Journal of Statistical Physics, Vol. 83, Nos. 5/6, 1996

Constrained Variational Problem with Applications to the Ising Model

Roberto H. Schonmann[1] and Senya B. Shlosman[2]

Received June 16, 1995; final October 24, 1995

We continue our study of the behavior of the two-dimensional nearest neighbor ferromagnetic Ising model under an external magnetic field h, initiated in our earlier work. We strengthen further a result previously proven by Martirosyan at low enough temperature, which roughly states that for finite systems with $(-)$-boundary conditions under a positive external field, the boundary effect dominates in the system if the linear size of the system is of order B/h with B small enough, while if B is large enough, then the external field dominates in the system. In our earlier work this result was extended to every subcritical value of the temperature. Here for every subcritical value of the temperature we show the existence of a critical value $B_0(T)$ which separates the two regimes specified above. We also find the asymptotic shape of the region occupied by the $(+)$-phase in the second regime, which turns out to be a "squeezed Wulff shape." The main step in our study is the solution of the variational problem of finding the curve minimizing the Wulff functional, which curve is constrained to the unit square. Other tools used are the results and techniques developed to study large deviations for the block magnetization in the absence of the magnetic field, extended to all temperatures below the critical one.

KEY WORDS: Ising model; Wulff shape; large deviations; boundary effects.

INTRODUCTION

In this paper we continue our studies of applications of variational techniques to the investigation of the qualitative behavior of the two-dimensional Ising model below the critical temperature in the presence of

[1] Mathematics Department, University of California at Los Angeles, Los Angeles, California 90024.

[2] Mathematics Department, University of California at Irvine, Irvine, California 92717, and Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia.

0022-4715/96/0600-0867$09.50/0

a magnetic field, initiated in ref. 17. The model is defined on $\mathbb{Z}^2$ by the formal Hamiltonian

$$H_h(\sigma) = -\frac{1}{2} \sum_{x,\, y \text{ n.n.}} \sigma(x)\, \sigma(y) - \frac{h}{2} \sum_x \sigma(x) \tag{1.1}$$

where $\sigma(x) = \pm 1$ is the spin at the site $x \in \mathbb{Z}^2$, and the first sum runs over pairs of sites which are nearest neighbors in $\mathbb{Z}^2$, each pair counted only once.

Two key ideas were crucial in ref. 17:

- The study of typical behavior of the Ising model in the magnetic field can be reduced to the study of large-deviation behavior of the Ising model without magnetic field (since by considering nonzero magnetic field we obtain what is called the "tilted distribution" in the theory of large deviations).
- The "typical large-deviation configurations" are described geometrically as the configurations of phase coexistence where one phase forms a droplet floating inside the other, with the shape of the droplet determined by the Wulff variational construction.

The second has been known for some time, restricted, however, to the range of low temperatures.[7] The possibility of extending it to all temperatures below the critical one is due to recent results by Ioffe,[10, 11] who built also on previous work by Alexander *et al.*,[2] Pfister,[13] and Pisztora.[14]

In the present paper we apply these ideas to the following problem: consider the Ising model in the presence of *positive* magnetic field h in a square box $\Lambda(l)$ of size l with *negative* boundary conditions (b.c.). We want to study the result of competing influences of the positive field and the negative b.c. on the Ising system. To give each of the contenders a comparable chance to influence the system, we have to impose the relation

$$l^2 h \sim l$$

(since in two dimensions the volume of the system is l^2, while the boundary has length of the order of l). That means that a reasonable choice is

$$l = B/h$$

with B a constant. To see a sharp transition, one has then to pass to the thermodynamic limit of the infinite volume, which in our case amounts to taking h to 0. The question then is: how does the limiting behavior of the system depend on the parameter B?

Partial answers to this question have been known for some time. Martirosyan showed that at low temperatures T and with $B > B_1(T)$ one would see the $(+)$-phase in the central part of the box $\Lambda(B/h)$ as $h \searrow 0$ (see Theorem 1 in ref. 12, which covers the case of arbitrary dimension d).

Martirosyan's result was reproven in ref. 15 with somewhat simpler methods in arbitrary dimension and with greatly simplified methods in two dimensions. It was also proven there that for some other value $B_2(T) < B_1(T)$ and for $B < B_2(T)$ one would obtain in the limit the $(-)$-phase; the proof given in ref. 15 also implies that the functions $B_1(T)$ and $B_2(T)$ can be taken arbitrarily close to the common value $2d$, provided the temperature is low enough. In ref. 17 this result was extended up to T_c in the two-dimensional case.

The aim of the present paper is to strengthen these results further in the $d=2$ case. Namely, we will show that there exists a unique function $B_0(T)$ which represents a sharp border between these two behaviors: in the above statements one can use the same function $B_0(T)$ instead of both $B_1(T)$ and $B_2(T)$. In Theorem 1 of ref. 17 the same result was obtained for the case of the box having the Wulff shape, corresponding to the temperature T. That result might appear somewhat artificial because of the extravagant choice of the shape of our box; it turns out, however, that the case of the Wulff-shaped box was singled out as the one which leads to the usual (unconstrained) Wulff variational problem, while all other cases require for their analysis information about certain constrained variational problems. Another important difference between the cases of the Wulff-shaped box and the square box is that for any value of $B > B_0(T)$, however large, a certain positive fraction of the square box $\Lambda(B/h)$ will be occupied with the $(-)$-phase, which is not the case for the Wulff-shaped box. The Wulff-shaped boxes, while artificial as objects of interest for their own right, proved nevertheless to be good enough tools to obtain the result stated in the title of ref. 17: namely, that complete analyticity for nice boxes holds everywhere in the interior of the uniqueness region in the phase diagram of the model. The analysis of the behavior in the square boxes was postponed to the present paper because it requires a substantial amount of extra work.

The variational problem that we have to solve here is the following *constrained Wulff problem*: we want to find the shape of a droplet with a given volume (area in 2D) which minimizes the boundary surface tension, under the additional restriction that the shape sought has to fit inside a given box. For the Wulff-shaped box the answer is evident: it has to be the Wulff shape itself. For boxes of different geometry the answer is not so simple. The solution, as a function of the droplet volume, might exhibit singularities, and these singularities make the behavior of the model interesting.

2. NOTATION AND TERMINOLOGY

The lattice and the space: The cardinality of a set $\Gamma \subset \mathbb{Z}^2$ will be denoted by $|\Gamma|$. The expression $\Gamma \subset\subset \mathbb{Z}^2$ will mean that Γ is a finite subset of $\mathbb{Z}^2$. Likewise the area (or the volume) of the measurable set $G \subset \mathbb{R}^2$ will be denoted by $|G|$. If $g = \partial G$ is the boundary of the set G, then we use $V(g)$ to denote the volume inside g:

$$V(g) = V(\partial G) = |G| \tag{2.1}$$

For each $x \in \mathbb{Z}^2$, we define the usual norms $\|x\|_p = (|x_1|^p + |x_2|^p)^{1/p}$, $p > 0$ finite, and $\|x\|_\infty = \max\{|x_1|, |x_2|\}$. The distance between two sets $A, B \in \mathbb{Z}^2$ in each one of these norms will be denoted by

$$\operatorname{dist}_p(A, B) = \inf\{\|x - y\|_p : x \in A, y \in B\}$$

In the case $A = \{x\}$, we also write $\operatorname{dist}_p(A, B) = \operatorname{dist}_p(x, B)$. For two bounded subsets $F, G \subset \mathbb{R}^2$ we define the Hausdorff distance between them as

$$\rho_{\mathrm{H}}(F, G) = \max\{\inf\{\varepsilon : F \subset U_\varepsilon(G)\}, \inf\{\varepsilon : G \subset U_\varepsilon(F)\}\}$$

where $U_\varepsilon(\cdot)$ stands for (Euclidean) ε-neighborhood.

The interior and exterior boundaries of a set $\Gamma \subset \mathbb{Z}^2$ will be denoted, respectively, by

$$\partial_{\mathrm{int}} \Gamma = \{x \in \Gamma : \|x - y\|_1 = 1 \text{ for some } y \notin \Gamma\}$$

and

$$\partial_{\mathrm{ext}} \Gamma = \{x \notin \Gamma : \|x - y\|_1 = 1 \text{ for some } y \in \Gamma\}$$

We denote by $S(l)$ the square in $\mathbb{R}^2$ with side l, centered at the origin:

$$S(l) = [-l/2, l/2]^2$$

For lattice squares centered at the origin, we will use the notation

$$\Lambda(l) = \mathbb{Z}^2 \cap S(l)$$

The set of bonds, i.e. (unordered) pairs of nearest neighbors, is defined as

$$\mathbb{B} = \{\{x, y\} : x, y \in \mathbb{Z}^2 \text{ and } \|x - y\|_1 = 1\} \tag{2.2}$$

Given a set $\Gamma \subset\subset \mathbb{Z}^2$ we define also

$$\mathbb{B}_\Gamma = \{\{x, y\} : x, y \in \Gamma \text{ and } \|x - y\|_1 = 1\} \tag{2.3}$$

$$\partial \mathbb{B}_\Gamma = \{\{x, y\} : x \in \Gamma, y \notin \Gamma, \text{ and } \|x - y\|_1 = 1\} \tag{2.4}$$

and

$$\bar{\mathbb{B}}_\Gamma = \{\{x, y\} : x \in \Gamma,\ y \in \mathbb{Z}^2, \text{ and } \|x - y\|_1 = 1\} = \mathbb{B}_\Gamma \cup \partial\mathbb{B}_\Gamma \qquad (2.5)$$

Notions from percolation: A chain is a sequence of distinct sites $x_1,..., x_n$, with the property that for $i = 1,..., n-1$, $\|x_i - x_{i+1}\|_1 = 1$. The sites x_1 and x_n are called the endpoints of the chain $x_1,..., x_n$, and n is its length. A (*)-chain, its endpoints, and its length are defined in the same way, but with $\|\cdot\|_1$ replaced by $\|\cdot\|_\infty$. Informally this means that while chains can only move along bonds of $\mathbb{Z}^2$, (*)-chains can also move along diagonals. A chain or (*)-chain is said to connect two sets if it has one endpoint in each set. A circuit is a chain such that $\|x_1 - x_n\|_1 = 1$. Similarly a (*)-circuit is a (*)-chain such that $\|x_1 - x_n\|_\infty = 1$.

