b}(u) = \delta_{a,b} \tag{2.7}$$

that can be represented in the so-called polymer form (though with labeled polymers, see Appendix B),

$$\mathscr{Z}_L^{a,b}(u) = \sum_{j=1}^{L} \sum_{\substack{m_i,\, i=1,\dots,j:\\ m_i \geqslant 1,\, \sum m_i = L}} \sum_{(\sigma_0, \sigma_1,\dots,\sigma_j) \in \{a\} \times \{1,2\}^{j-1} \times \{b\}} \prod_{i=1}^{j} \mathfrak{z}_{m_i}^{\sigma_{i-1},\sigma_i}(u)$$

with the weights of irreducible components given by

$$\mathfrak{z}_\ell^{a,b}(u) = \sum_{A \in \mathbf{A}_\ell^{a,b}} e^{h(A)\,\beta u} \mathfrak{z}(A). \tag{2.8}$$

Here $\theta_k \in (-\pi/2, \pi/2)$ is defined through $\tan\theta_k = k/L$ and $\mathbf{A}_\ell^{a,b}$ is the set of all aggregates compatible with the b.c. (a, b) that have a fixed initial point and the width ℓ. Then, for any pair (a, b) of labels we introduce the power series

$$\mathscr{F}_{a,b}(u, w) = \sum_{\ell=1}^{\infty} \mathfrak{z}_\ell^{a,b}(u)\, w^\ell \quad \text{and} \quad \mathscr{Z}_{a,b}(u, w) = \sum_{L=0}^{\infty} \mathscr{Z}_L^{a,b}(u)\, w^L, \tag{2.9}$$

and consider the matrices

$$\mathscr{F} = (\mathscr{F}_{a,b}(u, w))_{a,b \in \{1,2\}} \quad \text{and} \quad \mathscr{Z} = (\mathscr{Z}_{a,b}(u, w))_{a,b \in \{1,2\}}. \tag{2.10}$$

Clearly, the matrix equation

$$\mathscr{Z} = \sum_{m=0}^{\infty} \mathscr{F}^m = [\mathbb{1} - \mathscr{F}]^{-1} \tag{2.11}$$

is valid whenever the corresponding series is absolutely convergent. In fact we will show below that for any u,

$$u \in \mathcal{O}_\delta \equiv \left\{ z \in \mathbb{C}^1 : |\Re z| < \mu - \frac{\delta}{\beta} \right\}, \qquad \delta > 0, \tag{2.12}$$

with

$$\mu = \mu(\lambda, J) \equiv 2J + \min(\lambda, 2J), \tag{2.13}$$

the series in (2.11) is absolutely convergent in w, $|w| < r(u)$. Moreover, the radius $r(u)$ of convergence is uniformly positive for u belonging to any compact subset of the region $\mathcal{O}_\delta$. Then the asymptotics of the coefficients $\mathscr{Z}_L^{a,b}(u)$ of the function $\mathscr{Z}$ follows in a standard way (see, e.g., ref. 20, App. A; ref. 21, p. 330; and refs. 22 and 23).

Let us begin by considering the limiting case $\lambda \to \infty$. Then the spin 0 is entirely suppressed and the model degenerates into the Ising model discussed in detail, in the same context as here, in Chapter 4 of ref. 10. In particular, only one type of ceiling occurs, all polymers (aggregates) are of the same type, and thus matrices $\mathscr{Z}$ and $\mathscr{F}$ are 1×1, i.e., we are in the framework of the classical renewal theory. We will recover the results from ref. 10 in Appendix A. For the reader familiar with the Ising case, it provides a simple example of the use of the renewal approach for proving the asymptotic behaviour (1.9).

Coming back to the case of finite $\lambda > \lambda_0 + \varepsilon$, our approach to proving Theorem 1 can be described as follows. In view of the "theory" of linear polymer models with labeled polymers developed in Appendix B, the free energy $f_\beta(u)$ corresponding to the sequence of matrix-valued partition functions $\mathscr{Z}_L(u)$ (see (2.7)) equals (for real u)

$$f_\beta(u) \equiv -\log w_0(u),$$

where $w_0(u) \equiv w_0(u, \beta)$ is the smallest positive solution to the characteristic equation

$$\det\left[\mathbb{1} - \mathscr{F}(u, w)\right]\Big|_{w = w_0(u)} = 0 \tag{2.14}$$

with $\mathbb{1}$ denoting the 2×2 identity matrix and $\mathscr{F}(u, w)$ being the generating function of the irreducible components (see (2.10)):

$$\mathscr{F}(u, w) = \sum_{\ell \geqslant 1} \mathscr{F}^{(\ell)}(u)\, w^\ell. \tag{2.15}$$

A straigthforward application of the analytic implicit function theorem together with the smooothenes properties of the LHS in (2.14) established below allow then to continue the free energy $f_\beta(u)$ in some complex neighbourhood of the real line.

The rigorous study of the generating function $\mathscr{F}(u, w)$ entering (2.14) is a non-trivial task. Fortunately, the low-temperature considerations suggest a relatively simple approximating model of SOS-type. For our analysis it is convenient to consider the cases of "almost-horizontal" directions (i.e., corresponding to $|\theta| < \pi/4 - \varepsilon$) and of "almost-diagonal" directions (with $\varepsilon < \theta < \pi/2 - \varepsilon$) separately.

In the region of "almost-horizontal" angles, $|\theta| < \pi/4 - \varepsilon$, a good approximative candidate is defined through the linear generating function

$$\widehat{\mathscr{F}}_1(u, w) = \mathscr{F}^{(1)}(u)\, w$$

with the condition on θ being transformed into the condition that $u \in \mathcal{O}_{\delta K}$, see the exact definition in (3.7)–(3.8) below. Indeed, as we shall see, the minimal positive solution $\widehat{w}_1(u)$ to the equation

$$\det\,[\mathbb{1} - \widehat{\mathscr{F}}_1(u, w)] \equiv 1 - w \cdot \mathrm{trace}\,\mathscr{F}^{(1)}(u) + w^2 \cdot \det \mathscr{F}^{(1)}(u) = 0 \tag{2.16}$$

yields a good approximation to the quantity of interest, $w_0(u)$.

Defining

$$\widetilde{\mathscr{F}}_1(u, v) \equiv \widehat{\mathscr{F}}_1(u, v/\bar{w}_1), \qquad \bar{w}_1 = \bar{w}_1(u, \beta) \equiv \mathrm{trace}\,\mathscr{F}^{(1)}(u), \tag{2.17}$$

we reduce our study of (2.16) to the investigation of the quadratic equation

$$1 - v + \alpha v^2 = 0 \tag{2.18}$$

through the identity $\widehat{w}_1(u) \equiv v_1(u)/\bar{w}_1(u)$, $v_1(u)$ being the principal positive solution of (2.18). The study of this quadratic equation is performed in Section 3.2.2 and is based on the fact that in the region of parameters under consideration we have $0 < \alpha(u) < 1/4$, see (3.11).

Finally, in Section 3.2.3 we establish the inequality

$$\left| \frac{\mathfrak{Z}_\ell^{a,b}(u)}{(\bar{w}_1(u))^\ell} \right| \leqslant \exp\{-\kappa(\beta - \beta_0)(\ell - 1)\} \tag{2.19}$$

with some finite positive constants κ and β_0. This estimate, being uniform in $\ell \geqslant 2$ (except the case $\ell = 2$ and $a = b = 2$), provides us with a good control of the difference

$$\mathscr{F}(u, v/\bar{w}_1) - \widetilde{\mathscr{F}}_1(u, v)$$

and makes possible to work with each entry of the matrix $\mathscr{Z}$ separately.

Indeed, denoting by $\delta_{a,b}$ the Kronecker delta function, we use the standard matrix inversion formula to get

$$\begin{aligned}\mathfrak{g}_{a,b}(u,w) &:= \mathscr{Z}_{a,b}(u,w)\cdot \det\,[\mathbb{1}-\mathscr{F}(u,w)]\\ &= \delta_{a,b}(1-\mathscr{F}_{3-a,3-b}(u,w))+(1-\delta_{a,b})\,\mathscr{F}_{a,b}(u,w).\end{aligned}$$

Next, since for β large enough the principal solution $w_0(u)$ to the characteristic equation (2.14) is simple and (thanks to the massgap condition) the function $\mathfrak{g}_{a,b}(u,w)$ is uniformly finite in some neighbourhood of $(u, w_0(u))$, $u \in \mathcal{O}'_\delta \subset \mathcal{O}_\delta$, we obtain

$$\mathscr{Z}_L^{a,b}(u) = \left.\frac{\mathfrak{g}_{a,b}(u,w)\, w^{-L-1}}{-\frac{d}{dw}\det\,[\mathbb{1}-\mathscr{F}(u,w)]}\right|_{w=w_0(u)}.$$

Finally, another application of the Laplace method gives us the asymptotics (1.9) with the constant $C_{a,b}$ defined as

$$\begin{aligned}C_{a,b} &\equiv \left.\frac{\mathfrak{g}_{a,b}(u,w)}{w(-\frac{d}{dw}\det\,[\mathbb{1}-\mathscr{F}(u,w)])\sqrt{2\pi(\log w)''_{uu}}}\right|_{\substack{w=w_0\\ u=u_\theta}}(1+o(1))\\ &= \left.\frac{\mathfrak{g}_{a,b}(u,w)}{w(-\frac{d}{dw}\det\,[\mathbb{1}-\mathscr{F}(u,w)])\sqrt{2\pi(\log w)''_{uu}}}\right|_{\substack{w=\bar{w}_1\\ u=u_\theta}}(1+o(1)) \qquad (2.20)\end{aligned}$$

for β large enough; here u_θ is the value of the external field conjugate to the direction of interest θ:

$$\beta\tan\theta \equiv \frac{d}{du} f_\beta(u) = -\frac{d}{du}\log w_0(u). \qquad (2.21)$$

Of course, the surface tension $\tau(\theta)$ defined in (1.8) and the free energy $f_\beta(u)$ are related via the Legendre transformation:

$$\tau(\theta) = \frac{1}{\beta} f^*(\tan\theta)\cos\theta, \qquad f^*(x) \equiv \sup_u\,(ux - f(u)). \qquad (2.22)$$

Since all these details are standard and well known[5] (see, e.g., ref. 20, App. A; ref. 21, p. 330 for the renewal part, ref. 22, Chap. 4; ref. 23, Section IV.5; and ref. 24, Section 2 for the Laplace method in the local limit theorem), we restrict our attention below to the study of the characteristic equation (2.14) and to the proof of the mass-gap condition (2.19).

[5] A recent exposition can be also found in ref. 9.

The technique described above works well everywhere outside any neigbourhood of the region of parameters where the main contribution comes from thick diagonal interfaces (i.e., outside $0<\lambda\leqslant 2J$ and the inclination angle $\theta=\pi/4$). A new feature occurring in the latter region is that for the corresponding values of u, all quantities $(\mathfrak{z}_\ell^{a,b}(u))^{1/\ell}$ are of the same order even in the limit $\beta\to\infty$, rendering thus impossible existence of a reference scale $\bar{w}_1(u)$ with which the exponential bounds (2.19) hold for all but a finite number of ℓ's. As a result, the LHS of the characteristic equation (2.14) is no longer a polynomial of finite degree and its study requires a certain generalization of the above method (say, along the lines described in Proposition A.6 under Assumption A.8 in Appendix A; see also the discussion in Section 3.3.1).

Fortunately, for the directions close to the diagonal one our model has another approximation of the SOS type. The latter shares all the mentioned properties of the horizontal SOS model and thus its analysis requires only to check the analogue of the massgap property (2.19). Another interesting feature of our diagonal SOS model is that the principal solution to the corresponding characteristic equation is known *exactly*, since the latter is of the second degree.

We refer the reader to Section 3.3 for further details.

3. PROOF

This section is devoted to the proof of our main result—Theorem 1. Since we expect that the polymer system under consideration at small but positive temperatures behaves like a small perturbation of the system at zero-temperature, we start by studying the asymptotic properties of the underlying (zero-temperature) Solid-On-Solid model.

Our approach is based on certain generalization of the classical renewal theory. Since we failed to find a suitable formulation in the literature, we had to collect necessary (known) facts from the classical renewal theory as well as their appropriate generalization in Appendices A and B.

3.1. Zero-Temperature Picture

Fix any θ, $|\theta|<\frac{\pi}{2}$, and a pair $(a,b)\in\{1,2\}^2$. At zero temperature, the Gibbs measure in $\Lambda_{L,M}$ with the b.c. $\sigma^{\theta,a,b}$ becomes the uniform distribution in the set of all configurations from Ω_L the boundary of which consists of a single contour I from $\mathscr{I}_{L,M}^{\theta,a,b}$ with minimal energy, i.e., I belongs to the set

$$\widehat{\mathscr{I}}_{L,M}^{\theta,a,b}=\{I\in\mathscr{I}_{L,M}^{\theta,a,b}:E_{L,M}(I)=\min_{I\in\mathscr{I}_{L,M}^{\theta,a,b}}E_{L,M}(I)\}.$$

A simple computation shows us that the set $\hat{\mathscr{I}}_{L,M}^{\theta,a,b} \subset \mathscr{I}_{L,M}^{\theta,a,b}$ is a piecewise constant function of the ratio λ/J. In fact, there are three different realizations of $\hat{\mathscr{I}}_{L,M}^{\theta,a,b}$ denoted by $\mathscr{I}_0$, $\mathscr{I}_1$, and $\mathscr{I}_2$ below that correspond to the ratio λ/J in the interval Δ_i, respectively:

$$\Delta_0 = (\varepsilon, 2), \qquad \Delta_1 = (2, 4), \qquad \Delta_2 = (4, +\infty).$$

Up to the obvious modification near the boundary $\partial\Lambda_{L,M}$ (depending on the b.c. $\sigma^{\theta,a,b}$) the sets $\mathscr{I}_i$ of contours can be characterized as follows (see Fig. 3):

$\mathscr{I}_0$: shortest $+0-$ contours;

$\mathscr{I}_1$: shortest $\pm$ contours with a spin 0 put inside every corner;

$\mathscr{I}_2$: shortest (monotone, staircase-like) contours between $+$ and $-$.

Note that $\mathscr{I}_2$ contains only the contours appearing in the 2D Ising model at zero temperature (see ref. 10). On the other hand, due to the minimal energy condition and the attractive character of the interaction between 0-spins, the contours from $\mathscr{I}_0$ and $\mathscr{I}_1$ are subject to additional constraints; as a result, they are horizontally-diagonal (for $|\theta| \leqslant \pi/4$; see Fig. 3a,b) or vertically-diagonal (for $\pi/4 \leqslant |\theta| < \pi/2$).

In intervals Δ_0 and Δ_2 our system contains essentially one type of contours and thus can be treated, at suficiently small temperatures, using a standard perturbation technique. However, in the intermediate region $\lambda/J \in \Delta_1$ the system smoothly iterpolates between two types of behaviour and a necessity of a new approach becomes evident already at zero temperature for macroscopically inclined interfaces (i.e., with inclination angles $|\theta| \in (0, \pi/4)$, see Fig. 3b). At the border points $\lambda = 2J$ and $\lambda = 4J$ the geometry of the zero-temperature phase boundaries is even reacher due to the additional combinatorial complexity (=entropy) of the interpolating interfaces in the whole region $|\theta| < \pi/4$, see Fig. 4.

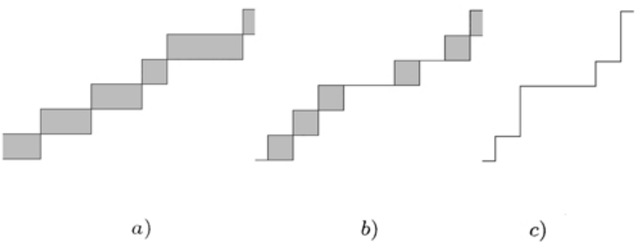

Fig. 3. Typical $\pm$ interfaces at zero temperature: (a) a horizontally-diagonal "thick" interface for $\lambda/J \in \Delta_0$, (b) a horizontally-diagonal "thin" interface for $\lambda/J \in \Delta_1$, (c) a "true" Ising-like interface for $\lambda/J \in \Delta_2$; the regions occupied by zeroes are shadowed.

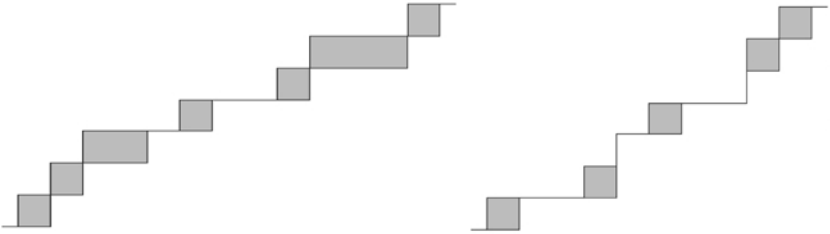

Fig. 4. Possible zero-temperature $\pm$ interfaces at the "transition" points $\lambda = 2J$ and $\lambda = 4J$.

To treat such situations, we develop in Appendix B a short "theory" of linear polymer models with labeled polymers that perfectly fits our needs. An additional advantage of this "theory" is that it makes possible to study our SOS-model in three regions Δ_i simultaneously (i.e., for any $\lambda > \lambda_0 + \varepsilon$) avoiding explicit combinatorial considerations.

In view of the lattice symmetries, it is enough to study the directions $\theta \in [0, \pi/2]$. In what follows, we will treat the cases of "almost horizontal" directions $|\theta| < \pi/4 - \eta$ and "almost diagonal" directions $\theta \in (\eta, \pi/2 - \eta)$ separately; here η stands for a fixed arbitrarily small positive constant independent of the inverse temperature β.

3.2. "Almost Horizontal" Directions $|\theta| < \frac{\pi}{4} - \eta$

3.2.1. Solid-On-Solid Approximation

Let first $\theta > 0$. A careful analysis of pictures in Figs. 3 and 4 shows that a natural candidate for describing the low-temperature behaviour of the $\pm$ interface under consideration is the SOS model constructed from the "irreducible" elements of horizontal projection one—1-*jumps*—that are shown in Fig. 5.

Being based on the "increasing" elements only, such a model is well adapted to describe the uniformly positive directions $\theta \in [\eta, \pi/4 - \eta]$; to

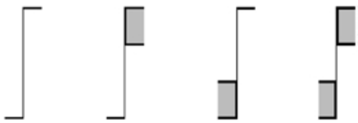

Fig. 5. Increasing irreducible 1-jumps; the thin vertical pieces can have zero length.

cover the whole region $|\theta| < \pi/4 - \eta$, an adequate low-temperature SOS-approximation should include also the corresponding "decreasing" elements (i.e., the symmetric counterparts of those depicted[6] in Fig. 5. Thus, a good choice of a low-temperature SOS-model is the one corresponding to the (matrix-valued) generating function

$$\hat{\mathscr{F}}_1(u, w) = \mathscr{F}^{(1)}(u)\, w = \begin{pmatrix} \mathfrak{z}_1^{1,1}(u) & \mathfrak{z}_1^{1,2}(u) \\ \mathfrak{z}_1^{2,1}(u) & \mathfrak{z}_1^{2,2}(u) \end{pmatrix} w \tag{3.1}$$

whose entries are computed in the following statement.

Lemma 3.1. For any $u \in \mathscr{O}_0$, we have

$$\mathfrak{z}_1^{1,1} = \frac{\exp\{-4\beta J\} \sinh(4\beta J)}{\cosh(4\beta J) - \cosh(\beta u)}, \tag{3.2}$$

$$\mathfrak{z}_1^{1,2} = \frac{\exp\{-\beta\lambda/2\}(1 - \exp\{-4\beta J\}) \cosh(\beta u/2)}{\cosh(4\beta J) - \cosh(\beta u)}, \tag{3.3}$$

$$\mathfrak{z}_1^{2,2} = \exp\{-\beta\lambda\} \left(\frac{\cosh(\beta u) - \exp\{-4\beta J\}}{\cosh(4\beta J) - \cosh(\beta u)} + e^{-2\beta J} \right). \tag{3.4}$$

Proof. Expression in the RHS of (3.2) coincides with Eq. (4.5.8) from ref. 10. Other formulas can be obtained by direct computation. ▮

Following the general scheme outlined in Section 2, the main reference scale is of order $\bar{w}_1 = \text{trace}\ \mathscr{F}^{(1)}$. We prefer however to use its simplified analogue

$$w_1 = \frac{1 + e^{-\beta\lambda}(e^{2\beta J} + 2\cosh \beta u)}{2(\cosh(4\beta J) - \cosh(\beta u))}, \tag{3.5}$$

which is equivalent to $\bar{w}_1$ in the limit $\beta \to \infty$. For future reference we note that

$$\frac{d}{\beta\, du} \log w_1 = \frac{2 \sinh(\beta u)}{e^{\beta\lambda} + e^{2\beta J} + 2\cosh(\beta u)} + \frac{\sinh(\beta u)}{\cosh(4\beta J) - \cosh(\beta u)}. \tag{3.6}$$

Remark 3.2. The second terms in Eqs. (3.6) correspond to the Ising-like interfaces (that are typical for $\lambda > 4J$) whereas the first ones reflect

[6] Observe however that in this system the neighbouring polymers obey the obvious matching rule (like in the dominoes game); such Markovian character of interaction produces a matrix-valued analogue of the renewal theory.

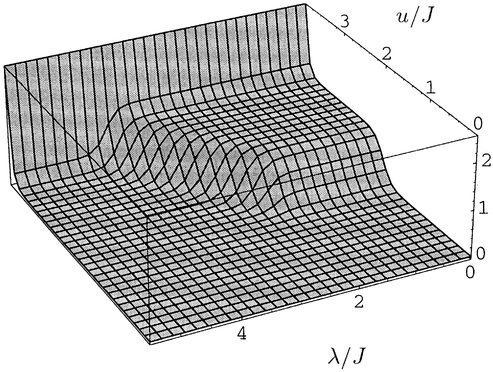

Fig. 6. Plot of $\frac{d}{\beta\, du} \log w_1(u)$ at $\beta = 10$.

the presence of the thick parts containing 0-spins. Using the second equation in (3.6) we translate the "small angles condition"

$$\left| \frac{d}{\beta\, du} \log w_1(u) \right| \leqslant \tan(\pi/4 - \eta)$$

into (recall (2.12)–(2.13))

$$u \in \mathcal{O}_{\delta K} \equiv \left\{ u \in \mathcal{O}_\delta : |u| < \mu_0 + \frac{K}{\beta} \right\} \tag{3.7}$$

$$\mu_0 \equiv \max(\lambda, 2J) \tag{3.8}$$

with some finite $K = K(\eta)$. Note that for $\lambda < 4J$ inclination of an interface comes exclusively from 12- and 22-elements (i.e., containing 0-spins):

$$\frac{d}{\beta\, du} \log w_1(u) = \frac{2 \sinh(\beta u)}{e^{\beta\lambda} + e^{2\beta J} + 2\cosh(\beta u)} + O(e^{-(\mu_0 - 4J)\,\beta})$$

uniformly in $u \in \mathcal{O}_{\delta K}$; see also Fig. 6.

By a direct calculation one verifies the following result:

Lemma 3.3. As $\beta \to \infty$, uniformly in $u \in \mathcal{O}_\delta$,

$$\text{trace } \mathscr{F}^{(1)}/w_1 = 1 + O(e^{-4\beta J}) + O(e^{-2\beta J - \beta\lambda}),$$

$$\det \mathscr{F}^{(1)}/(w_1)^2 = \frac{e^{2\beta J + \beta\lambda}}{(e^{\beta\lambda} + e^{2\beta J} + 2\cosh(\beta u))^2} + O(e^{-2\beta J}).$$

Denoting by $\alpha = \alpha_\beta(u)$ the approximation to the rescaled determinant of $\mathscr{F}^{(1)}$,

$$\alpha = \frac{e^{2\beta J + \beta\lambda}}{(e^{\beta\lambda} + e^{2\beta J} + 2\cosh(\beta u))^2}, \tag{3.9}$$

we immediately obtain the (limiting) characteristic equation for the SOS-model under consideration:

$$1 - v + \alpha v^2 = 0. \tag{3.10}$$

Its principal solution will be investigated in details in Section 3.2.2; here we observe only that in view of the obvious inclusion $\alpha \in (0, 1/4)$ this solution is bounded uniformly in real u.