The configurations and observables: At each site in $\mathbb{Z}^2$ there is a spin which can take values -1 and $+1$. The configurations will therefore be elements of the set $\{-1, +1\}^{\mathbb{Z}^2} = \Omega$. Given $\sigma \in \Omega$, we write $\sigma(x)$ for the spin at the site $x \in \mathbb{Z}^2$. Two configurations are especially relevant, the one with all spins -1 and the one with all spins $+1$. We will use the simple notation $(-)$ and $(+)$ to denote them. The single spin space $\{-1, +1\}$ is endowed with the discrete topology and Ω is endowed with the corresponding product topology. The following definition will be important when we introduce finite systems with boundary conditions later on; given $\Gamma \subset\subset \mathbb{Z}^2$ and a configuration $\eta \in \Omega$, we introduce

$$\Omega_{\Gamma,\eta} = \{\sigma \in \Omega : \sigma(x) = \eta(x) \text{ for all } x \notin \Gamma\}$$

Real-valued functions with domain in Ω are called observables. For each observable f, we use the notation $\|f\|_\infty = \sup_{\eta \in \Omega} |f(\eta)|$. Local observables are those which depend only on the values of finitely many spins, more precisely, $f: \Omega \to \mathbb{R}$ is a local observable if there exists a set $S \subset\subset \mathbb{Z}^2$ such that $f(\sigma) = f(\eta)$ whenever $\sigma(x) = \eta(x)$ for all $x \in S$. The smallest S with this property is called the support of f, denoted $\operatorname{supp}(f)$. The topology introduced above on Ω has the nice feature that it makes the set of local observables be dense in the set of all continuous observables.

We introduce the notation $\theta_x(f)$ for the translate by $x \in \mathbb{Z}^2$ of the function f, i.e., $(\theta_x(f))(\sigma) = f(\theta_x^*(\sigma))$, with $\theta_x^*(\sigma)(y) = \sigma(y + x)$.

For the average spin in a set $\Gamma \subset\subset \mathbb{Z}^2$ we will use

$$X_\Gamma(\sigma) = \frac{1}{|\Gamma|} \sum_{x \in \Gamma} \sigma(x)$$

In Ω the following partial order is introduced:

$$\eta \leqslant \zeta \quad \text{if} \quad \eta(x) \leqslant \zeta(x) \qquad \text{for all} \quad x \in \mathbb{Z}$$

The probability measures: We endow Ω also with the Borel σ-algebra corresponding to the topology introduced above. In this fashion, each probability measure μ in this space can be identified by the corresponding expected values $\int f\,d\mu$ of all the local observables f. A sequence of probability measures, $(\mu_n)_{n=1,2,\dots}$ is said to converge weakly to the probability measure ν in case

$$\lim_{n\to\infty}\int f\,d\mu_n=\int f\,d\nu \qquad \text{for every continuous observable } f \tag{2.6}$$

The family of probability measures on Ω will be partially ordered by the following relation: $\mu\leqslant\nu$ if

$$\int f\,d\mu\leqslant\int f\,d\nu \qquad \text{for every continuous nondecreasing observable } f \tag{2.7}$$

Because the local observables are dense in the set of continuous observables, we can restrict ourselves to the local ones in (2.6) and (2.7).

The Gibbs measures: We will consider always the formal Hamiltonian (1.1). In order to give precise definitions, we define, for each set $\Gamma\subset\subset\mathbb{Z}^2$ and each boundary condition $\eta\in\Omega$,

$$\begin{aligned}H_{\Gamma,\eta,h}(\sigma)={}&-\frac{1}{2}\sum_{\{x,y\}\in\mathbb{B}_\Gamma}\sigma(x)\,\sigma(y)-\frac{1}{2}\sum_{\substack{\{x,y\}\in\partial\mathbb{B}_\Gamma\\ y\notin\Gamma}}\sigma(x)\,\eta(y)\\&-\frac{h}{2}\sum_{x\in\Gamma}\sigma(x)\end{aligned}\tag{2.8}$$

where $h\in\mathbb{R}$ is the external field and $\sigma\in\Omega$ is a generic configuration.

Given $\Gamma\subset\subset\mathbb{Z}^2$, $\eta\in\Omega$ and $E\subset\Omega$, we write

$$Z_{\Gamma,\eta,T,h}(E)=\sum_{\sigma\in\Omega_{\Gamma,\eta}\cap E}\exp(-\beta H_{\Gamma,\eta,h}(\sigma))$$

where $\beta=1/T$. We abbreviate $Z_{\Gamma,\eta,T,h}=Z_{\Gamma,\eta,T,h}(\Omega)$.

The Gibbs (probability) measure in Γ with boundary condition η under external field h and at temperature T is now defined on Ω as

$$\mu_{\Gamma,\eta,T,h}(\sigma)=\begin{cases}\dfrac{\exp(-\beta H_{\Gamma,\eta,h}(\sigma))}{Z_{\Gamma,\eta,T,h}} & \text{if } \sigma\in\Omega_{\Gamma,\eta}\\ 0 & \text{otherwise}\end{cases}$$

The Gibbs measures satisfy the following monotonicity relations, to which we will refer as the FKG-Holley inequalities:

If $\eta \leqslant \zeta$, $T_1 \geqslant T_2$, and $h_1 \leqslant h_2$, then for each $\Gamma \subset\subset \mathbb{Z}^2$, $\mu_{\Gamma, \eta, T_1, h_1} \leqslant \mu_{\Gamma, \zeta, T_2, h_2}$.

A Gibbs measure for the infinite system on $\mathbb{Z}^2$ is defined now as any probability measure μ which satisfies the DLR equations in the sense that for every $\Gamma \subset\subset \mathbb{Z}^2$ and μ-almost all $\eta \in \Omega$

$$\mu(\cdot \mid \Omega_{\Gamma, \eta}) = \mu_{\Gamma, \eta, T, h}(\cdot) \tag{2.9}$$

Alternatively and equivalently, Gibbs measures can be defined as elements of the closed convex hull of the set of weak limit points of sequences of the form $(\mu_{\Gamma_i, \eta_i, T, h})_{i=1,2,\ldots}$, where each Γ_i is finite and $\Gamma_i \to \mathbb{Z}^2$, as $i \to \infty$, in the sense that $\bigcup_{i=1}^{\infty} \bigcap_{j=i}^{\infty} \Gamma_j = \mathbb{Z}^2$.

For each value of T and h, $\mu_{\Lambda(l), -, T, h}$ (resp., $\mu_{\Lambda(l), +, T, h}$) converges weakly, as $l \to \infty$, to a probability measure that we will denote by $\mu_{-, T, h}$ (resp., $\mu_{+, T, h}$). If $h \neq 0$, it is known that $\mu_{-, T, h} = \mu_{+, T, h}$, which will then be denoted simply by $\mu_{T, h}$; it is also known that this is the only Gibbs measure for the infinite system in this case. If $h = 0$, the same is true if the temperature is larger than or equal to a critical value $T_c > 0$, and is false for $T < T_c$, in which case one says that there is phase coexistence.

We use the following abbreviations and names:

$$\mu_{-, T, 0} := \mu_{-, T} = \text{the minus phase} = \text{the } (-)\text{-phase}$$

$$\mu_{+, T, 0} := \mu_{+, T} = \text{the plus phase} = \text{the } (+)\text{-phase}$$

For the expected value corresponding to a Gibbs measure $\mu_{\ldots}$ in finite or infinite volume we will use the notation

$$\langle f \rangle_{\ldots} = \int f \, d\mu_{\ldots}$$

where the dots stand for arbitrary subscripts. The spontaneous magnetization at temperature T is defined as

$$m_T^* = \langle \sigma(0) \rangle_{+, T}$$

(Here we are using a common and convenient form of abuse of notation: $\sigma(x)$ is being used to denote the observable which associates to each configuration the value of the spin at the site x in that configuration. This notation will also be used in other places.) It is known that $m_T^* > 0$ if and only if $\mu_{-, T} \neq \mu_{+, T}$.

Surface tension and Wulff shape: The direction-dependent surface tension is defined in the following way. First consider on $\mathbb{R}^2 \times \mathbb{R}^2$ the usual inner product $(x, y) = x_1 y_1 + x_2 y_2$. Let $\mathbb{S}^1 = \{x \in \mathbb{R}^2 \colon \|x\|_2 = 1\}$, and for each vector $\mathbf{n} \in \mathbb{S}^1$, consider the following configuration, to be used as a boundary condition:

$$\eta(\mathbf{n})(x) = \begin{cases} +1 & \text{if } (x, \mathbf{n}) \geqslant 0 \\ -1 & \text{if } (x, \mathbf{n}) < 0 \end{cases}$$

The surface tension in the direction perpendicular to $\mathbf{n}$ is given by

$$\tau_T(\mathbf{n}) = \lim_{l \to \infty} -\frac{1}{\beta \, \|y(l) - z(l)\|_2} \log \frac{Z_{\Lambda(l),\, \eta(\mathbf{n}),\, T,\, 0}}{Z_{\Lambda(l),\, +,\, T,\, 0}}$$

where $y(l)$ and $z(l) = -y(l)$ are the points where the straight line $\{x \in \mathbb{R}^2 \colon (x, \mathbf{n}) = 0\}$ intersects the boundary of the square $\Lambda(l)$. It is known that for each $T < T_c$ the surface tension $\tau_T(\cdot)$ is a continuous strictly positive and finite function.

Some results below are valid for the general surface tension function $\tau(\cdot)$. We will denote the corresponding quantities by the subscript τ. When we will talk about the Ising model quantities, we will use the subscript T instead of the full subscript τ_T.