Remark 3.4. Observe that in the limit $\beta \to \infty$ the rescaled determinant of $\mathscr{F}^{(1)}$ survives only for $\lambda = 2J$; otherwise, uniformly in $u \in \mathscr{O}_\delta$,

$$\det \mathscr{F}^{(1)}/(w_1)^2 = O(e^{-\beta\,|2J-\lambda|}) + O(e^{-2\beta J}).$$

3.2.2. Characteristic Equation of the Horizontal Polymer Model

The principal (i.e., the smallest positive for real u) solution to the characteristic equation (3.10), is given by

$$v_1 = v_1(u) = \frac{2}{1 + \sqrt{1 - 4\alpha}}$$

with α defined in (3.9). Observe that in view of the elementary bound

$$0 < \alpha(u) \leqslant \frac{1}{4(1 + \exp\{-\beta\mu_0\}\cosh(\beta u))^2} \leqslant \frac{1}{4(1 + \exp\{-\beta\mu_0\})^2} < \frac{1}{4} \tag{3.11}$$

valid uniformly in all real u, this solution is well defined for all such u and satisfies $v_1 \in (1, 2)$. Moreover, in view of the simple inequality

$$\left| \frac{e^{\beta\lambda} + e^{2\beta J} + 2\cosh(\beta u)}{e^{\beta\lambda} + e^{2\beta J} + 2\cosh(\Re \beta u)} \right| \geqslant 1 - |\Im \beta u|$$

this solution can be extended analytically into the region

$$\left\{ u \in \mathbb{C} : |\Im u| < \frac{\delta}{\beta} \exp\{-\beta\mu_0\} \right\}$$

with any $0 < \delta < 1$ provided β is large enough.

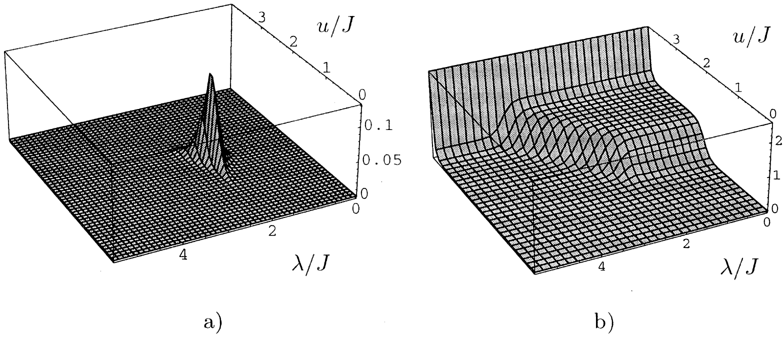

Fig. 7. (a) Plot of $\frac{d}{\beta\, du} \log v_1(u)$; (b) plot of $\frac{d}{\beta\, du} \log \hat{w}_1(u)$ at $\beta = 10$.

By a direct computation, we obtain

$$\frac{1}{\beta}(\log v_1)'_u = \frac{1}{1+\sqrt{1-4\alpha}} \cdot \frac{4\alpha}{\sqrt{1-4\alpha}} \cdot \frac{-2\sinh(\beta u)}{e^{\beta\lambda}+e^{2\beta J}+2\cosh(\beta u)},$$

and thus, in view of Lemma 3.3, the solution

$$\hat{w}_1(u) = \frac{v_1(u)}{\bar{w}_1(u)} = \frac{v_1(u)}{w_1(u)}(1+O(e^{-\beta\mu}))$$

to the approximating characteristic equation (2.18) satisfies the identity

$$-\frac{d}{\beta\, du}\log \hat{w}_1(u) = \frac{2\sinh(\beta u)}{\sqrt{(e^{\beta\lambda}-e^{2\beta J})^2+4(e^{\beta\lambda}+e^{2\beta J}+\cosh(\beta u))\cosh(\beta u)}}$$
$$+\frac{\sinh(\beta u)}{\cosh(4\beta J)-\cosh(\beta u)}+O(e^{-\beta\mu}).$$

As one might have already guessed (recall Remark 3.2), the functions $\hat{w}_1(u)$ and $\bar{w}_1(u)$ are close to each other; indeed, the effect of $v_1(u)$ is visible only in the vicinity of the point $u \sim \lambda \sim 2J$ (cf. Figs. 6 and 7).

It is straightforward to verify that for any $\varepsilon > 0$ there exist two positive constants $K_1(\varepsilon)$ and $K_2(\varepsilon)$ such that $K_1(\varepsilon) \leqslant K_2(\varepsilon)$, $K_1(\varepsilon) \uparrow \infty$ as $\varepsilon \downarrow 0$, and

$$u \in \mathcal{O}_{\delta, K_1} \Rightarrow \left|\frac{d}{\beta\, du}\log \hat{w}_1(u)\right| \leqslant 1-\varepsilon \Rightarrow u \in \mathcal{O}_{\delta, K_2}.$$

3.2.3. The Massgap Condition

We recall that the activity of an aggregate $A = (I_A, \mathscr{C}_A)$ is given by (2.5),

$$\mathfrak{z}(A) = \exp\{-\beta E(I_A)\} \prod_{C \in \mathscr{C}_A} \Psi(C),$$

where the weights $\Psi(C) \equiv \exp\{-\Phi^T(C)\} - 1$ satisfy the estimate (2.3):

$$|\Psi(C)| \leqslant \exp\{-\kappa_0(\beta - \beta')\,|C|\}$$

with some fixed $\beta' < \infty$ and all β large enough. Also, the partition function $\mathfrak{z}_\ell^{a,b}(u)$ is given by (2.8):

$$\mathfrak{z}_\ell^{a,b}(u) = \sum_{A \in \mathbf{A}_\ell^{a,b}} e^{h(A)\,\beta u} \mathfrak{z}(A).$$

We start with the case of small angles. Recall that this corresponds to the fulfilling the condition $u \in \mathcal{O}_{\delta K}$ (see (3.7)–(3.8), (2.12)–(2.13)):

$$|u| < v \equiv v(\delta, K) := \min\left(\mu - \frac{\delta}{\beta}, \mu_0 + \frac{K}{\beta}\right). \tag{3.12}$$

With $\bar{w}_1(u)$ defined as in (2.17) we verify the massgap property (2.19):

Lemma 3.5. Let $\ell \geqslant 3$, $a, b \in \{1, 2\}$ or $\ell = 2$ and $(a, b) \neq (2, 2)$. There exist finite $\beta_0 = \beta_0(J, \lambda, \delta, K)$, $\bar{\beta}_0 = \bar{\beta}_0(J, \lambda, \delta, K)$, and positive $\kappa = \kappa(J, \lambda)$ such that the inequality

$$\left| \frac{\mathfrak{z}_\ell^{a,b}(u)}{(\bar{w}_1(u))^\ell} \right| \leqslant \exp\{-\kappa(\beta - \beta_0)(\ell - 1)\} \tag{3.13}$$

holds, for $\beta \geqslant \bar{\beta}_0$, uniformly in $u \in \mathcal{O}_{\delta K}$.

Remark 3.6. For u under consideration one has $|\cosh(\beta u)|/\cosh(\beta v) \leqslant e^{-\delta/2}$ once $\beta \geqslant \tilde{\beta}_0(\delta) > 0$ with a suitable chosen $\tilde{\beta}_0(\delta)$ and therefore, for some positive constant $C(\delta)$, the approximation $w_1(u)$ to $\bar{w}_1(u) = \mathsf{trace}\, \mathscr{F}^{(1)}(u)$ (recall (3.5)) satisfies

$$\exp\{-C(\delta)\} \leqslant |w_1(u)| \exp\{\beta\mu\} \leqslant \exp\{C(\delta)\}$$

uniformly in $u \in \mathcal{O}_{\delta K}$. Thus, in view of Lemma 3.3, we need only to verify the following simplified inequality

$$\left|\frac{\mathfrak{z}_n^{a,b}(u)}{\exp\{-n\beta\mu\}}\right| \leqslant \exp\{-\kappa(\beta-\tilde{\beta}_0(\delta))(n-1)\} \tag{3.14}$$

uniformly in u under consideration (and possibly with a larger constant $\tilde{\beta}_0$).

The remaining part of this section contains the proof of the bound (3.14). Our argument is close in spirit to that of Lemma 4.7 in ref. 10 and consists of several steps. The main idea is to rewrite the partition function $\mathfrak{z}_\ell^{a,b}(u)$ as (recall (2.8))

$$\mathfrak{z}_\ell^{a,b}(u) = \sum_{I \in \mathscr{I}_\ell^{a,b}} \sum_{A \in \mathbf{A}_\ell^{a,b} : I_A = I} e^{h(A)\beta u} \mathfrak{z}(A)$$

with $\mathscr{I}_\ell^{a,b}$ denoting the collection of all (a, b)-interfaces of horizontal projection ℓ,

$$\mathscr{I}_\ell^{a,b} = \bigcup_{|\theta| < \pi/2,\, M \geqslant 1} \mathscr{I}_{\ell,M}^{\theta,a,b},$$

followed by performing first the inner summation over all agregates $A = (I_A, \mathscr{C}_A)$ with fixed support I_A and then verifying the inequality (3.14) depending on the "regularity" properties of the support I_A.

We start with the following simple observation. Let $A = (I_A, \mathscr{C}_A)$ be any aggregate from $\mathbf{A}_\ell^{a,b}$. We say that a column is regular for the support I_A if it is regular for $(I_A, \varnothing)$ (i.e., it becomes regular after erasing all the "decorations" coming from $\mathscr{C}_A$) and denote by $r = r(I_A)$ and $j = j(I_A)$ the numbers of regular and irregular columns of the support I_A correspondingly. Clearly, the width $w(A)$ of A satisfies

$$w(A) = r(I_A) + j(I_A) + 1.$$

According to our definition of irreducibility, every regular column m of I_A intersects a cluster $C_m \in \mathscr{C}_A$ attached to the support I_A. Therefore,

$$\sum_{C \in \mathscr{C}_A} |C| \geqslant r(I_A). \tag{3.15}$$

Step 1. Let any $I \in \mathscr{I}_\ell^{a,b}$ be fixed. We claim that for any $\varepsilon > 0$ there is a finite $\bar{\beta} = \bar{\beta}(\varepsilon)$ such that

$$\sum_{A \in \mathbf{A}_\ell^{a,b} : I_A = I} |e^{\beta h(A) u} \mathfrak{z}(A)| \leqslant e^{-\beta E(I) + \beta h(I) \Re u - 2\kappa_0 r(I)(\beta - \bar{\beta})} e^{\varepsilon |\Delta(I)|}, \tag{3.16}$$

where, as before, $\Delta(I)$ denotes the collection of all attached points to the support $I_A = I$ of the aggregate A.

Indeed, due to (2.5) and (2.3),

$$\sum_{A \in \mathbf{A}_\ell^{a,b} : I_A = I} |e^{\beta h(A) u} \mathfrak{z}(A)| \leqslant e^{-\beta E(I) + \beta h(I) \Re u} X(I)$$

with

$$X(I) = \sum e^{-\kappa_0 (\beta - \beta') |C|},$$

where β' is, for our model, an absolute constant and the sum goes over all aggregates $A = (I_A, C_A) \in \mathbf{A}_\ell^{a,b}$ such that each $C \in \mathscr{C}_A$ satisfies $C \cap \Delta(I) \neq \varnothing$ and the condition (3.15) is fulfilled. Consequently, for β_1 large enough (for details, see ref. 10, Lemma 4.7),

$$\begin{aligned} X(I) &\leqslant e^{-2\kappa_0 (\beta - \beta_0 - \beta_1) r(I)} \left(1 + \sum_{C \ni 0} e^{-\kappa_0 \beta_1 |C|} \right)^{|\Delta(I)|} \\ &\leqslant e^{-2\kappa_0 (\beta - \bar{\beta}) r(I)} e^{\varepsilon(\beta_1) |\Delta(I)|}, \end{aligned}$$

where $\bar{\beta} = \beta_0 + \beta_1$ and $\varepsilon(\beta_1) \searrow 0$ as $\beta_1 \nearrow \infty$. Inequality (3.16) follows.