About the general surface tension function $\tau(\mathbf{n})$ we suppose that it is symmetric with respect to the reflections in coordinate axes, that $\tau(x, y) = \tau(y, x)$, and that it satisfies the Sharp Triangle Inequality [see (2.20.1) of ref. 7:

$$|AB| \, \tau(\mathbf{n}_{AB}) + |BC| \, \tau(\mathbf{n}_{BC}) > |AC| \, \tau(\mathbf{n}_{AC}) \tag{2.10}$$

where, for any triangle ABC on $\mathbb{R}^2$, $|AB|$, $|BC|$, and $|AC|$ are the lengths of its sides and $\mathbf{n}_{AB}$, $\mathbf{n}_{BC}$, and $\mathbf{n}_{AC}$ are unit vectors orthogonal to the corresponding sides, the first two oriented toward the interior of the triangle, while the third one is oriented outside the triangle. This last property follows from the stiffness positivity condition:

$$\tau(\mathbf{n}) + \tau''(\mathbf{n}) > 0. \tag{2.11}$$

In fact, the infinitesimal version of the inequality (2.10) will give the last inequality (of course, with $\geqslant$ instead of $>$), so (2.10) follows from (2.11) by integration. It follows from the exact expression for the stiffness coefficient in the lhs of (2.11), obtained in ref. 1, that stiffness positivity holds for the 2D Ising model at any temperature below the critical one. The function τ_T has also the above-mentioned symmetry properties.

We shall use $\mathscr{D}$ to denote the set of all closed self-avoiding rectifiable curves $\gamma \in \mathbb{R}^2$ that are a boundary of a bounded region, $\gamma = \partial V$, $V \subset \mathbb{R}^2$. Let us recall that a curve is called rectifiable if the supremum of the lengths of polygons, with edges connecting sequentially arbitrary collections of points chosen on the curve, is finite (and equals then the length of the curve), and that a rectifiable curve has a tangent at almost every point. It is easy to verify that a curve γ that is the boundary of a convex bounded region belongs to $\mathscr{D}$. We can assign to each curve $\gamma \in \mathscr{D}$ the quantity

$$\mathscr{W}(\gamma) = \mathscr{W}_\tau(\gamma) = \int_\gamma \tau(\mathbf{n}_s)\, ds \tag{2.12}$$

where s parametrizes the curve γ according to Euclidean length measured along this curve, and $\mathbf{n}_s$ is the unit outward normal vector to the curve at the point $s \in \gamma$ (i.e., the vector orthogonal to the tangent in the considered point and oriented outward from the region bounded by γ). The functional $\mathscr{W}_\tau$ will be called the Wulff functional associated to the surface tension function $\tau(\cdot)$. Sometimes we will refer to it also as the integrated surface tension. In the isotropic case when $\tau = 1$, the corresponding Wulff functional is just the Euclidean length of the curve γ; it will be denoted by $|\gamma|$. The temperature-T Ising-model Wulff functional will be denoted by $\mathscr{W}_T$, according to our convention.

To every vector $\mathbf{n} \in \mathbb{S}^1$ and $\lambda > 0$ we assign the half-plane

$$L_{\tau, \mathbf{n}, \lambda} = \{x \in \mathbb{R}^2 \colon (x, \mathbf{n}) \leqslant \lambda \tau(\mathbf{n})\}$$

Let us consider the intersections

$$W_{\tau, \lambda} = \bigcap_{\mathbf{n} \in \mathbb{S}^1} L_{\tau, \mathbf{n}, \lambda} \tag{2.13}$$

These sets clearly satisfy the scaling relation $W_{\tau, \lambda} = \lambda W_{\tau, 1}$. In particular they keep the same shape as λ varies; this shape is called *the Wulff shape.* The parameter λ will be called the radius of the Wulff shape. *The Wulff body* is defined as $W_\tau = W_{\tau, \lambda_0}$, where the radius $\lambda_0 = \lambda_0(\tau)$ is chosen so that its volume is 1. W_τ is clearly convex and thus its boundary $\partial W_\tau \in \mathscr{D}$. The following is therefore well defined:

$$w_\tau = \mathscr{W}_\tau(\partial W_\tau)$$

For each τ, the boundary of the Wulff body satisfies the following variational principle. For all $\gamma \in \mathscr{D}$ which are boundaries of regions of volume 1,

$$w_\tau \leqslant \mathscr{W}_\tau(\gamma)$$

with equality only in case γ is a translation of ∂W_τ. The curve $\gamma_\tau = \partial W_\tau$ will be called *the Wulff curve* corresponding to the function τ.

We define the width l_τ of the Wulff curve γ_τ to be the length l of the side of the smallest square $S(l)$ into which the curve γ_τ can fit. In our case, when the function $\tau(\mathbf{n})$ is symmetric with respect to the reflections in coordinate axes, and also $\tau(x, y) = \tau(y, x)$, the width l_τ is given by

$$l_\tau = \frac{4\bar{\tau}}{\omega_\tau} \tag{2.14}$$

where the quantity $\bar{\tau}$ is given by

$$\bar{\tau} = \tau(0, 1) \tag{2.15}$$

The relation (2.14) follows from convexity and symmetry of γ_τ and from the relation (2.7.6) of ref. 7. The quantity $\bar{\tau}$ happens to be equal to one-half of the width of the Wulff shape $W_{\tau,1}$ of radius one.

3. MAIN RESULT

In this section we will formulate the main result of the present paper, Theorem 2 below, which states the existence of the threshold value $B_0(T)$ such that if the size l of our box $\Lambda(l)$ is less than B_0/h, than for small h we are in the $(-)$-phase, while if it is bigger than B_0/h, than we are predominantly in the $(+)$-phase. We start with the heuristic calculation of what the critical value $B_0(T)$ should be, since it is entering in the formulation of Theorem 2.

To do the calculation, we need to know the shape of the droplet of the $(+)$-phase which would develop in a system mixed from $(+)$- and $(-)$-phases, provided the $(+)$-phase occupies a given fraction α of the total volume, and the whole zero-field system is constrained to the square box with $(-)$-b.c. Let us scale down the whole picture in such a way that the square box would be a 1 by 1 square. As we shall learn from Lemma 1 below, the resulting droplet $\gamma_T(\alpha)$ (of volume α) has the following shape: (i) the usual Wulff shape, $\alpha^{1/2}\gamma_T$, provided it can fit the square box; (ii) *the squeezed Wulff shape* otherwise. The latter can be obtained from the square by rounding off its four corners in such a way, that each rounded part is congruent to one quarter of the (same) Wulff shape, see Fig. 1 below. Note that the resulting curve $\gamma_T(\alpha)$ is smooth, and consists of four segments and four quarters of the curve

$$\frac{(1-\alpha)^{1/2}}{l_T(1-l_T^{-2})^{1/2}}\,\gamma_T$$

The proof of that statement will be given in the next section; here we explain the heuristic behind it.

Heuristic Proof. To get an idea of how the sought curve has to look, the following picture is helpful. (It corresponds to the case of isotropic surface tension.) Consider a (one-dimensional) balloon, placed inside a box $S(1)$, which starts to be pumped. At first it is not touching the walls, and so it is not different from the unconstrained shape, which is, of course, a circle in the isotropic case. At some moment, however, it starts to touch the walls, and from that moment on a part of its surface is pressed flat against them. Since the pressure inside the balloon stays constant over its entire interior, the balloon's curvature at any point where it does not touch the wall has to be the same, being proportional to the pressure. That means that the free surface of the balloon is that of a part of a circle. Since at any moment the pumping can be stopped and the balloon would then rest, its surface has to have no corners; otherwise the corner point has to move, being subjected to the two surface tension forces which add to nonzero resultant force. Hence each curved part of our curve is congruent to one-quarter of a circle. ∎

The curves described above specify the asymptotic shape of the bubble of the $(+)$-phase over the background of the $(-)$-phase. It is well known that the probability to observe such a bubble is of the order of the exponential of the surface tension along its boundary times the inverse temperature. With such a picture in mind, the following result should not be a surprise.

Theorem 1. Suppose that $T < T_c$ and let $-1 \leqslant m_1 < m_2 \leqslant +1$; then

$$\lim_{l \to \infty} -\frac{1}{l} \log \mu_{\Lambda(l),\,-,\,T,\,0}(X_{\Lambda(l)} \in (m_1, m_2)) = \inf_{m \in (m_1, m_2)} \psi(m) \tag{3.1}$$

where

$$\psi(m) = \begin{cases} \beta \mathscr{W}_T(\gamma_T(\alpha(m))) & \text{if } \; m \in [-m_T^*, +m_T^*] \\ +\infty & \text{otherwise} \end{cases} \tag{3.2}$$

and

$$\alpha(m) = \frac{m + m_T^*}{2m_T^*}$$

The function $\alpha(m)$ above gives the relative volume of the droplet of the $(+)$-phase on the background of the $(-)$-phase, such that the resulting magnetization of the system is m.

The theorem just stated extends the results obtained recently by Ioffe.[10, 11] The difference lies in the fact that we are interested here in all values m_1, m_2, and not just in those close to $-m_T^*$. Hence we have to look for the solution of the constrained Wulff problem, with squeezed Wulff shape to enter in the statement of the theorem. Ioffe's proofs need some modifications to work in our case, and we present them in Section 7.

The function $\psi(m)$ given by (3.2) appeared in ref. 16, though there it was written less explicitly. The context of ref. 16 was somewhat different: there the question of large deviation was studied on the (more precise) level of the local limit theorem at low temperatures, while (3.1) is a statement about large deviation on the level of the integral limit theorem, valid, though, for all temperatures below the critical. The statement made in ref. 16 about the asymptotic shape of the droplet of the $(+)$-phase in the box with $(-)$ boundary condition is erroneous; the correct shapes are given by the curves $\gamma_T(\alpha)$. The low-temperature results of ref. 16 for periodic and free boundary conditions can be extended with some extra work to all temperatures below the critical, again at the price of going to the level of the integral limit theorem. They will be subject of a forthcoming publication by Cesi *et al.*[6] The proof of the Theorem 1 will be given in Section 7.

To proceed, we need to know the value of the Wulff functional $\mathscr{W}_T(\gamma_T(\alpha))$ of the shape $\gamma_T(\alpha)$ as a function of the volume α it occupies in the box $S(1)$. For $\alpha \leqslant l_T^{-2}$ we have $\gamma_T(\alpha) = \alpha^{1/2}\gamma_T$, and so

$$\mathscr{W}_T(\gamma_T(\alpha)) = \alpha^{1/2} w_T \tag{3.3}$$

For the remaining values of α a direct calculation gives the formula

$$\begin{aligned}\mathscr{W}_T(\gamma_T(\alpha)) &= 4\bar{\tau}_T - \frac{(1-\alpha)^{1/2}}{(1-l_T^{-2})^{1/2}}\,(4\bar{\tau}_T - l_T^{-1} w_T)\\ &= 4\bar{\tau}_T - (1-\alpha)^{1/2}\,(16\bar{\tau}_T^2 - w_T^2)^{1/2}\end{aligned}$$

where in the last step we used the identity (2.14). The function $\mathscr{W}_T(\gamma_T(\alpha))$ is concave for $0 \leqslant \alpha \leqslant l_T^{-2}$, is convex for $l_T^{-2} \leqslant \alpha \leqslant 1$, and is smooth everywhere on $[0,1]$ (its derivative at the inflection point $\alpha = l_T^{-2}$ is equal to $2\bar{\tau}_T$). See Fig. 1 for the graph of this function.

Accepting the above droplet picture, we can now calculate the free energy of a system of size B/h in a positive magnetic field h under the condition that it has a droplet of $(+)$-phase of relative volume α, as a function of α, and then look for its minimum. If the minimum is attained at $\alpha = 0$, then the system prefers to stay in the $(-)$-phase; otherwise there has to be a droplet of $(+)$-phase occupying the central part of our box. The

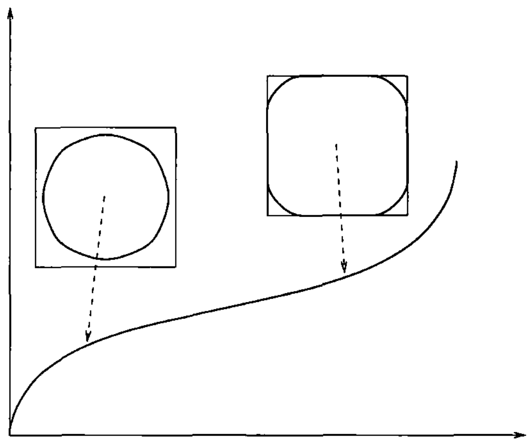

Fig. 1. Graph of $\mathscr{W}_T(\gamma_T(\alpha))$ and various shapes $\gamma_T(\alpha)$.

key observation is that if the relative volume α of $(+)$-phase—and thus the total magnetization—is fixed, then the conditional Gibbs distribution is not affected by the presence of the magnetic field, and so the droplet picture has to be the same as in the zero-field case. Without a magnetic field the free energy of the system [calculated with respect to the level of $(-)$-phase] comes only from the presence of the interface, and the minimal cost our system can pay for having an interface surrounding the relative volume α is given by

$$\beta \mathscr{W}_T(\gamma_T(\alpha)) \frac{B}{h}$$

The positive magnetic field favors the $(+)$-phase, and the energy reward the system gets after the field is turned on is given by

$$-\beta m_T^* h \left[\alpha \left(\frac{B}{h} \right)^2 \right] = -\beta m_T^* \alpha \frac{B^2}{h}$$

The above speculations tell us to look for the minima of the function (see Fig. 2)

$$g(\alpha, B) = \mathscr{W}_T(\gamma_T(\alpha))\, B - m_T^* \alpha B^2$$

For B small it has exactly one minimum at $\alpha = 0$ (which is a local minimum for all B). After the threshold $B = 2\bar{\tau}_T / m_T^*$ [at which value of B the function

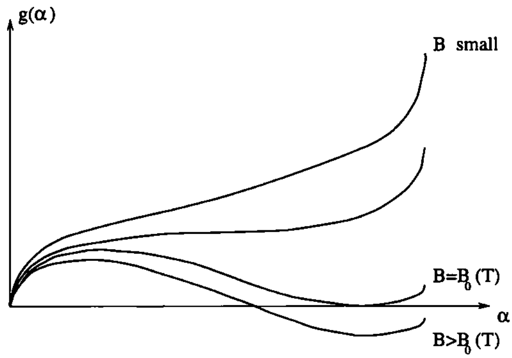

Fig. 2. Graphs of $g(\alpha, B)$ for different B.

$g(\alpha, B)$ has zero derivative at the inflection point], a second local minimum appears at the point

$$\alpha_B = 1 - \frac{16\bar{\tau}_T^2 - w_T^2}{4(Bm_T^*)^2}$$

The value $g(\alpha_B, B)$ is decreasing in B, and we are interested in the value of B for which

$$g(\alpha_B, B) = 0 \tag{3.4}$$

The solution of that equation is the critical value $B_0(T)$ we are looking for. Indeed, for smaller values of B the system prefers to be near the value $\alpha = 0$, which is then the global minimum of the function $g(\alpha, B)$, while for bigger values of B the global minimum is attained at α_B. The solution to Eq. (3.4) is given by the formula

$$B_0(T) = \frac{4\bar{\tau}_T + w_T}{2m_T^*}$$

The corresponding value of α_{B_0} is given by

$$\alpha_{B_0} = 1 - \frac{4\bar{\tau}_T - w_T}{4\bar{\tau}_T + w_T}$$

The corresponding (smallest) curve $\gamma_T(\alpha_{B_0})$, which gives the shape of the $(+)$-phase droplet, does have flat pieces along the boundary—which does not come as a surprise, since otherwise the droplet would have been able to appear at lower values of B; the length of the corresponding four straight segments of $\gamma_T(\alpha_{B_0})$ is given by

$$\frac{w_T}{4\bar{\tau}_T + w_T}$$

The reader would guess correctly that the curves $\gamma_T(\alpha)$, which have shorter length of the flat parts of the boundary are not observed in the Ising system regardless of the parameter B (and the theorem below confirms this guess). For $B > B_0$ the mean magnetization of our system is given by (see Fig. 3)

$$m_T(B) = m_T^* - \frac{16\bar{\tau}_T^2 - w_T^2}{2B^2 m_T^*} \tag{3.5}$$

This function goes asymptotically to m_T^* as $B \to \infty$. When $B \searrow B_0$, the mean magnetization goes to

$$m_T(B_0) = m_T^* \left(1 - 2\frac{4\bar{\tau}_T - w_T}{4\bar{\tau}_T + w_T}\right) \tag{3.6}$$

while for $B < B_0$ we have $m_T(B) = -m_T^*$.

One would expect the expression (3.6) to be positive, since the corresponding droplet touches the walls of the box, and therefore it occupies more than half of the box (by convexity and symmetry). Positivity boils down to the inequality

$$3w_T > 4\bar{\tau}_T \tag{3.7}$$

This inequality is indeed correct. To see it, observe, that the r.h.s. is the value of the Wulff functional calculated for the boundary of the unit

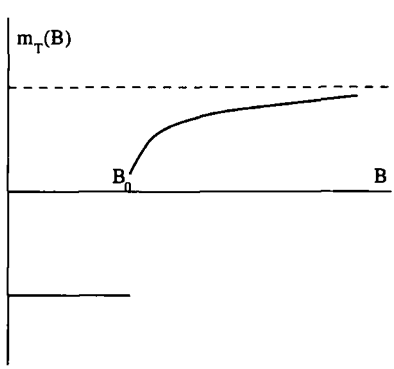

Fig. 3. Graph of $m_T(B)$.

square. The width of the Wulff body of unit volume W_T is larger than 1, hence the magnified body $3W_T$ contains the unit square, and that proves (3.7) by the triangle inequality.

Now we can formulate our main result.

Theorem 2. Suppose that $T < T_c$ is fixed and define

$$B_0 = B_0(T) = \frac{4\bar{\tau}_T + w_T}{2m_T^*}$$

Then:

a. If $B < B_0$, the following hold:

1. Given $\varepsilon > 0$, there exist $\delta > 0$ and $C < \infty$ such that for all $h > 0$

$$\mu_{\Lambda(B/h), -, T, h}(X_{\Lambda(B/h)} \in (-m_T^* - \varepsilon, -m_T^* + \varepsilon)) \geqslant 1 - Ce^{-\delta/h}$$

2. For each local function f and each function $x(\cdot)$: $(0, \infty) \to \mathbb{Z}^2$ such that $x(h) \in \Lambda(B/h)$ and

$$\text{dist}_2(x(h), \partial_{\text{ext}} \Lambda(B/h)) \to \infty \qquad \text{as} \quad h \searrow 0 \tag{3.8}$$

we have for the shifted function

$$\langle \theta_{x(h)} f \rangle_{\Lambda(B/h), -, T, h} \to \langle f \rangle_{-, T} \qquad \text{as} \quad h \searrow 0 \tag{3.9}$$

b. If $B > B_0$, the following hold:

1. Given $\varepsilon > 0$, there exist $\delta > 0$ and $C < \infty$ such that for all $h > 0$

$$\mu_{\Lambda(B/h), -, T, h}(X_{\Lambda(B/h)} \in (m_T(B) - \varepsilon, m_T(B) + \varepsilon)) \geqslant 1 - Ce^{-\delta/h}$$

where $m_T(B)$ is given by (3.5).