Step 2. We observe next that any $I = I_A \in \mathscr{I}_\ell^{a,b}$ splits in a natural way into $r(I) + 1$ irreducible components I_j. Using the simple estimate

$$E(I) \geqslant \kappa_1 |\Delta(I)|$$

that holds with some fixed positive constant κ_1 (uniformly in λ under consideration), we see that the right hand side of (3.16) is bounded by

$$\exp\{2\kappa_0 (\beta - \bar{\beta})\} \prod_{j=1}^{r(I)+1} \mathfrak{z}(I_j),$$

where positive activities $\mathfrak{z}(I)$ are given by (cf. (2.5))

$$\begin{aligned} \mathfrak{z}(I) &= \exp\{-\beta E(I) + \beta h(I) \Re u + \varepsilon |\Delta(I)|\} e^{-2\kappa(\beta - \bar{\beta})} \\ &\leqslant \exp\{-(\beta - \varepsilon_1) E(I) + \beta v |h(I)|\} e^{-2\kappa(\beta - \bar{\beta})} \end{aligned} \tag{3.17}$$

with $\varepsilon_1 \equiv \varepsilon / \kappa_1$. In what follows we will assume that ε (and thus also $\bar{\beta}(\varepsilon)$) is chosen in such a way that $\varepsilon\mu < \delta / 2\kappa_1$ holds (recall (3.12)). This generates another one-dimensional polymer system with the weights given by the RHS of (3.17). Note that the corresponding polymers are labeled, the latter determined by the very geometry of the polymers according to the types of the regular columns at their ends.

Let $\bar{\mathscr{I}}_\ell^{ab} \subset \mathscr{I}_\ell^{a,b}$ denote the collection of all (geometrically) irreducible interfaces from $\mathscr{I}_\ell^{ab}$. In view of (3.17), the proof of the lemma follows directly from the following estimate:

There exist finite $\beta_0 = \beta_0(J, \lambda, \delta, K)$, $\bar{\beta}_0 = \bar{\beta}_0(J, \lambda, \delta, K)$, and positive $\kappa = \kappa(J, \lambda)$ such that the inequality

$$\sum_{I \in \bar{\mathscr{I}}_\ell^{ab}} e^{-(\beta - \varepsilon_1)\, E(I) + \beta\nu\, |h(I)|} \leqslant e^{-(\beta - \beta_0)(\mu\ell + \kappa(\ell - 1))} \tag{3.18}$$

holds for all $\beta \geqslant \bar{\beta}_0$.

From now on we shall concentrate ourselves on the proof of the bound (3.18). The idea behind our formal argument below is as follows. For any fixed boundary conditions $a, b \in \{1, 2\}$ and an integer $\ell \geqslant 2$, we introduce an auxiliary subclass $\tilde{\mathscr{I}}_\ell^{ab} \subset \bar{\mathscr{I}}_\ell^{ab}$ of "reduced" interfaces that possess the next two properties

(a) for some finite constant $C = C(\delta) > 1$, the inequality

$$\sum_{I \in \bar{\mathscr{I}}_\ell^{ab}} e^{-(\beta - \varepsilon_1)\, E(I) + \beta\nu\, |h(I)|} \leqslant \sum_{I \in \tilde{\mathscr{I}}_\ell^{ab}} e^{-(\beta - \varepsilon_1)\, E(I) + \beta\nu\, |h(I)|} C^{|h(I)|} \tag{3.19}$$

holds true uniformly for all integer $\ell \geqslant 2$, $a, b \in \{1, 2\}$, and β large enough;

(b) for any positive λ, J, and δ, there exist finite $\beta_0 = \beta_0(\lambda, J, \delta, K) > 0$, $\bar{\beta}_0 = \bar{\beta}_0(\lambda, J, \delta, K) > 0$, and positive $\kappa = \kappa(\lambda, J)$ such that the estimate

$$\sum_{I \in \tilde{\mathscr{I}}_\ell^{ab}} e^{-(\beta - \varepsilon_1)\, E(I) + \beta\nu\, |h(I)|} C^{|h(I)|} \leqslant e^{-(\beta - \beta_0)(\mu\ell + \kappa(\ell - 1))} \tag{3.20}$$

holds uniformly for all integer $\ell \geqslant 2$, $a, b \in \{1, 2\}$, and any $\beta \geqslant \bar{\beta}_0$.

The massgap bound (3.18) is an immediate corollary from these two properties. We proceed now to the formal proof.

Step 3. Fix any $a, b \in \{1, 2\}$, an integer $\ell \geqslant 2$, and consider arbitrary interface $I \in \bar{\mathscr{I}}_\ell^{ab}$. Recall that an interface I is (geometrically) reducible in a column $x = m$ if the restriction of I to this column consists either of one horizontal bond or of a pair of horizontal bonds and a single zero spin between them; we call any such column $x = m$ *regular* (for the interface I). An interface is called (geometrically) irreducible, if it contains no regular columns.

In a similar way (by rotating the coordinate axes by $\pi/2$) we define regular rows. We shall call two neigbouring (regular) rows *connected* if one of them is the vertical translate of another and the interface has no

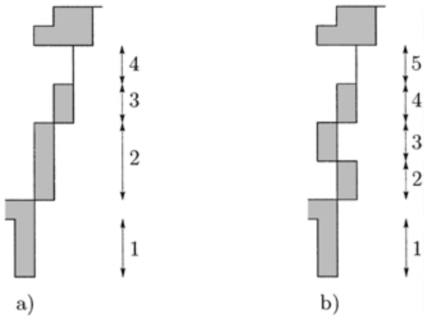

Fig. 8. Blocks of regular rows.

horizontal bonds in the common borderline of these rows; in particular, both rows are of the same type. Clearly, this "connectivity" relation splits the set of all regular rows of the interface I into equivalence classes; we shall call the latter *blocks* (of regular rows). A block is called *large* if it contains more than two rows; a block that is not large is called *small*. According to this definition, the interface in Fig. 8a has four blocks of regular columns, and the one in Fig. 8b—five such blocks.

In addition, we shall call a block *internal*, if it lies inside the horizontal strip determined by the endpoints of the interface; otherwise the block is called *external*. As we shall see below, only internal large blocks require our attention.

We say that an interface $I \in \bar{\mathscr{I}}_\ell^{ab}$ is (vertically) *reduced* if all its internal blocks of regular rows are small; the subclass of all (vertically) reduced interfaces in $\bar{\mathscr{I}}_\ell^{ab}$ is denoted $\tilde{\mathscr{I}}_\ell^{ab}$. Thus, the interface in Fig. 8b is reduced whereas the one in Fig. 8a is not as its second block has size four.

Finally, we define a projection $\mathscr{R}$ from $\bar{\mathscr{I}}_\ell^{ab}$ to $\tilde{\mathscr{I}}_\ell^{ab}$, $\mathscr{R} : I \mapsto \mathscr{R}(I)$ by reducing **each** internal large block of $I \in \bar{\mathscr{I}}_\ell^{ab}$ to two rows only; see Fig. 9 for an example.

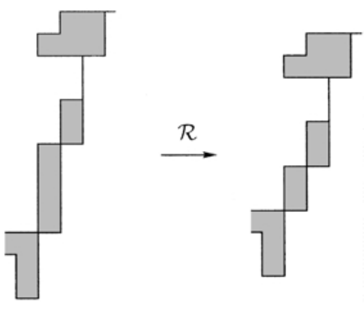

Fig. 9. Reduction of the interface from Fig. 8a.

Now, fix any $I \in \tilde{\mathscr{J}}_\ell^{ab}$, and let $R = R(I) = R_1 + R_2$ be the total number of its internal blocks containing exactly two regular rows, R_1 blocks of type 1 and R_2 blocks of type 2. If $R(I) = 0$, the mapping $\mathscr{R}$ is one-to-one; otherwise $\mathscr{R}$ is infinity-to-one and acts by reducing the size $k+2$, $k > 0$, of any internal large block to two (extremal) rows, see Fig. 9. As a result, for any interface $I \in \tilde{\mathscr{J}}_\ell^{ab}$ we obtain

$$\sum_{\substack{I' \in \bar{\mathscr{J}}_\ell^{ab} \\ \mathscr{R}(I') = I}} e^{-(\beta-\varepsilon_1)\, E(I') + \beta v h(I')} = e^{-(\beta-\varepsilon_1)\, E(I) + \beta v h(I)} C_1^{R_1} C_2^{R_2},$$

where

$$C_1 := \sum_{k \geqslant 0} e^{-[(\beta-\varepsilon_1)\, 4J - \beta v]\, k} \leqslant \sum_{k \geqslant 0} e^{-\delta k/2} = \frac{1}{1-e^{-\delta/2}},$$

$$C_2 := \sum_{k \geqslant 0} e^{-[(\beta-\varepsilon_1)(\lambda+2J) - \beta v]\, k} \leqslant \sum_{k \geqslant 0} e^{-\delta k/2} = \frac{1}{1-e^{-\delta/2}}.$$

Observing that $R(I) = R_1 + R_2 \leqslant |h(I)|/2$ we deduce (3.19).

Step 4. Our proof of (3.20) is based on two facts. The first of them is a simple combinatorial observation (see, e.g., Lemma 5.4 of ref. 25):

> Let $G = (V, E)$ be a graph such that every vertex $v \in V$ of G is of finite degree $d_v \leqslant d < \infty$. Fix any $v_0 \in V$ and denote by $\mathscr{G}_n(v_0)$ the collection of connected subgraphs G_n of G on n vertices such that $V(G_n) \ni v_0$. Then, for some positive $A \leqslant ed$, the collection $\mathscr{G}_n(v_0)$ contains no more than A^n elements.

Our second ingredient is the uniform energetic bound on $I \in \tilde{\mathscr{J}}_\ell^{ab}$. As before, we use here the following notations (see (2.13), (3.8), and (3.12))

$$\mu \equiv 2J + \min(\lambda, 2J), \qquad \mu_0 \equiv \max(\lambda, 2J), \qquad v_0 \equiv \min(\mu, \mu_0).$$

Lemma 3.7. For any $\lambda > 0$, $J > 0$, there exists

$$\bar{\kappa} = \bar{\kappa}(\lambda, J) \geqslant \min(\lambda, 2J)/13 > 0$$

such that the energy $E(I)$ of any geometrically irreducible interface $I \in \tilde{\mathscr{J}}_\ell^{ab}$ satisfies the inequality

$$E(I) \geqslant v_0\, |h(I)| + \mu w(I) + \bar{\kappa}[(|h(I)|-1)_+ + w(I) - 1]. \tag{3.21}$$

Here, as before, $h(I)$ and $w(I)$ denote the height and the width of I, respectively; also, for real x we write $(x)_+ \equiv \max(x, 0)$.

We postpone the proof of Lemma 3.7 till Section 3.2.4 and finish here our derivation of (3.14). First, we note that for any $\beta_0 > 0$ and any interface $I \in \tilde{\mathscr{I}}_\ell^{ab}$, we have

$$\begin{aligned}
-(\beta-\beta_0)\,E(I)+\beta v\,|h(I)|+(\beta-\beta_0)[\mu\ell+\bar{\kappa}(\ell-1)]\\
\leqslant [\beta_0 v_0+\beta(v-v_0)]\,|h(I)|-(\beta-\beta_0)\,\bar{\kappa}(|h(I)|-1)_+\\
\leqslant (\beta_0 4J+K)\,|h(I)|-(\beta-\beta_0)\,\bar{\kappa}(|h(I)|-1)_+\\
\leqslant \beta_0 4J+K \leqslant (\beta-\beta_0)\,\bar{\kappa}(\ell-1)/2
\end{aligned}$$

provided β is large enough, $\beta \geqslant \bar{\beta}_0$ to satisfy $(\beta-\beta_0)\,\bar{\kappa}/2 \geqslant \beta_0 4J+K$. Thus, to prove (3.20), it remains to find an upper bound of the sum

$$\sum_{I \in \tilde{\mathscr{I}}_\ell^{ab}} e^{-(\beta_0-\varepsilon_1)\,E(I)} C^{|h(I)|}. \tag{3.22}$$

To this end, observe that each $I \in \tilde{\mathscr{I}}_\ell^{ab}$ is a connected subgraph of the graph $G=(V, E)$ such that:

- its set V of vertices consists of all unit bonds between the sites of the integer lattice $\mathbb{Z}^2$ as well as the vertices of the dual lattice $(\mathbb{Z}^2)^* = \mathbb{Z}^2+(1/2, 1/2)$;
- its set E of edges is generated by the following adjacency relation: two bonds from $\mathbb{Z}^2$ are neighbors if they share a point; a dual site is adjacent to a bond from $\mathbb{Z}^2$ if the euclidean distance between them equals $1/2$.