2. Given $\varepsilon > 0$, there exist $\delta > 0$ and $C > \infty$ such that if we denote by $\mathscr{E}_{h, \varepsilon}$ the event that inside the box $\Lambda(B/h)$ there is an external $(-)$-contour lying in the annulus between the curves $\gamma_T(\alpha_B)\, B(1-\varepsilon)/h$ and $\gamma_T(\alpha_B)\, B(1+\varepsilon)/h$ and encircling the origin, while the length of each of the remaining contours is less than ε/h, then for all $h > 0$

$$\mu_{\Lambda(B/h), -, T, h}(\mathscr{E}_{h, \varepsilon}) \leqslant 1 - Ce^{-\alpha/h}$$

3. For each local function f and function $x(\cdot)$: $(0, \infty) \to \mathbb{Z}^2$ such that the points $hx(h)/B$ lie inside $\gamma_T(\alpha_B)$ and

$$\liminf_{h \searrow 0} \text{dist}_2(hx(h)/B, \gamma_T(\alpha_B)) > 0 \tag{3.10}$$

we have for the shifted function

$$\langle \theta_{x(h)} f \rangle_{\Lambda(B/h),\, -,\, T,\, h} \to \langle f \rangle_{+,\, T} \qquad \text{as} \quad h \searrow 0 \tag{3.11}$$

4. For each local function f and function $x(\cdot)$: $(0, \infty) \to \mathbb{Z}^2$ such that the points $hx(h)/B$ lie inside $S(1)$ but outside $\gamma_T(\alpha_B)$ and

$$\liminf_{h \searrow 0} \operatorname{dist}_2(hx(h)/B,\, \gamma_T(\alpha_B) \cup \partial S(1)) > 0 \tag{3.12}$$

we have for the shifted function

$$\langle \theta_{x(h)} f \rangle_{\Lambda(B/h),\, -,\, T,\, h} \to \langle f \rangle_{-,\, T} \qquad \text{as} \quad h \searrow 0 \tag{3.13}$$

The proof of the theorem is close to the proof of Theorem 1 of ref. 17. The main difference lies in the fact that here we have to consider the constrained Wulff problem, which was avoided in ref. 17 by considering the Wulff-shaped box. So in what follows we give the necessary geometric constructions which have to replace those of ref. 17. Namely, we first solve the constrained Wulff problem, then we establish the stability properties of the solution, and finally we state the results about the constrained Wulff problem for the more general case of the families of curves with possible intersections and self-intersections. These changes plus Theorem 1 above are all essential extra ingredients needed to modify the proof of Theorem 1 of ref. 17 to work in our case.

4. CONSTRAINED VARIATIONAL PROBLEM

Let $0 < \alpha < 1$ be a real number. We want to study the Wulff variational problem of finding a closed curve $\gamma_T(\alpha)$ with a given area $V(\gamma_T(\alpha)) = \alpha$ inside, which minimizes the Wulff functional (2.12) under the additional restriction that the domain of the functional is restricted to the curves belonging to a unit square $S(1) = \{(x, y), -1/2 \leqslant |x|, |y| \leqslant 1/2\}$. We shall call this problem the *constrained Wulff problem.*

It is clear that for small values of the parameter α the corresponding curve $\gamma_T(\alpha)$ is congruent to the curve $\alpha^{1/2}\gamma_\tau$, where γ_τ is the Wulff curve which solves the unconstrained Wulff variational problem for curves of unit area. However, when α is close to one, the scaled curve $\alpha^{1/2}\gamma_\tau$ would not fit into the square $S(1)$, so we have to look for a different solution to our problem. We have

$$\gamma_\tau(\alpha) = \alpha^{1/2}\gamma_\tau \tag{4.1}$$

for $\alpha \leqslant l_\tau^{-2}$. As we shall see soon, for the remaining values of α the answer is given by the following construction:

Let $0<\rho<1$. Consider the curve $\rho l_\tau^{-1}\gamma_\tau$; it lies inside the square $S(1)$. Moreover, the distance between the curve $\rho l_\tau^{-1}\gamma_\tau$ and the boundary of $S(1)$ is $(1-\rho)/2$. Hence each of the four shifted curves

$$\rho l_\tau^{-1}\gamma_\tau+\left(\pm\frac{1-\rho}{2},\pm\frac{1-\rho}{2}\right)$$

is tangent to two corresponding sides of the square $S(1)$. Let Δ_ρ be the convex closure of the union of these four shifted curves and δ_ρ be its boundary,

$$\delta_\rho=\partial\Delta_\rho=\partial\operatorname{Conv}\left[\rho l_\tau^{-1}\gamma_\tau+\left(\pm\frac{1-\rho}{2},\pm\frac{1-\rho}{2}\right)\right] \tag{4.2}$$

It is clear that as ρ varies between 0 and 1, the area $V(\delta_\rho)$ decreases continuously from 1 down to $V(l_\tau^{-1}\gamma_\tau)$, so the definitions (4.1), (4.2) specify a family of curves with the area taking all values between 0 and 1.

Lemma 1. The curve δ_ρ is the solution to our variational problem:

$$\gamma_\tau(V(\delta_\rho))=\delta_\rho$$

That means that for any rectifiable, closed, self-avoiding curve $\gamma\subset S(1)$ with $V(\gamma)=V(\delta_\rho)$ for some ρ we have

$$\mathscr{W}(\gamma)\geqslant\mathscr{W}(\delta_\rho)$$

with equality only if $\gamma=\delta_\rho$.

Proof. For $\gamma\subset S(1)$ a rectifiable closed self-avoiding curve, let $\Delta_\gamma=\operatorname{Conv}(\gamma)$ be its convex closure. Then $V(\gamma)\leqslant|\Delta_\gamma|$, while $\mathscr{W}(\gamma)\geqslant\mathscr{W}(\partial\Delta_\gamma)$ because of the triangle inequality. Therefore, we can restrict our search for the curves $\gamma_\tau(\alpha)$ to the subset of all convex curves $\mathscr{C}$ in $S(1)$. If we include in $\mathscr{C}$ all segments and all single points of $S(1)$, then in the topology induced on it by the Hausdorff metrics, $\mathscr{C}$ is compact (Blaschke selection theorem).[4] The functional $\mathscr{W}$ is continuous on $\mathscr{C}$, which implies the existence of the minimizing curves $\gamma_\tau(\alpha)$. Hence each of the curves $\gamma_\tau(\alpha)$ is convex and contains at most four straight-line segments (which might be reduced to a single point in the limiting case), which are parts of the (different) sides of the square $S(1)$, plus a corresponding number of arcs $\kappa_1,\dots,\kappa_I$, $1\leqslant I\leqslant 4$, joining the endpoints of these segments.

(I) We will show first that these arcs should be congruent to arcs of a certain Wulff shape $b\gamma_\tau$ with corresponding well-defined value of the dilatation parameter b, the same for all arcs $\kappa_1,\dots,\kappa_I$, $I\leqslant 4$. (Here and in the following we always are talking about the congruence according to the group of all shifts only.)

To prove our statement, let us choose M to be an inner point of the arc κ_i, $1 \leqslant i \leqslant I$, let $r > 0$ be any real number, and let A, B be two points on the arc κ_i at the distance r from M. We denote by κ_{AB} the part of κ_i between A, B. Let $[A, B]$ be the corresponding chord, and K_{AB} be the corresponding segment (i.e., region bounded by the arc and its chord). We suppose that the point M is such that the arc κ_{AB} is not a straight segment for any value of r, which implies that the area $|K_{AB}|$ is positive. Such a choice of the point M is possible provided the whole arc κ_i is not a straight segment. As we shall see in the next step of the proof, this possibility can be ruled out. The choice of the value of the parameter $r = r(M)$ will be made later, in such a way that both the length of the chord $[A, B]$ and the area $|K_{AB}|$ will be small enough. This is possible since both go to 0 with r.

Now, it is easy to see that one can find a value b_i of the dilatation parameter such that the Wulff curve $b_i \gamma_\tau$ passes through two points A', B' with the following properties:

(i) The chords $[A, B]$ and $[A', B']$ are congruent (i.e., equal and parallel). Because of the convexity of the curve γ_τ, the property required holds for any Wulff shape $b\gamma_\tau$ provided only that b is not too small.

(ii) If we denote by $K_{A'B'}$ the corresponding segment of the Wulff body $b_i W_\tau$, then the areas $|K_{AB}|$ and $|K_{A'B'}|$ are the same. (It is easy to see that the possible values of the area $|K_{A'B'}|$ cover all positive real numbers as b varies.)

We claim now that the arc κ_{AB} is congruent to the arc $\kappa_{A'B'}$ of the Wulff shape $b_i \gamma_\tau$, at least when r is small enough. To see this, let us attach the arc κ_{AB} to the curve $b_i \gamma_\tau$ along the chord $[A', B']$ instead of the arc $\kappa_{A'B'}$, and, vice versa, attach the arc $\kappa_{A'B'}$ to the curve $\gamma_\tau(\alpha)$ along the chord $[A, B]$ instead of the arc κ_{AB} (Fig. 4). In this way we still have two closed self-avoiding (because of the convexity) curves. If the congruence claimed does not hold, then the modified Wulff shape has the value of the Wulff functional strictly bigger than $b\mathscr{W}(\gamma_\tau)$ (because of the uniqueness of the solution of the Wulff variational problem). That implies that the value of the Wulff functional for the curve $\gamma_\tau(\alpha)$ after modification is strictly lower than $\mathscr{W}(\gamma_\tau(\alpha))$. That might happen only if the curve $\gamma_\tau(\alpha)$, being modified, does not fit inside the square $S(1)$. However, if both the length of the chord $[A', B']$ and the area of the segment $K_{A'B'}$ of the Wulff body are small, then the diameter of the segment $K_{A'B'}$ is small as well, which rules out the last possibility for $r = r(M)$ small enough. Since the curve γ_τ is strictly

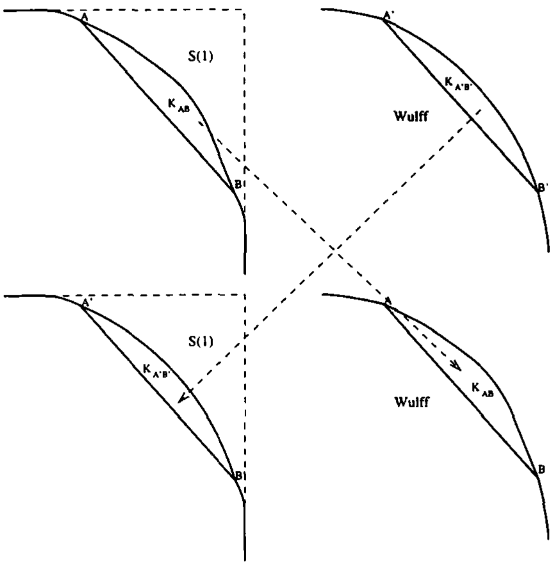

Fig. 4. The result of the surgery.

convex and smooth, the dilatations $b\gamma_\tau$ and $b'\gamma_\tau$ can not have congruent pieces for $b \neq b'$. That means that the value of the dilatation parameter b_i is well defined. Thus far we have proven that if the arc κ_i is not a straight segment, then it does not contain any straight segment (since the Wulff curve does not contain any). Hence we also know by now that the whole arc κ_i is congruent to a part of the Wulff shape $b_i\gamma_\tau$, since the arc κ_i can be covered by smaller arcs, to which the above reasoning can be applied.