Clearly, the degree of any "vertex" $v \in V$ is bounded above by 8; moreover, each contour I consists of at least $w(I)+|h(I)|$ such vertices. Observe also that if I contains n "vertices," its energy $E(I)$ is bounded below by $\mathbf{e}n$ with $\mathbf{e} := \min(\lambda, J) > 0$. An upper bound on the sum (3.22) thus is

$$\sum_{n \geqslant \ell} \exp\{-(\beta_0-\varepsilon_1)\,\mathbf{e}n\}\, C^n A^n < 1,$$

where β_0 is such that $\exp\{-\mathbf{e}(\beta_0-\varepsilon_1)\}\, CA < 1/2$. We finally obtain (cf. (3.20))

$$\sum_{I \in \tilde{\mathscr{I}}_\ell^{ab}} e^{-(\beta-\varepsilon_1)\,E(I)} e^{\beta v\,|h(I)|} C^{|h(I)|} \leqslant \exp\{-(\beta-\beta_0)[\mu\ell+\kappa(\ell-1)]\}$$

with $\kappa = \bar{\kappa}/2$.

The proof of (3.14) is finished.

Remark 3.8. Observe that the argument above is quite general. Some particular properties of the model at hand were used only when proving the estimates (3.19) and (3.20); see the very definition of the ensemble $\widetilde{\mathcal{I}}_{\ell}^{ab}$ together with the projection mapping $\mathcal{R}$: $\overline{\mathcal{I}}_{\ell}^{ab} \to \widetilde{\mathcal{I}}_{\ell}^{ab}$ in Step 3 and the energy bound (3.21) used in Step 4. Thus, when studying the diagonal SOS model we may and shall present only the model-dependent part of the proof.

3.2.4. Proof of Lemma 3.7

Our aim here is to prove the energy bound (3.21) for irreducible interfaces from $\widetilde{\mathcal{I}}_{\ell}^{ab}$. Note that it is enough to verify it for the "cheapest" connections only. We start with several preparatory remarks.

Recalling our surgery procedure from Section 2 we see that any intersection of a vertical line $x = k \in \mathbb{Z}$ with a contour I consists of one or several pieces each of them being either a single bond between $+$ and $-$, or a connected component of zero sites together with two attached bonds. Thus, the value of the "energy of a vertical intersection" belongs to the set

$$\mathcal{E} \equiv \left\{ 4Jk_0 + \sum_{j=1}^{\infty} (2J + j\lambda)\, k_j \right\} \Big\backslash \{8J\},$$

where $k_0, k_1, \ldots$ are nonnegative integers (with at least one of them being positive) and j denoting the number of zeroes. The regularity condition reads:

$$k_0 + k_1 = 1, \qquad k_2 = k_3 = \cdots = 0 \tag{3.23}$$

and the energy of the "cheapest" column configuration is

$$\mathbf{e}_1 = \min(4J, \lambda + 2J) = \mu.$$

If the condition (3.23) fails, the column configuration is non-regular, its energy being bounded below by

$$\mathbf{e}_2 = \min\left[\mathcal{E} \setminus \{0, 4J, \lambda + 2J\}\right] = \begin{cases} 2J + 2\lambda, & 0 < \lambda \leqslant 4J, \\ 6J + \lambda, & 4J \leqslant \lambda \leqslant 6J, \\ 12J, & 6J \leqslant \lambda. \end{cases} \tag{3.24}$$

Once the "horizontal energy" has been taken into account, we are left with the vertical bonds only and for each row $y = k \in \mathbb{Z}$ the minimal price is either $2J$ (in case the row contains zeroes) or $4J$ otherwise. We note however, that sometimes it will be preferable to interchange the role of the vertical and the horizontal "energies" by attaching the zeroes to the former.

For future references, let us recall (see (3.12)) that

$$0<\lambda\leqslant 2J: \qquad v_0=2J \quad \text{and} \quad \mu=2J+\lambda, \tag{3.25}$$

$$2J\leqslant\lambda\leqslant 4J: \qquad v_0=\lambda \quad \text{and} \quad \mu=4J, \tag{3.26}$$

$$4J\leqslant\lambda: \qquad v_0=4J \quad \text{and} \quad \mu=4J. \tag{3.27}$$

In what follows we consider different cases depending on the values of the parameter λ, the height $h(I)$ and the width $w(I)$ of an interface I. Since the energy $E(I)$ is an additive functional, we can consider separately the energy of the extremal half-columns of I and the energy of the "interior" I^* of I. If the height Δh of an extremal half-column satisfies $\Delta h=0$ (see Fig. 3c and the left end in Fig. 3a, then the corresponding energy is $2J$ for the configuration σ^1 and $J+\lambda/2$ for σ^2; in both cases we have

$$2J\geqslant\frac{\mu}{2}\equiv\mu\,\Delta w+v_0\,|\Delta h|, \qquad J+\lambda/2\geqslant\frac{\mu}{2}\equiv\mu\,\Delta w+v_0\,|\Delta h|.$$

Otherwise $|\Delta h|=1/2$ and the corresponding energy equals $2J+\lambda/2$ (see Fig. 3a, right end) thus giving

$$2J+\frac{\lambda}{2}\geqslant\frac{1}{2}\,(v_0+\mu)\equiv v_0\,|\Delta h|+\mu\,\Delta w.$$

As a result, we need only to verify the bound

$$E(I^*)\geqslant(v_0+\kappa)\,|h(I^*)|+(\mu+\kappa)\,w(I^*). \tag{3.28}$$

In the remaining part of the proof we will work only with the interior I^* of any interface I and will use the simplified notations I, h, and w for I^*, $h(I^*)$, and $w(I^*)$ respectively. Note that in view of symmetry it is enough to consider only non-decreasing interfaces (i.e., with $h\equiv h(I)\geqslant 0$).

The energy E_0 of a cheapest connection is a piece-wise continuous linear function; in view of finiteness of the entropy it is sufficient to study only the open intervals of its linearity (the increased entropy at the border points of such intervals will be eventually suppressed by $\beta\kappa$ with β large enough).

Case 1. $0<\lambda<4J$, $h\leqslant 12w$.

The cheapest irreducible horizontal connection here costs $\mathbf{e}_2=2J+2\lambda$ (corresponding to exactly two zeroes in each column); on the other hand, the absolute lowest price for the vertical bonds is $2J$ for every row (provided there is at least one zero in each of them; otherwise it is at least $4J$). Consequently,

$$E(I) \geqslant (2J+2\lambda)\, w + 2Jh.$$

Let $0 < \lambda < 2J$; using (3.25), we obtain

$$E(I) - v_0 h - \mu w \geqslant \lambda w \geqslant \frac{\lambda}{13}(w+h). \tag{3.29}$$

For $2J < \lambda < 4J$, we use (3.26) and get

$$E(I) - v_0 h - \mu w \geqslant 2Jw \geqslant \frac{2J}{13}(w+h). \tag{3.30}$$

Case 2. $4J < \lambda$, $h \leqslant 12w$.

For $j \in \mathbb{N}$, let l_j denote the number of columns with exactly j zeroes. Then the width w of the interface and the total number N of zeroes in it satisfy

$$w = \sum_{j=0}^{\infty} l_j \quad \text{and} \quad N = \sum_{j=1}^{\infty} j l_j.$$

Let $m = m(I) \leqslant N$ be the number of rows in I containing zeroes; then the total energy of zeroes and the vertical bonds is bounded below by

$$\lambda N + 2Jm + 4J(h-m)_+ \geqslant \lambda N - 2Jm + 4Jh \geqslant 2JN + 4Jh.$$

Next, the energy of the horizontal bonds is not smaller than

$$12Jl_0 + 6Jl_1 + 2J \sum_{j=2}^{\infty} l_j = 10Jl_0 + 4Jl_1 + 2Jw.$$

Using the obvious inequality $2w \leqslant 2l_0 + l_1 + N$, we get

$$E(I) \geqslant 2JN + 4Jh + 2Jw + 10Jl_0 + 4Jl_1 \geqslant 6Jw + 4Jh$$

and thus (3.27) implies

$$E(I) - v_0 h - \mu w \geqslant 2Jw \geqslant \frac{2J}{13}(w+h). \tag{3.31}$$

Case 3. $0 < \lambda < 2J$, $h > 12w$.

Here we count zeroes in horizontal sections (and use the absolute lower bound $\mathbf{e}_1 = \mu$ for each row configuration). Since the contribution of the horizontal edges in each column is not smaller than $2J$, we obtain $E(I) \geqslant 2Jw + \mu h$. As a result (recall (3.25)),

$$E(I) - v_0 h - \mu w \geqslant \lambda h - \lambda w \geqslant \frac{11\lambda}{13}(w+h). \tag{3.32}$$

Case 4. $2J < \lambda < 4J$, $h > 12w$.

If the interface I contains exactly k regular rows, the total energy of zeroes and the vertical bonds is bounded below by

$$\mathbf{e}_1 k + \mathbf{e}_2 (h-k)_+ \geqslant 4Jk + (2J+2\lambda)(h-k)_+.$$

On the other hand, if any $I \in \widetilde{\mathscr{I}}_\ell^{ab}$ contains k regular columns, there are at least $k/2$ blocks of regular columns and thus the total energy of the horizontal bonds is at least $\frac{J}{2}k$. As a result, we obtain

$$E(I) \geqslant 4Jk + (2J+2\lambda)(h-k)_+ + \frac{J}{2}k \geqslant \frac{9J}{2}h$$

and therefore (recall (3.26))

$$E(I) - v_0 h - \mu w \geqslant \left(\frac{9J}{2} - \lambda\right)h - 4Jw \geqslant \frac{J}{2}h - 4Jw \geqslant \frac{2J}{13}(w+h). \tag{3.33}$$

Case 5. $4J < \lambda$, $h > 12w$.

Arguing as in the preceding case, we get

$$E(I) \geqslant \mathbf{e}_1 k + \mathbf{e}_2 (h-k)_+ + \frac{J}{2}k \geqslant \frac{9J}{2}k + 10J(h-k)_+ \geqslant \frac{9J}{2}k$$

and deduce

$$E(I) - v_0 h - \mu w \geqslant \left(\frac{9J}{2} - 4J\right)h - 4Jw \geqslant \frac{J}{2}h - 4Jw \geqslant \frac{2J}{13}(w+h). \tag{3.34}$$

Clearly, the estimates (3.29)–(3.34) imply (3.28) with $\kappa = \min(2J, \lambda)/13$. The proof of the lemma is finished.

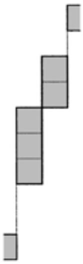

Fig. 10. A 22-jump of width 3.

3.3. "Almost Diagonal" Directions $\theta \in (\eta, \frac{\pi}{2} - \eta)$

3.3.1. Solid-On-Solid Approximation

The SOS-model corresponding to (3.1) is well adapted to describe the low-temperature $\pm$-interfaces with inclination angles $|\theta| < \pi/4 - \eta$ for $\lambda \leqslant 2J$ and $|\theta| < \pi/2 - \eta$ for $\lambda > 2J$ (with uniform estimates on compact subsets in this region). Otherwise (i.e., for $\lambda \leqslant 2J$ *and* $\theta \sim \pi/4$) it has to be modified by including the whole family of *wide* (i.e., of horizontal projection $k \geqslant 2$) irreducible 22-jumps. A typical example of such a jump is given in Fig. 10.