What remains to be seen is that the dilatation factors b_i are the same for each of the arcs κ_i, $1 \leqslant i \leqslant I$.

To see this, let C, D be another pair of points on the arc $\kappa_{i'}$, $i' \leqslant I$, different from κ_i, which has the same properties as the pair A, B. Consider the corresponding chord $[C, D]$ and the corresponding segment K_{CD}. The arcs κ_{AB}, κ_{CD} are, of course, disjoint. It is easy to see that if these arcs are small enough, one can find a value b of the dilatation parameter such that the Wulff curve $b\gamma_\tau$ passes now through four points A', B', C', D' with the following properties:

(i) The chords $[A, B]$, $[C, D]$ and $[A', B']$, $[C', D']$ are congruent.

(ii) If we denote by $K_{A'B'}$, $K_{C'D'}$ the corresponding segments of the Wulff body bW_τ, then the two sums of the areas $|K_{AB}| + |K_{CD}|$ and $|K_{A'B'}| + |K_{C'D'}|$ are the same.

(iii) The segments $K_{A'B'}$, $K_{C'D'}$ are disjoint.

Indeed, consider the Wulff shape which corresponds to the value b equal to the smaller value of the two factors b_i and $b_{i'}$. Then (i) and (iii) are satisfied, provided only that A, B are close enough, as well as C, D. The property (ii) is violated if $b_i \neq b_{i'}$; moreover, $|K_{A'B'}| + |K_{C'D'}| > |K_{AB}| + |K_{CD}|$, since for fixed A, B the area $|K_{A'B'}|$ is a decreasing function of the parameter b. Increasing now the value of b would result in continuous decrease of the value of the sum $|K_{A'B'}| + |K_{C'D'}|$, while (i) and (iii) would remain satisfied. That proves the existence of the sought value of b.

We claim now that the arcs κ_{AB}, κ_{CD} are congruent to the corresponding arcs $\kappa_{A'B'}, \kappa_{C'D'}$ of the Wulff shape $b\gamma_\tau$. To see this, we have to attach the arcs κ_{AB}, κ_{CD} to the curve $b\gamma_\tau$ along the chords $[A', B']$, $[C', D']$ instead of the arcs $\kappa_{A'B'}, \kappa_{C'D'}$, and, vice versa, attach the arcs $\kappa_{A'B'}, \kappa_{C'D'}$ to the curve $\gamma_\tau(\alpha)$ along the chords $[A, B]$, $[C, D]$ instead of the arcs κ_{AB}, κ_{CD}. In this way we again have two closed self-avoiding curves, and the same argument based on the optimality of the Wulff curve shows the congruence claimed, which finishes the proof of (I).

(II) Our next claim is that if A, B are now the endpoints of the arc κ_1, then the tangents to κ_1 at these points coincide with the sides of the square $S(1)$. Supposing that this is not the case, and that the angle between the tangent line to κ_1 at A and the side of $S(1)$ is $\phi < \pi$, we will construct a curve $\tilde{\gamma}$ with $V(\tilde{\gamma}) = V(\gamma_\tau(\alpha)) = \alpha$, while $\mathscr{W}(\tilde{\gamma}) < \mathscr{W}(\gamma_\tau(\alpha))$; this contradiction will imply our statement. The construction of the curve $\tilde{\gamma}$ is performed in two steps: first we cut off a small portion of the curve $\gamma_\tau(\alpha)$ around the point A and replace it by a straight segment, obtaining thus the intermediate curve $\bar{\gamma}$. Since $\phi < \pi$, that would decrease the value of the Wulff functional $\mathscr{W}$ because of the Sharp Triangle Inequality. However, that also would make the area smaller, so we then enlarge the curve $\bar{\gamma}$ in a certain way, to bring its area back to the starting value of α. An easy check will then show that for the final curve $\tilde{\gamma}$ the value of the Wulff functional $\mathscr{W}(\tilde{\gamma})$ is still smaller than the initial one.

Actually we will construct the whole family of intermediate curves $\bar{\gamma}_r$, depending on the small parameter r, the value of which will be chosen later. Let C, D be two points at which a circle centered at A with radius r intersects the curve $\gamma_\tau(\alpha)$; for r small enough there are indeed exactly two points in this intersection. Suppose first that the intersection of the curve $\gamma_\tau(\alpha)$ with the square $S(1)$ contains a component which is a segment $[A, A']$, with the point A as its endpoint. Let C be the point belonging to the boundary of the square $S(1)$, while D be the one belonging to the arc κ_1. The curve $\bar{\gamma}_r$ is obtained from $\gamma_\tau(\alpha)$ by removing from it the union of the arc $\kappa_{A,D}$ and the straight-line segment $[A, C]$ and by adding the straight line segment $[C, D]$. We denote by $\bar{\kappa}_1$ the union $[C, D] \cup \kappa_{D,B}$. It follows now from the Sharp Triangle Inequality that

$$\mathscr{W}(\gamma_\tau(\alpha)) - \mathscr{W}(\bar{\gamma}_r) = C_1(\phi)\, r + O(r^2) \tag{4.3}$$

as $r \to 0$, with $C_1(\phi) > 0$ for $\phi < \pi$. Also

$$V(\gamma_\tau(\alpha)) - V(\bar{\gamma}_r) = C_2(\phi)\, r^2 + O(r^3) \tag{4.4}$$

with $C_2(\phi) > 0$.

Now we are going to enlarge the curve $\bar{\gamma}_r$ back to the area α. To do this we introduce another family of curves $\bar{\gamma}_r^s$ in the following way. We suppose additionally at this moment that the point A belongs to the top horizontal side of $S(1)$, while B belongs to the right vertical side; the remaining cases are studied in the same manner. Consider now the shift of the arc $\bar{\kappa}_1$ by a vertical vector $(0, s)$ with $s > 0$ small; let E be the intersection of the shifted arc $\bar{\kappa}_1 + (0, s)$ with the top horizontal side of $S(1)$, and $\bar{\kappa}_1^s$ be the part of this shifted curve inside $S(1)$. The curve $\bar{\gamma}_r^s$ is defined now by

$$\bar{\gamma}_r^s = (\bar{\gamma}_r \backslash \bar{\kappa}_1) \cup [B, B + (0, s)] \cup \bar{\kappa}_1^s \cup [C, E]$$

Let $C_3 = 1/2 - A_x$, where A_x is the horizontal coordinate of the point A; this positive quantity C_3 is just the distance between A and the upper-right corner of $S(1)$. Then

$$V(\bar{\gamma}_r^s) - V(\bar{\gamma}_r) = (C_3 + r)\, s + O(s^2) \tag{4.5}$$

and also

$$0 \leqslant \mathscr{W}(\bar{\gamma}_r^s) - \mathscr{W}(\bar{\gamma}_r) \leqslant C_4 s \tag{4.6}$$

where C_4 can be taken equal to $2 \max \tau$, as can be seen from the Sharp Triangle Inequality. Our choice for $\tilde{\gamma}$ is that (see Fig. 5)

$$\tilde{\gamma} = \bar{\gamma}_r^{s(r)}$$

where $s = s(r)$ solves the equation $V(\bar{\gamma}_r^s) = \alpha$. Hence by (4.4) and (4.5) the function $s = s(r)$ should satisfy the relation

$$(C_3 + r)\, s + O(s^2) = C_2(\phi)\, r^2 + O(r^3)$$

which implies that $s(r) = O(r^2)$ since $C_3 > 0$. Together with (4.3) and (4.6) this implies that $\mathscr{W}(\tilde{\gamma}) - \mathscr{W}(\gamma_\tau(\alpha)) = -C_1(\phi)\, r + O(r^2)$, with the last expression negative for small values of r, which provides the contradiction that proves our statement.

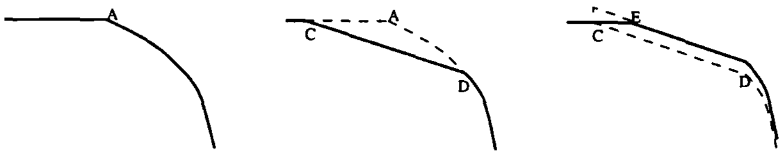

Fig. 5. $\gamma \to \bar{\gamma} \to \tilde{\gamma}$.

The case when the segment $[A, A']$ is reduced to the point A can be treated completely analogously.

(III) The same cutting and enlarging argument shows that no corner of $S(1)$ can belong to $\gamma_\tau(\alpha)$ for $\alpha < 1$.

Combining (I)–(III), we obtain that $I = 4$. That ends the proof of Lemma 1. ∎

For future use in the proof of Lemma 3 we note that the proof above can be modified in the following way: instead of the point B, we were able to use any other point B' on the arc κ_1 between A and B.

5. STABILITY PROPERTY OF THE CURVES $\gamma_\tau(\alpha)$

In this section we prove the analog of the stability property of the Wulff functional given by the formula (2.4.1) of ref. 7, for the case of the constrained variational problem treated in the preceding section. Theorem 2.4 of ref. 7 states that for any curve $\gamma \in \mathscr{D}$ surrounding a region V of unit area, $|V| = 1$, $\gamma = \partial V$, there exists a point $x = x(\gamma) \in \mathbb{R}^2$ for which

$$\rho_H(\gamma, \gamma_\tau + x) \leqslant 8 \frac{[\mathscr{W}(\gamma)^2 - \mathscr{W}(\gamma_\tau)^2]^{1/2}}{\mathscr{W}(\gamma_\tau)^2} \max_{\mathbf{n}} \tau(\mathbf{n})$$

That estimate implies that the curve γ is close to a translate of γ_τ, provided the value of the Wulff functional $\mathscr{W}(\gamma)$ is close to $\mathscr{W}(\gamma_\tau)$. We prove a similar statement for the solution of the constrained variational problem.

Lemma 2. Let $1 > \alpha > l_\tau^{-2}$ and let ε, δ be positive numbers. Let $\gamma \subset S(1)$ be a rectifiable closed self-avoiding curve with $\alpha - \varepsilon \leqslant V(\gamma) \leqslant \alpha + \varepsilon$, $\mathscr{W}[\gamma_\tau(\alpha)] - \delta \leqslant \mathscr{W}(\gamma) \leqslant \mathscr{W}[\gamma_\tau(\alpha)] + \delta$. Then

$$\rho_{\mathrm{H}}(\gamma, \gamma_\tau(\alpha)) \leqslant \mathscr{C}(\alpha, \varepsilon, \delta)$$

where $\mathscr{C}(\alpha, u, v)$ is a continuous function and $\mathscr{C}(\alpha, 0, 0) = 0$.

Note 1. In fact, Lemma 3 below implies after some straightforward calculations that

$$\mathscr{C}(\alpha, u, v) \leqslant \mathrm{const} \cdot (u + \sqrt{v}) \tag{5.1}$$

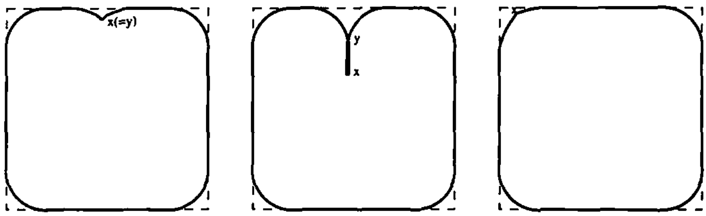

Fig. 6. Diagrams of $\gamma_\tau(\alpha, x)$ for different x.

Note 2. The same statement is of course true for the remaining values $\alpha \leqslant l_\tau^{-2}$, but then one needs to allow the curve γ to be shifted. In our case we do not need any shifts since the position of γ is fixed by the sides of the square $S(1)$.

To prove Lemma 2 we are going first to solve another variational problem, also concerning "τ-shortest" curves within $S(1)$, surrounding a given area. Namely, let α have the same meaning as above, and $x \in S(1)$ be a fixed inner point. We are looking for the curve $\gamma_\tau(\alpha, x)$ which solves the same variational Wulff problem formulated in the first paragraph of the preceding section, under the constraints that the curves over which the minimum is taken (i) belong to $S(1)$, (ii) surround the area which is equal to α, and (iii) pass through the point x.

We call the last problem the *point constrained Wulff problem.*

It turns out that for certain values of a and x the point constrained Wulff problem has no solution in the class $\mathscr{D}$, so one has to extend this space to get one. We obtain the relevant extension $\bar{\mathscr{D}}$ of $\mathscr{D}$ by relaxing the requirement of self-avoidness. Namely, the family $\bar{\mathscr{D}}$ consists of curves $\delta \subset \mathbb{R}^2$ which can be represented as unions

$$\delta = \gamma \cup \pi \tag{5.2}$$

with $\gamma \subset \mathbb{R}^2$ to be a closed self-avoiding rectifiable curve, and π to be a path, attached to γ, so that the intersection $\gamma \cap \pi$ consists of exactly one point, which is an endpoint of π. If π consists of just one point, then we recover the initial space $\mathscr{D}$. In general the points of the path π are the double points of the curve δ. Such curves are continuous—but not homeomorphic—images of the circle $\mathbb{S}^1$ (see the middle part of Fig. 6). Naturally, the set of double points should contribute twice to the value of the Wulff functional, so we define the extension of the Wulff functional to this larger space $\bar{\mathscr{D}}$ by

$$\mathscr{W}(\delta) = \mathscr{W}_\tau(\delta) = \int_\gamma \tau(\mathbf{n}_s)\, ds + 2 \int_\pi \tau(\mathbf{n}_s)\, ds \tag{5.3}$$

Lemma 3. Let $\alpha > l_\tau^{-2}$. The curve $\gamma_\tau(\alpha, x)$ has the following structure:

(i) For some point $y \in S(1)$ the curve $\gamma_\tau(\alpha, x)$ contains the segment $[x, y]$ as the set of its double points (the case $x = y$ is not excluded).

(ii) The complement $\tilde{\gamma}_\tau(\alpha, x) = \gamma_\tau(\alpha, x) \backslash [x, y)$ is a simple closed Jordan curve, smooth everywhere except possibly at y.

(iii) The curve $\tilde{\gamma}_\tau(\alpha, x)$ contains segments $I_1, ..., I_k$ of the sides of the square $S(1)$, with $k \leqslant 5$ [the side closest to y can contribute two disjoint segments to $\gamma_\tau(\alpha, x)$].

(iv) The complement $\tilde{\gamma}_\tau(\alpha, x) \backslash (\bigcup_1^k I_j \cup \{y\})$ consists of $k+1$ arcs which are congruent to arcs of the Wulff curve $\lambda\gamma_\tau$ with the same dilatation factor $\lambda = \lambda(\alpha, x)$.

(v) If $x \neq y$, then the segment $[x, y]$ is tangent to the two Wulff arcs which terminate at y.

We first give the proof of Lemma 2, using Lemma 3, and then prove Lemma 3, presenting first a heuristic proof, followed by the proper one.

Proof of Lemma 2. It is easy to see that the functional $\mathscr{W}(\gamma_\tau(\alpha, x))$ is a continuous function of the parameters α and x, which vary over a compact set. For example, the continuity in x follows from the simple estimate

$$\mathscr{W}(\gamma_\tau(\alpha, x)) - \mathscr{W}(\gamma_\tau(\alpha, z)) \leqslant 2 \max \tau \, \|x - z\|_2 \tag{5.4}$$

which one gets by connecting the point y with the curve $\gamma_\tau(\alpha, x)$ by a double segment $[x, y]$. [Note, however, that as a function of x the curve $\gamma_\tau(\alpha, x)$ is discontinuous. It has discontinuities at some x on the diagonals of $S(1)$, where the curve $\gamma_\tau(\alpha, x)$ is not unique.] Because the solution of the constrained Wulff problem is unique (contrary to the point constrained problem), the difference $\mathscr{W}(\gamma_\tau(\alpha, x)) - \mathscr{W}(\gamma_\tau(\alpha))$ stays positive, provided $x \notin \gamma_\tau(\alpha)$. Hence there exists a function $D(\alpha, \rho)$ which is continuous jointly in α and ρ and which is positive for positive ρ such that $\mathscr{W}(\gamma_\tau(\alpha, x)) - \mathscr{W}(\gamma_\tau(\alpha)) > D(\alpha, \rho)$ as soon as $\operatorname{dist}(x, \gamma_\tau(\alpha)) > \rho$. Without loss of generality we can assume that $D(\alpha, \rho)$ is monotone in ρ for each α. From the optimality of the curves $\gamma_\tau(\alpha, x)$ it follows in particular that if

$$\rho_{\mathrm{H}}(\gamma, \gamma_\tau(V(\gamma))) \geqslant \rho$$

then

$$\mathscr{W}(\gamma) - \mathscr{W}(\gamma_\tau(V(\gamma))) > D(V(\gamma), \rho)$$

Let us introduce the functions

$$d_1(\alpha, \alpha') = \rho_{\mathrm{H}}(\gamma_\tau(\alpha), \gamma_\tau(\alpha'))$$

$$d_2(\alpha, \alpha') = |\mathscr{W}(\gamma_\tau(\alpha)) - \mathscr{W}(\gamma_\tau(\alpha'))|$$

It is clear that they are continuous and that $d_1(\alpha, \alpha')$, $d_2(\alpha, \alpha') \sim |\alpha - \alpha'| \to 0$ as $\alpha' \to \alpha$.

Let now the curve γ satisfy the conditions of Lemma 2. Then

$$|\mathscr{W}(\gamma) - \mathscr{W}(\gamma_\tau(V(\gamma)))| \leqslant \delta + d_2(\alpha + \varepsilon, \alpha - \varepsilon)$$

Hence

$$\rho_{\mathrm{H}}(\gamma, \gamma_\tau(V(\gamma))) \leqslant D^{-1}(V(\gamma), \delta + d_2(\alpha + \varepsilon, \alpha - \varepsilon))$$

where $D^{-1}(\cdot, \cdot)$ is the inverse function in the second argument. Finally

$$\rho_{\mathrm{H}}(\gamma, \gamma_\tau(\alpha)) \leqslant \rho_{\mathrm{H}}(\gamma, \gamma_\tau(V(\gamma))) + d_1(\alpha - \varepsilon, \alpha + \varepsilon)$$

which proves the lemma.

Note 1. Actually, the only information from Lemma 3 which is used in the proof of Lemma 2 is the continuity of the function $\mathscr{W}(\gamma_\tau(\alpha, x))$ and the positivity of the difference $\mathscr{W}(\gamma_\tau(\alpha, x)) - \mathscr{W}(\gamma_\tau(\alpha))$ for $x \notin \gamma_\tau(\alpha)$. Both statements seem to be almost evident, and the natural approach is to try to prove Lemma 2 without using Lemma 3, by introducing instead of the function $\mathscr{W}(\gamma_\tau(\alpha, x))$ the function

$$\tilde{\mathscr{W}}(\gamma_\tau(\alpha, x)) = \inf_{\substack{\gamma \in \mathscr{D}:\, \gamma \subset S(1), \\ x \in \gamma,\, V(\gamma) = \alpha}} \mathscr{W}(\gamma)$$

The continuity of this function follows from the analog of the estimate (5.4). However, the positivity of the difference $\tilde{\mathscr{W}}(\gamma_\tau(\alpha, x)) - \mathscr{W}(\gamma_\tau(\alpha))$ is not at all obvious, due to the infinite dimensionality of the space of all curves $\mathscr{D}$. Lemma 3 allows us to pass to the finite-dimensional submanifold of $\mathscr{D}$ formed by all curves $\gamma_\tau(\alpha, x)$, and the positivity statement follows then from continuity of the function $\mathscr{W}(\gamma_\tau(\alpha, x))$ and the uniqueness of the curve $\gamma_\tau(\alpha)$.