The corresponding partition function is easy to compute:

$$\mathfrak{z}_\ell^{2,2}(u) = G_\ell(u) + G_\ell(-u), \qquad \ell \geqslant 2,$$

with

$$G_\ell(u) = e^{-(4\beta J + \beta\lambda) + \beta u} [\bar{g}_1(u)]^\ell \, [\bar{g}_2(u)]^{\ell - 1},$$

where[7]

$$\bar{g}_1(u) = \sum_{k \geqslant 0} [e^{-4\beta J + \beta u}]^k = \frac{1}{1 - e^{-4\beta J + \beta u}},$$

$$\bar{g}_2(u) = \sum_{k \geqslant 2} e^{-2\beta J} [e^{-(2\beta J + \beta\lambda) + \beta u}]^k = \frac{e^{-(6\beta J + 2\beta\lambda) + 2\beta u}}{1 - e^{-(2\beta J + \beta\lambda) + \beta u}}$$

[7] Note that the condition $u \in \mathcal{O}_\delta$ with some $\delta > 0$ (recall (2.12)–(2.13)) is necessary and sufficient for the series to converge.

correspond to the thin and the thick vertical pieces respectively. Now, denote

$$F_2(u, w) = \sum_{\ell \geqslant 2} G_\ell(u)\, w^\ell = \frac{e^{-(4\beta J + \beta\lambda) + \beta u} (\bar{g}_1(u))^2\, \bar{g}_2(u)\, w^2}{1 - \bar{g}_1(u)\, \bar{g}_2(u)\, w},$$

the series being convergent for w small enough. Clearly, $F_2(u, w) + F_2(-u, w)$ represents the generating function of the wide 22-elements; thus, a good SOS-approximation is given by (cf. (3.1))

$$\hat{\mathscr{F}}(u, w) = \begin{pmatrix} \mathfrak{z}_1^{1,1}(u) & \mathfrak{z}_1^{1,2}(u) \\ \mathfrak{z}_1^{2,1}(u) & \mathfrak{z}_1^{2,2}(u) \end{pmatrix} w + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} (F_2(u, w) + F_2(-u, w)).$$

This model provides an adequate low-temperature approximation of $\pm$-interfaces in the whole region $\{(\lambda, \theta) : \lambda > \varepsilon, |\theta| < \pi/2\}$ and can be analyzed by a method similar to that in the previous section and based on the multidimensional analogue of Proposition A.6 under Assumption A.8. The latter, however, requires an accurate analysis of analytic properties of the corresponding generating functions—not entirely easy task. However, for the directions close to the diagonal (say, for $|\theta - \pi/4| < \pi/4 - \eta$) our model has an equivalent "diagonal" representation analogous (and even simpler) to the one used for almost horizontal interfaces.

To define the "diagonal" polymer model (corresponding to $\theta \approx \pi/4$) we rotate the coordinate axes by $\pi/4$ anti-clockwise. In the new coordinate system all zero-temperature interfaces can be decomposed into irregular components by the "vertical" lines (with the step $1/\sqrt{2}$ of course). A direct analysis of Fig. 11 gives us the complete list of regular jumps of our diagonal SOS-approximation, see Fig. 12.

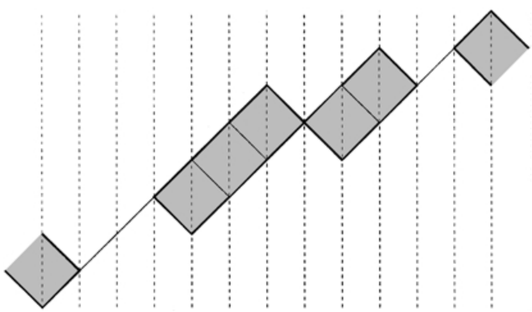

Fig. 11. Canonical decomposition of a diagonal interface.

Fig. 12. Regular diagonal jumps of width 1.

We stress that in the "new" coordinate system the basic lattice step is changed by a factor $1/\sqrt{2}$ in both direction; on the other hand, since the asymptotics of the surface tension for fixed directions is scale-invariant (while keeping the same energetic price of the dual edges), the result does not change if we scale our picture by a factor $\sqrt{2}$.

As a result, our list of regular jumps consists of: two 11-jumps of height ± 1, one 12-jump of height 0, one 21-jump of height 0, and two 22-jumps of height ± 1, see Fig. 12; then the corresponding diagonal generating function is given by

$$\mathscr{F}_{\mathrm{diag}}(u, w) \equiv \begin{pmatrix} 2e^{-4\beta J} \cosh \beta u & e^{-2\beta J - \beta\lambda/2} \\ e^{-2\beta J - \beta\lambda/2} & 2e^{-2\beta J - \beta\lambda} \cosh \beta u \end{pmatrix} w \tag{3.35}$$

and the related characteristic equation (2.14), $\det[\mathbb{1} - \mathscr{F}_{\mathrm{diag}}(u, w)] = 0$, reads

$$(1 - 2we^{-4\beta J} \cosh(\beta u))(1 - 2we^{-2\beta J - \beta\lambda} \cosh(\beta u)) = w^2 e^{-4\beta J - \beta\lambda}. \tag{3.36}$$

3.3.2. Characteristic Equation of the Diagonal Polymer Model

We turn now to the study of the characteristic equation (3.36). This quadratic equation is easy to solve, its principal solution w_2 being given by

$$w_2 = \frac{e^{2\beta J}}{(e^{-\beta\lambda} + e^{-2\beta J}) \cosh(\beta u) + \sqrt{e^{-\beta\lambda} + (e^{-\beta\lambda} - e^{-2\beta J})^2 \cosh^2(\beta u)}}.$$

As a result,

$$-\frac{d}{\beta\, du} \log w_2 = \tanh \beta u - \frac{e^{-\beta\lambda}}{e^{-\beta\lambda} + (e^{-\beta\lambda} - e^{-2\beta J})^2 \cosh^2(\beta u)} \times \left(1 + \frac{(e^{-\beta\lambda} + e^{-2\beta J}) \cosh(\beta u)}{\sqrt{e^{-\beta\lambda} + (e^{-\beta\lambda} - e^{-2\beta J})^2 \cosh^2(\beta u)}}\right)^{-1} \tanh \beta u$$

implying that

$$\left|\frac{d}{\beta\, du} \log w_2(u)\right| < 1$$

for any real u; moreover,

$$\left|\frac{d}{\beta\, du} \log w_2\right| = 1-\varepsilon \quad \Leftrightarrow \quad |u| \sim \min(\lambda, 2J) - \frac{\lambda}{2} + \frac{K}{\beta}$$

with $K = K(\varepsilon)$ satisfying $K \uparrow \infty$ as $\varepsilon \downarrow 0$.

In our study of the diagonal model we shall restrict ourselves to the region

$$\mathcal{O}_K^{\mathsf{diag}} := \left\{ u\colon |\Re u| \leqslant \min(\lambda, 2J) - \frac{\lambda}{2} + \frac{K}{\beta} \right\}; \tag{3.37}$$

in addition, we consider only the values of $\lambda \in (0, 4J]$. The region $\lambda > 4J$ is covered by horizontal case for all angles. For future reference we observe that the bound

$$0 < C(K) \leqslant w_2 e^{-2\beta J - \beta\lambda/2} \leqslant 1 \tag{3.38}$$

holds uniformly for $u \in \mathcal{O}_K^{\mathsf{diag}}$ with a constant C depending only on K.

3.3.3. The Massgap Condition for the Diagonal Polymer Model

To finish our study of the diagonal model, we shall verify the following massgap condition:

> Let a finite K be fixed; then, for any $\lambda \in (0, 4J)$, there exists a positive constant $\kappa = \kappa(\lambda, J, K)$ such that uniformly in a, $b \in \{1, 2\}$ and in $u \in \mathcal{O}_K^{\mathsf{diag}}$, one has the bound
>
> $$\left|\frac{\mathfrak{z}_n^{a,b}(u)}{(w_2(u))^n}\right| \leqslant \exp\{-\kappa(\beta - \beta_0)(n-2)\}.$$

Here, as everywhere below in this section, $\mathfrak{z}_n^{a,b}(u)$ denotes the diagonal analogue of the partition function for the irreducible animals of horizontal projection n (note that accordingly to our surgery rule, such irreducible animals exist only when $n \geqslant 3$).

Combining the uniform bound (3.38) and Remark 3.8 we easily see that the diagonal massgap condition is equivalent to

$$|\mathfrak{z}_n^{a,b}(u)\, e^{(2\beta J+\beta\lambda/2)\,n}| \leqslant \exp\{-\kappa(\beta-\beta_0)(n-2)\}. \tag{3.39}$$

In the sequel, we shall concentrate ourselves on the proof of (3.39).

Observe that the energy of each column of a diagonal interface belongs to the set

$$\mathscr{E}_{\mathsf{diag}} \equiv \left\{4Jk_0 + \sum_{j=1}^{\infty}\left(2J+\frac{\lambda 2}{j}\right)k_j : k_0, k_j \geqslant 0\right\} \setminus \{0, 8J\};$$

also, for $J>0$, $0<\lambda<4J$, we have

$$\begin{aligned}
\mathbf{e}_1 &\equiv \min \mathscr{E}_{\mathsf{diag}} = 2J+\lambda/2,\\
\mathbf{e}_2 &\equiv \min\,[\mathscr{E}_{\mathsf{diag}} \setminus \{4J, 2J+\lambda/2, 2J+\lambda\}] = 2J+3\lambda/2,\\
\upsilon_0 &\equiv \min(2J, \lambda)-\lambda/2.
\end{aligned}$$

By repeating the arguments of the first two steps in Section 3.2.3, we easily reduce (3.39) to the following property:

Let $\bar{\mathscr{I}}^{ab}_{\mathsf{diag},\,n}$ denote the collection of all irreducible diagonal interfaces of horizontal projection n and the boundary conditions a, b.

Then there exist finite $\beta_0 = \beta_0(J, \lambda, K)$, $\bar{\beta}_0 = \bar{\beta}_0(J, \lambda, K)$, and a positive $\kappa = \kappa(J, \lambda)$ such that the inequality

$$\sum_{I \in \bar{\mathscr{I}}^{ab}_{\mathsf{diag},\,n}} e^{-(\beta-\varepsilon_1)\,E(I)+(\beta\upsilon_0+K)\,|h(I)|} \leqslant e^{-(\beta-\beta_0)(\mu n+\kappa(n-1))}$$

holds for all $\beta \geqslant \bar{\beta}_0$.

The proof of this property is a literal repetition of the argument in step 4 of Section 3.2.3, once the following energy bound is verified.

Lemma 3.9. For any $J>0$, $0<\lambda<4J$, there exists

$$\bar{\kappa} = \bar{\kappa}(J, \lambda) \geqslant \frac{\min(\lambda, 2J)}{7} > 0$$

such that the energy $E(I)$ of any geometrically irreducible interface $I \in \bar{\mathscr{I}}^{ab}_{\mathsf{diag},\,n}$ satisfies the inequality

$$E(I) \geqslant \upsilon_0\,|h(I)| + \mathbf{e}_1 w(I) + \bar{\kappa}[(|h(I)|-1)_+ + w(I)-1].$$

Proof. Fix any $I \in \bar{\mathscr{J}}^{ab}_{\mathrm{diag},n}$.

First, we verify easily that the energy of any of the jumps in Fig. 12 is bounded from below by $\mathbf{e}_1 \Delta w + v_0 |\Delta h|$; thus, exactly as in Section 3.2.4, we are left with the internal parts of the irreducible interfaces only. As for the latter, we shall consider two cases,

$$|h| \leqslant \tfrac{4}{3} w \qquad \text{and} \qquad |h| \geqslant \tfrac{4}{3} w$$

separately. Here and below, h and w denote the internal height and internal width of the interface I correspondingly (cf. Section 3.2.4).

Case 1. $|h| \leqslant 4w/3$.

We use the absolute bound $\mathbf{e}_2$ for the energy of an irreducible column configuration and obtain:

$$E(I) - \mathbf{e}_1 w - v_0 \, |h| \geqslant \lambda w - v_0 \, |h| \geqslant \frac{\lambda}{3} w \geqslant \frac{\lambda}{7} (w + |h|).$$

Case 2. $|h| \geqslant 4w/3$.

Here we turn our interface in $\pi/4$ clockwise and observe that each diagonal (h, w)-interface is mapped onto a horizontal (w', h')-interface such that

$$w' = \frac{w + |h|}{2}, \qquad |h'| \geqslant \frac{|h - w|}{2}.$$

According to the argument in Case 3 of Section 3.2.4, the energy of any horizontal (w', h')-connection is bounded below by

$$E(I) \geqslant 2J(w' + |h'|) + \lambda \max(w', |h'|) \geqslant 2J \max(w, |h|) + \lambda \frac{w + |h|}{2}.$$

Using now the bound $|h| \geqslant 4w/3$, we easily obtain

$$E(I) - \mathbf{e}_1 w - v_0 \, |h| \geqslant \left(2J + \frac{\lambda}{2} - v_0 \right) |h| - 2Jw$$

$$\geqslant 2J(|h| - w) \geqslant \frac{2J}{7} (w + |h|).$$

The proof of the lemma is finished. ∎

APPENDIX A. LINEAR POLYMER MODELS: THE SIMPLEST CASE

Models of polymer type[26] appear in a natural way in many interesting situations. Together with cluster expansions (that were created and adapted mainly to study such models) they form a powerful tool with a large area of applicability—from statistical mechanics and probability to combinatorics and number theory. Though, due to their perturbation nature, cluster expansions are usually limited to certain areas (eg, sufficiently small or large temperatures, large values of external fields etc.) they nevertheless provide there complete information about the corresponding partition functions, distributions etc.

One particular case—the so-called linear (or "one-dimensional") polymer models are well adapted to work with "interfaces:" trajectories of self-avoiding walks (SAW), percolation clusters, phase boundary in the 2D Ising model etc. Here the cluster expansions have their analogue in the renewal theory and sometimes (SAW, percolation) this makes possible to prove the results in the whole subcritical region.[6–9, 27]

The aim of this section is to reformulate known results about the cluster expansions from the point of view of the renewal theory in the way that makes possible their further generalization to polymer systems with polymers of different type (in particular, labeled polymers; see Appendix B).

We say that a sequence of (complex) numbers $Z(n)$ forms a sequence of partition functions of the linear polymer model with the (complex) weights z_l, $l=1,2,\ldots$, if for any $n \geqslant 0$ on has (by definition, we put $Z(0)=1$)

$$Z(n)=\sum_{k=1}^{n} \sum_{\substack{m_i,\, i=1,\ldots,k:\\ m_i \geqslant 1,\, \sum m_i = n}} \prod_{i=1}^{k} z_{m_i}.$$

A simple combinatorics shows then that the generating functions of the sequences z_l and $Z(n)$ are related,

$$\mathscr{F}(w)=\sum_{l=1}^{+\infty} z_l w^l, \qquad \mathscr{Z}(w) \equiv \sum_{n=0}^{+\infty} Z(n)\, w^n = \frac{1}{1-\mathscr{F}(w)}. \tag{A.1}$$

We start with the following simple observation.

Proposition A.1. Assume that

$$\limsup_{l\to\infty} \sqrt[l]{|z_l|} = \frac{1}{R} < +\infty,$$

i.e., the ball $B(0, R)$ is the circle of analyticity of the function $\mathscr{F}(\cdot)$ from (A.1). Then the radius of analyticity r of $\mathscr{Z}(\cdot)$ is nonvanishing, $r \leqslant R$, and for any ρ, $0 < \rho < r$, we have the relation

$$Z(n) = \frac{1}{2\pi i} \oint_{|w| = \rho} \frac{w^{-n-1}}{1 - \mathscr{F}(w)} \, dw \tag{A.2}$$

Moreover, if

$$0 < r < R, \tag{A.3}$$

then

$$\limsup_{n \to \infty} \frac{\log |Z(n)|}{n} = \log \limsup_{n \to \infty} \sqrt[n]{|Z(n)|} = -\log r < +\infty.$$

Remark A.2. Condition (A.3), known as the mass separation condition,[6, 8, 9] guarantees the important analytic properties of the generating function necessary to develop a consistent perturbation theory; as such, it plays the central rôle in our considerations. Of course, the formulated results hold under less stringent assumptions, see, e.g., ref. 28.

Corollary A.3. Assume, in addition, that all z_l are non-negative numbers such that the greatest common divisor of all those l such that $z_l > 0$ equals 1 (in particular, at least one of z_l's is positive). Then for all sufficiently large n the numbers $Z(n)$ are positive and the following limit exists,

$$f \equiv \lim_{n \to \infty} \frac{\log Z(n)}{n} = \sup_{n} \frac{\log Z(n)}{n}, \tag{A.4}$$

the function f satisfies the relation $f = -\log r$, and

$$\lim_{n \to \infty} Z(n) \, e^{-nf} = (e^{-f} \mathscr{F}'(e^{-f}))^{-1}.$$

Proof. The proof is a standard one; for details, see ref. 20, App. A and ref. 21, p. 330. ▮

Remark A.4. Let $a \neq 0$ be any number. Then (A.2) can be rewritten in the form

$$Z(n) = \frac{a^n}{2\pi i} \oint_{|w| = \tilde{\rho}} \frac{w^{-n-1}\, dw}{1 - \mathscr{F}(w/a)} \tag{A.5}$$

with any $\tilde{\rho}$, $0 < \tilde{\rho} < |a|\, r$.

Corollary A.3 identifies the free energy f from (A.4) with the (minus) logarithm of the radius r of analyticity of the generating function $\mathscr{Z}(\cdot)$ from (A.1), which, in its turn is equal to the distance from the origin to the closest solution of the equation

$$1 - \mathscr{F}(w) = 0.$$

In general such an equation cannot be solved explicitly. In certain cases however, when the weights z_k depend on other parameters (temperature, external fields etc.), the proper change of variables (see Remark A.4 above) makes possible to approximate the function $\mathscr{F}(w/a)$ by a simpler one. We illustrate this approach for the interface of the 2D Ising model in Remark A.7 later.

Suppose now that the weights z_l depend on the inverse temperature $\beta < \infty$ and (complex-valued, multidimensional) "external field" u, i.e., $z_l = z_l(\beta, u)$. Let $z_0(\beta, u)$ be a function such that the limits

$$\lim_{\beta \to \infty} \frac{z_l(\beta, u)}{(z_0(\beta, u))^l} = a_l(u)$$

exist and are nonnegative for real u. Then, for $a = z_0(\beta, u)$, the function $\widetilde{\mathscr{F}}(w) \equiv \mathscr{F}(w/a)$ is well approximated by the function

$$\widetilde{\mathscr{F}}_0(w) = \sum_{l=1}^{\infty} a_l w^l.$$

For our future considerations we need the following assumption.

Assumption A.5. There exist $\beta_0 < \infty$, an integer number $m \geqslant 1$, non-negative sequence a_l, $l = 1, 2, \ldots$, a constant $C = C(\beta_0) > 0$, and a positive function $\alpha = \alpha(\beta)$, $\alpha(\beta) \searrow 0$ as $\beta \nearrow \infty$, such that:

(i) $a_m > 0$ and $a_l = 0$ for all $l > m$;

(ii) for any $l \geqslant 1$ and all β, $\beta_0 \leqslant \beta < \infty$, the estimate

$$\left|\frac{z_l(\beta, u)}{(z_0(\beta, u))^l} - a_l\right| \leqslant C\alpha(\beta)^l$$

holds uniformly in u from some (complex) neighbourhood $\mathcal{O}_0 = \mathcal{O}(\beta_0)$.

Under such an assumption the function $\tilde{\mathcal{F}}_0(\cdot)$ is a polynomial,

$$\tilde{\mathcal{F}}_0(w) = \sum_{l=1}^{m} a_l w^l,$$

that approximates the function of interest $\tilde{\mathcal{F}}(\cdot)$ well enough:

$$|\tilde{\mathcal{F}}(w) - \tilde{\mathcal{F}}_0(w)| \leqslant C \sum_{l \geqslant 1} |\alpha(\beta)\, w|^l = \frac{C\, |\alpha(\beta)\, w|}{1 - |\alpha(\beta)\, w|}, \tag{A.6}$$

the RHS approaching zero uniformly in w from any compact set in $\mathbb{C}^1$, provided only that β is sufficiently large to guarantee the estimate $|\alpha(\beta)\, w| < 1$. Observe also that for $\tilde{\mathcal{F}}_0$ one has $0 < r < +\infty \equiv R$, i.e., the mass separation condition (A3) holds.

Proposition A.6. Under Assumption A.5 there exists $\beta' \in [\beta_0, \infty)$, such that for all $\beta \geqslant \beta'$: the mass separation condition $0 < r(\beta, u) < R(\beta, u)$ is valid uniformly in $u \in \mathcal{O}_0$, the free energy

$$f(\beta, u) \equiv \lim_{n \to \infty} \frac{\log Z(n, \beta, u)}{n}$$

exists and satisfies the relation

$$-f(\beta, u) = \log z_0(\beta, u) + \log r + \tilde{f}(\beta, u),$$

where r stands for the smallest positive solution to the equation $\tilde{\mathcal{F}}_0(u) = 1$ and $\tilde{f}(\beta, u)$ is an analytic function satisfying, as $\beta \nearrow \infty$, the condition

$$\tilde{f}(\beta, u) = O(\alpha(\beta)).$$

Remark A.7. The quantity $-\log z_0(\beta)$ plays the role of the free energy in the ensemble of tame animals. Recall that one can rederive the results from ref. 10 concerning the surface tension of the low-temperature

Ising model by taking $m=1$, $z_0(\beta)\equiv z_1(\beta)$, $\tilde{\mathscr{F}}_0(w)\equiv w$, $C=\alpha(\beta)^{-1}$, $\alpha(\beta)=\exp\{4(\beta_0-\beta)\}$, with the summation in (A.6) starting with $l=2$, and thus the RHS of (A.6) getting an additional factor w.

Finally, we note that the claim of Proposition A.6 holds also under the following conditions:

Assumption A.8. There exists a function $\tilde{\mathscr{F}}_0(w)=\sum_{l=1}^{\infty} a_l w^l$ with non-negative coefficients such that

1. $\tilde{\mathscr{F}}_0(w)$ is analytic in a ball $|w|<R_0$.

2. The smallest positive solution r_0 to the equation $\tilde{\mathscr{F}}_0(w)=1$ satisfies $0<r_0<R_0$.

3. For all β large enough, the function $\tilde{\mathscr{F}}(w)=\sum_{l=1}^{\infty} z_l(z_0)^{-l}\,w^l$ is analytic in a ball $|w|<r$ with some $r>r_0$; moreover, there exists r', $r_0<r'<\min(r, R_0)$, such that the difference $|\tilde{\mathscr{F}}(w)-\tilde{\mathscr{F}}_0(w)|$ converges to zero (as $\beta\to\infty$) uniformly in $w\in[0, r']$.

APPENDIX B. LINEAR POLYMER MODELS WITH LABELED POLYMERS

Let $\mathscr{B}=\{1, 2,..., M\}$ be a set of labels (boundary conditions) where $M>1$ is some fixed natural number. For any pair $\{i, j\}$, $i, j\in\mathscr{B}$, consider the generating function

$$F_{ij}(w)=\sum_{n\geqslant 1} f_n^{(i,j)}w^n \tag{B.1}$$

of a sequence $f_n^{(i,j)}\geqslant 0$ and assume that this function is analytic in the circle $|w|<R_{ij}$, $R_{ij}>0$. Consider also the (matrix-valued) generating function

$$\mathscr{F}(w)=(F_{ij}(w))_{i,j=1}^{M}, \qquad \mathscr{F}(0)=\mathbf{0}.$$

Then the quantity $R=\min_{i,j} R_{ij}>0$ represents its radius of analyticity.

Consider also the function[8] (with $\mathbb{1}$ denoting the identity matrix)

$$\mathscr{Z}(w)\equiv\sum_{m\geqslant 0}[\mathscr{F}(w)]^m=[\mathbb{1}-\mathscr{F}(w)]^{-1}=(Z_{ij}(w))_{i,j=1,M}, \tag{B.2}$$

[8] Observe that $\mathscr{F}$ plays the role of the usual transfer-matrix of a one-dimensional system with the only difference that in our case the (irreducible) jumps can be of horizontal projection $n\geqslant 1$.

where

$$Z_{ij}(w) = \sum_{n \geqslant 0} Z_n^{(i,j)} w^n$$

is the generating function of the sequence $Z_n^{(i,j)}$ of (finite volume) partition functions with boundary conditions $\{i, j\}$. We are interested in studying the asymptotic behaviour of the sequence $Z_n^{(i,j)}$. In particular, we would like to investigate the existence of the free energies

$$f_{ij} \equiv \lim_{n \to \infty} \frac{\log Z_n^{(i,j)}}{n}. \tag{B.3}$$

Lemma B.1. Let all $f_n^{(i,j)} \geqslant 0$ for all i, j and the conditions

$$Z_n^{(i,j)} > 0 \tag{B.4}$$

hold for all $i, j \in \mathscr{B}$ and all sufficiently large n, $n \geqslant n_0$. Assume also that the smallest positive root r of the equation

$$\det [\mathbb{1} - \mathscr{F}(w)]|_{w=r} = 0$$

satisfies the relation

$$0 < r < R, \tag{B.5}$$

where R is the radius of analyticity of the function $\mathscr{F}(w)$.

Then for all pairs $\{i, j\}$ one has

$$f_{ij} \equiv f := -\log r, \tag{B.6}$$

i.e., the free energies defined in (B.3) do not depend on the boundary conditions and coincide with the logarithm of the (inverse) radius of analyticity of the matrix-valued generating function $\mathscr{Z}(w)$ from (B.2).

Proof. Due to assumption (B.4) all entries of the matrix $\mathscr{Z}(w)$ have the same radius of analyticity r. Thus, relation (B.6) is a direct consequence of the mass separation condition (B.5) and the general theorem of the renewal theory (ref. 16, p. 331). ∎

Remark B.2. The radius r of analyticity of the matrix-valued function $\mathscr{Z}(w)$ has another characterization in terms of the spectral radius of the matrix $\mathscr{F}(w)$. Namely, if **spr** A denotes the spectral radius of a square matrix A,

$$\mathbf{spr}(A) = \limsup_{n\to\infty} \|A^n\|^{1/n},$$

then the radius of analyticity of the function $\mathscr{Z}(w)$ is given by the only positive solution r to the equation $\mathbf{spr}(\mathscr{F}(w)) = 1$. Now, for positive w, due to non-negativity of the coefficients $Z_n^{(i,j)}$ and the Perron–Frobenius theorem, the maximal positive eigenvalue $\lambda(w)$ of the matrix $\mathscr{F}(w)$ is simple and depends analytically on w, the rest of the spectrum being inside the ball $|z| < \lambda(w)$. Since the spectral radius of a matrix gives the maximal absolute value of its eigenvalues, we get $\lambda(|w|) \equiv \mathbf{spr}\ \mathscr{F}(|w|)$, and therefore the radius r of analyticity of $\mathscr{Z}(w)$ solves the equation $\lambda(r) = \mathbf{spr}\ \mathscr{F}(r) = 1$. For future reference we note also that $\lambda'(r) > 0$ for positive r.

Remark B.3. Condition (B.5), known also as the condition of mass separation (see refs. 6, 8, and 9), holds automatically if all $F_{ij}(w)$ are polynomials of bounded degree. Another sufficient condition is given by the following statement.[9]

Lemma B.4. Assume that the coefficients $f_n^{(i,j)}$ in (B.1) are analytic functions of some positive parameter β, i.e., $f_n^{(i,j)} = f_n^{(i,j)}(\beta) \geqslant 0$, $\beta \geqslant \beta_0 > 0$, and for all such β and all $i, j \in \mathscr{B}$ one has

$$Z_n^{(i,j)}(\beta) > 0 \tag{B.7}$$

provided n is large enough. Assume, further, that there exist a polynomial matrix-valued function

$$\mathscr{F}_0(w) = (\bar{F}_{ij}(w))_{i,j=1,M}, \qquad \mathscr{F}_0(0) = \mathbf{0}, \tag{B.8}$$

of finite degree $m = \max_{ij} \deg \bar{F}_{ij}(w)$ with non-negative coefficients not depending on β, an analytic function $z_0(\beta) > 0$, and a positive function $\gamma = \gamma_\varepsilon(\beta)$, $\gamma_\varepsilon(\beta) \searrow 0$ as $\beta \nearrow +\infty$ for all ε small enough, $\varepsilon \in (0, \varepsilon_0)$, such that for $\beta \geqslant \beta_0$ and $\varepsilon \in (0, \varepsilon_0)$ one has

$$\max_{i,j} \sup_{|w| < (1+\varepsilon) r_0} |F_{ij}(w/z_0(\beta)) - \bar{F}_{ij}(w)| \leqslant \gamma_\varepsilon(\beta), \tag{B.9}$$

[9] This can be easily generalized by the reader.

where as before

$$r_0 = \min\{r > 0 : \det [\mathbb{1} - \mathscr{F}_0(w)]|_{w=r} = 0\}. \tag{B.10}$$

Then for all sufficiently large $\beta > 0$ the mass separation condition $0 < r(\beta) < R(\beta)$ holds and the corresponding free energy

$$f_{ij}(\beta) \equiv \lim_{n\to\infty} \frac{\log Z_n^{(i,j)}(\beta)}{n}$$

exists and satisfies the relation

$$-f_{ij}(\beta) = \log z_0(\beta) + \log r_0 + \tilde{f}(\beta),$$

where $\tilde{f}(\beta)$ is an analytic function of β satisfying the condition $\tilde{f}(\beta) = O(\gamma_\varepsilon(\beta))$ as $\beta \to \infty$.

Proof. Denote $\tilde{\mathscr{F}}(w) \equiv \mathscr{F}(w/z_0(\beta))$. The quantity

$$\tilde{r}(\beta) = \min\{r > 0 : \det [\mathbb{1} - \tilde{\mathscr{F}}(w)]|_{w=r} = 0\}$$

clearly satisfies the relation $\tilde{r}(\beta) = z_0(\beta)\, r(\beta)$. On the other hand, it presents the unique positive solution to the equation $\mathbf{spr}(\tilde{\mathscr{F}}(w)) = 1$. Here again the spectral radius $\mathbf{spr}(\tilde{\mathscr{F}}(|w|))$ identically coincides with the maximal positive (simple) eigenvalue $\tilde{\lambda}(|w|, \beta)$ of the matrix $\tilde{\mathscr{F}}(|w|)$. The claim of the lemma now easily follows from the analytic implicit function theorem, the identity $\tilde{\lambda}(|w|, 0) \equiv \lambda_0(|w|)$, and the a priori estimate (B.9) taking into account that in view of Remark B.2 one has the inequality

$$\left.\frac{\partial\tilde{\lambda}(w,\beta)}{\partial w}\right|_{(w,\beta)=(r,0)} \equiv \left.\frac{d\lambda_0(w)}{dw}\right|_{w=r} > 0. \qquad \blacksquare$$

Using the same arguments one can prove also the following statement.

Corollary B.5. Let the coefficients $f_n^{(i,j)}$ in (B.1) be analytic functions of $\beta > 0$ and a complex variable $H \in \mathbb{C}^k$, $(H, \beta) \in \mathcal{O} \times (\beta_0, +\infty)$, where $\mathcal{O}$ denotes some region in $\mathbb{C}^k$. Assume also that for all real H from $\mathcal{O}$ and all $\beta \geqslant \beta_0$ the condition (B.7) holds and for some polynomial matrix-valued function $\mathscr{F}_0(w)$ (recall (B.8)) with constant non-negative coefficients the approximability condition (B.9) holds for some $z_0 = z_0(H, \beta)$ (which is positive for real H) uniformly in $H \in \mathcal{O}$.

Then for all sufficiently large $\beta \geqslant \beta_0$ the mass separation condition $0 < r(H, \beta) < R(H, \beta)$ holds, the corresponding free energy

$$f_{ij}(H, \beta) \equiv \lim_{n \to \infty} \frac{\log Z_n^{(i,j)}(H, \beta)}{n}$$

exists, it is an analytic function of H and β, and can be rewritten in the form

$$-f_{ij}(H, \beta) \equiv f(H, \beta) = \log z_0(H, \beta) + \log r_0 + \tilde{f}(H, \beta),$$

where $\tilde{f}(H, \beta)$ is an analytic function of H and β satisfying the condition

$$\tilde{f}(H, \beta) = O(\gamma_\varepsilon(\beta))$$

as $\beta \to \infty$.

ACKNOWLEDGMENTS

O.H. thanks Dima Ioffe and Rostyslav Hryniv for interesting discussions. We also thank the referees for useful comments. O.H. was partly supported by DFG postdoctoral fellowship. R.K. was partly supported by the Grants GA ČR 201/00/1149 and MSM 110000001.

REFERENCES

1. R. L. Dobrushin, Gibbs state describing coexistence of phases for a three-dimensional Ising model, *Teor. Ver. Pril.* **17**:619–639 (1972); English transl. in *Theor. Probability Appl.* **17**:582–600 (1972).
2. P. Holický, R. Kotecký, and M. Zahradník, Rigid interfaces for lattice models at low temperatures, *J. Statist. Phys.* **50**:755–812 (1988).
3. S. Pirogov and Ya. Sinai, Phase diagrams of classical lattice systems, *Theoret. and Math. Phys.* **25**:1185–1192 (1975) and **26**:39–49 (1976).
4. P. Holický, R. Kotecký, and M. Zahradník, Phase diagram of horizontally invariant Gibbs states for lattice models, *Ann. Henri Poincaré*, in press.
5. G. Gallavotti, The phase separation line in the two-dimensional Ising model, *Commun. Math. Phys.* **27**:103–136 (1972).
6. M. Campanino, J. T. Chayes, and L. Chayes, Gaussian fluctuations of connectivities in the subcritical regime of percolation, *Probab. Theory Related Fields* **88**:269–341 (1991).
7. M. Campanino and D. Ioffe, Ornstein–Zernike theory for the Bernoulli bond percolation on $\mathbb{Z}^d$, preprint (1999).
8. J. T. Chayes and L. Chayes, Ornstein–Zernike behaviour for self-avoiding walks at all non-critical temperatures, *Commun. Math. Phys.* **105**:221–238 (1986).
9. D. Ioffe, Ornstein–Zernike behaviour and analyticity of shapes for self-avoiding walks on $\mathbb{Z}^d$, *Markov Process. Related Fields* **4**:323–350 (1998).

10. R. L. Dobrushin, R. Kotecký, and S. B. Shlosman, *Wulff Construction: A Global Shape from Local Interaction.* (Translations of Mathematical Monographs, Vol. 104) (Amer. Math. Soc., 1992).
11. J. Bricmont and J. Slawny, Phase transitions in systems with a finite number of dominant ground states, *J. Statist. Phys.* **54**:89–161 (1989).
12. J. Bricmont and J. Lebowitz, Wetting in Potts and Blume–Capel models, *J. Statist. Phys.* **46**:1015–1029 (1987).
13. R. L. Dobrushin, A statistical behaviour of shapes of boundaries of phases, in *Phase Transitions: Mathematics, Physics, Biology ...*, R. Kotecký, ed. (World Scientific, Singapore, 1993), pp. 60–70.
14. R. L. Dobrushin and O. Hryniv, Fluctuations of the phase boundary in the 2D Ising ferromagnet, *Commun. Math. Phys.* **189**:395–445 (1997).
15. O. Hryniv, On local behaviour of the phase separation line in the 2D Ising model, *Probab. Theory Related Fields* **110**:91–107 (1998).
16. R. Kotecký and D. Preiss, Cluster expansions for abstract polymer models, *Commun. Math. Phys.* **103**:491–498 (1986).
17. D. B. Abraham and P. Reed, Interface profile of the Ising ferromagnet in two dimensions, *Commun. Math. Phys.* **49**:35–46 (1976).
18. Y. Higuchi, On some limit theorems related to the phase separation line in the two-dimensional Ising model, *Z. Wahrsch. Verw. Gebiete* **50**:287–315 (1979).
19. J. Bricmont, J. L. Lebowitz, and C. E. Pfister, On the local structure of the phase separation line in the two-dimensional Ising system, *J. Statist. Phys.* **26**:313–332 (1981).
20. E. Seneta, *Non-negative Matrices and Markov Chains* (Springer, 1981).
21. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed. (Wiley, 1970).
22. N. G. de Bruijn, *Asymptotic Methods in Analysis*, Bibliotheca Mathematica, Vol. IV (North-Holland, 1970).
23. M. V. Fedoriuk, *Asymptotics: Integrals and Series* (Nauka, Moscow, 1987).
24. R. L. Dobrushin and S. B. Shlosman, Large and moderate deviations in the Ising model, in *Probability Contributions to Statistical Mechanics*, pp. 91–219, Adv. Soviet Math., Vol. 20 (Amer. Math. Soc., 1994).
25. R. L. Dobrushin, Perturbation methods of the theory of Gibbsian fields, in *Lectures on Probability Theory and Statistics (Saint-Flour, 1994)*, pp. 1–66, Lecture Notes in Math., Vol. 1648 (Springer, Berlin, 1996).
26. C. Gruber and H. Kunz, General properties of polymer systems, *Commun. Math. Phys.* **22**:133–161 (1971).
27. O. Hryniv and D. Ioffe, Self-avoiding polygons: Sharp asymptotics of canonical partition functions under the fixed area constraint, preprint (2001).
28. M. Pinsky, A note on the Erdös–Feller–Pollard theorem, *Amer. Math. Monthly* **83**:729–731 (1976).