Note 2. The explicit information about the curves $\gamma_\tau(\alpha, x)$ provided by Lemma 3, implies easily that

$$D(\alpha, \rho) \geqslant \mathrm{const} \cdot \rho^2$$

Heuristic Proof of Lemma 3. The physical picture mentioned before the proof of Lemma 1 in the previous section is helpful here as well. We can get the shapes of the curves $\gamma_\tau(\alpha, x)$ by the same pumping picture; this time, however, the box $S(1)$ should be supplied with a nail, hammered in at the point x. So from the moment when the balloon starts to touch the nail, it begins to envelope it. If the pressure is high enough, the parts of the balloon meet each other behind the nail and form a double straight layer (to the point y). However, it might happen that the pressure would be already high enough for the volume of the balloon to have the necessary value while the nail at x is still outside the balloon. Then the procedure has to be reversed: one should start from the infinite value of the pressure, when the balloon is forced to occupy the whole box $S(1)$, and then lower the pressure to the value needed to give the necessary value to the volume. Meanwhile the balloon would meet the nail and would hang on it for the lower values of the pressure (this time without the double layer).

Proof of Lemma 3. As in the proof of Lemma 1, we have to start with a statement about the existence of the minimizing curves. The idea is again to show first that there exists a compact subspace $\overline{\mathscr{C}} \subset \overline{\mathscr{D}}$ such that for every $\delta \in \overline{\mathscr{D}}$ one can find some $\bar{\delta} \in \overline{\mathscr{C}}$ such that

$$\mathscr{W}(\bar{\delta}) \leqslant \mathscr{W}(\delta), \qquad V(\bar{\delta}) \geqslant V(\delta) \tag{5.5}$$

with $\mathscr{W}$ continuous on $\overline{\mathscr{C}}$. That will imply the existence of the minimizer. This time the trick of replacing δ by $\partial\Delta_\delta$, where $\Delta_\delta = \mathrm{Conv}(\delta)$, does not work in general, since the curve $\partial\Delta_\delta$ might not pass through the point x. Nevertheless, if $x \in \partial\Delta_\delta$, then we define $\bar{\delta} = \partial\Delta_\delta$.

Suppose now that $x \notin \partial\Delta_\delta$. Then the curve $\partial\Delta_\delta$ contains a segment $[A, B]$ with $A, B \in \delta$ such that the removal of the segment $[A, B]$ from $\delta \cup \partial\Delta_\delta$ disconnects the point x from the rest of $\partial\Delta_\delta$. Let $\kappa_{AB} \subset \delta$ be the arc of δ between A and B, containing x, and $K_{AB} = \partial\Delta_\delta \setminus [A, B]$. The arc κ_{AB} can be split further into $\kappa_{AB} = \kappa_{Ax} \cup \kappa_{xB}$. The curve $\bar{\delta}$ is built now from three pieces:

$$\bar{\delta} = K_{AB} \cup K_{Ax} \cup K_{xB}, \tag{5.6}$$

where the arcs K_{Ax}, K_{xB} are convexifications of the arcs κ_{Ax}, κ_{xB}. Their construction, as well as the validity of the properties (5.5) are most easily seen from Fig. 7. Note that the intersection $K_{Ax} \cap K_{xB}$, apart from the point x, consists in general from the nonzero straight segment $[x, y]$, so in the representation (5.2) for $\bar{\delta}$ we will have in the obvious notations, that

$$\gamma = \gamma(\bar{\delta}) = K_{AB} \cup K_{Ay} \cup K_{yB} \tag{5.7},$$

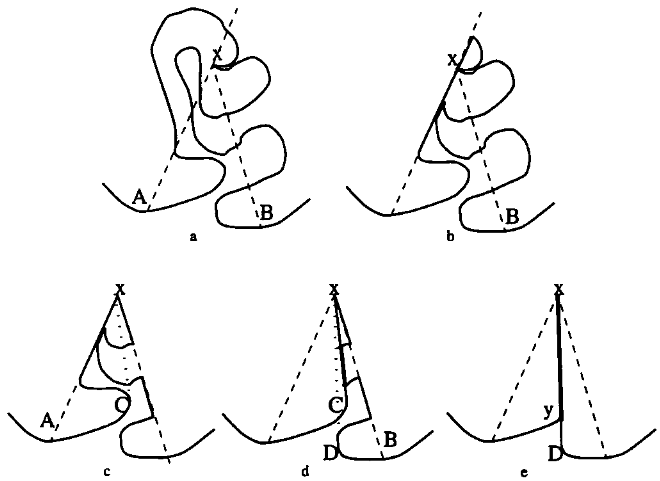

Fig. 7. Convexification of the arcs κ_{Ax}, κ_{xB}. As we proceed from (a) to (e) both the length and the Wulff functional of the curve decrease, while the area inside grows. In going from (c) to (d) we convexify the arc ACx, using the support line Cx. We then convexify the arc BDx, using the line Dx. Note, that part of the (convex) arc ACx disappears.

while $\pi = [x, y]$. In the limiting case $y = x$, and we will also use it as a convention in the case when $x \in \partial \Delta_\delta$, which was described above. We also define $\gamma(\bar{\delta}) = \bar{\delta}$ in that case.

In order to show that the resulting family $\bar{\mathscr{C}}$ of curves $\bar{\delta}$, defined by (5.6), is compact, we are going to construct an embedding E of it into another compact space as a closed subspace. This new compact space will be the product $\Pi = \mathscr{I} \times \mathscr{C} \times \mathscr{C}$, where $\mathscr{C}$ was defined in the beginning of the proof of Lemma 1 as the space of all convex curves in $S(1)$, together with their limits, which are segments and single points of $S(1)$, and $\mathscr{I}$ is the space of all segments and single points of $S(1)$. The compactness of Π follows from the Blaschke selection theorem. To specify the embedding $E: \bar{\mathscr{C}} \to \Pi$ we will specify three maps E_i from $\bar{\mathscr{C}}$ to the corresponding factors. The map E_1 is easy:

$$E_1(\bar{\delta}) = [x, y] \in \mathscr{I}$$

We construct the remaining two maps by cutting the shape surrounded by $\gamma(\bar{\delta})$ into two convex bodies. To do it in a continuous manner, let us consider all support rays to $\gamma(\bar{\delta})$ at y and take the two extremal ones, r_1 and r_2. [In the case when $\gamma(\bar{\delta})$ is not convex, we take the support rays to the convex arcs K_{Ay} and K_{yB}.] The line bisecting the angle between r_1 and

r_2 cuts the shape surrounded by $\gamma(\bar{\delta})$ into the left and right halves which are convex. That ends the proof of the compactness. The existence of the curves $\gamma_\tau(\alpha, x)$ is thus established.

Let now A, B be two points on $\gamma_\tau(\alpha, x)$ which lie on the same side of the square $S(1)$. Suppose that the portion of $\gamma_\tau(\alpha, x)$ between A and B does not pass through x. Then the triangle inequality implies immediately that the above-mentioned portion has to coincide with the segment $[A, B]$. That proves (iii).

Statement (iv) is a repetition of the first half of the proof of Lemma 1 from the previous section.

The smoothness properties (ii) and (v)—i.e., the statements that the curve $\gamma_\tau(\alpha, x)$ has no angles at its joints—are obtained in the same manner as the corresponding statement in the same lemma of the previous section.

Statement (i) that the double point portion of the curve $\gamma_\tau(\alpha, x)$ between x and y has to be a straight segment follows immediately from the triangle inequality.

6. CONSTRAINED WULFF PROBLEM FOR FAMILIES OF CURVES

In this section we consider the modification of the above problem to the case of several curves with possible intersections. Let $\bar{\gamma} = \{\gamma_1, ..., \gamma_n\}$ be a family of closed rectifiable curves in $S(1)$. We define its Wulff functional as the sum

$$\mathscr{W}(\bar{\gamma}) = \mathscr{W}(\gamma_1) + \cdots + \mathscr{W}(\gamma_n)$$

In order to define the volume of the family $\bar{\gamma}$, we suppose additionally that the curves γ_i are piecewise smooth and that the complement $\mathbb{R}^2 \setminus \cup \gamma_i$ has finitely many connected components. The relevant notion of volume in our context turns out to be that of the "phase volume" as defined in Section 2.10 of ref. 7. This definition is as follows. The set $\mathbb{R}^2 \setminus \cup \gamma_i$ splits up into a collection of connected components Q_α with exactly one unbounded component among them. A component Q_α will be called a minus component if any path that connects its interior points with points of the unbounded component and intersects the curves from $\bar{\gamma}$ in a finite number of points of smoothness intersects them in an odd number of points (counted with multiplicities). The phase volume of $\bar{\gamma}$, denoted by $\hat{V}(\bar{\gamma})$, is defined as the joint volume of all the minus components. We are now in a position to state the analog of Lemma 1 for the case of the collection of curves.

Lemma 4. Let $1 > \alpha > l_\tau^{-2}$, and let ε, δ be positive numbers. Let $\bar{\gamma}$ be a collection of rectifiable closed curves in $S(1)$ with $\alpha - \varepsilon \leqslant \hat{V}(\bar{\gamma}) \leqslant \alpha + \varepsilon$, $\mathscr{W}[\gamma_\tau(\alpha)] - \delta \leqslant \mathscr{W}(\bar{\gamma}) \leqslant \mathscr{W}[\gamma_\tau(\alpha)] + \delta$. Then there exists an index j, $1 \leqslant j \leqslant n$, such that for the corresponding curve γ_j

$$\rho_{\mathrm{H}}(\gamma_j, \gamma_\tau(\alpha)) \leqslant \mathscr{C}(\alpha, \varepsilon, \delta)$$

while for the remaining curves we have

$$\sum_{i \neq j} |\gamma_i| \leqslant \mathscr{C}'(\alpha, \varepsilon, \delta)$$

Here the function $\mathscr{C}$ satisfies the same bound (5.1) of the order of $\varepsilon + \sqrt{\delta}$, while the function $\mathscr{C}'$ satisfies

$$\mathscr{C}'(a, u, v) \leqslant \mathrm{const}_\tau \cdot (u + v)$$

The proof of this statement is a repetition of the proofs of Theorems 2.9 and 2.10 from ref. 7, and will be omitted.

7. VARIATIONS ON THE RESULTS OF IOFFE

The proof of Theorem 1 in ref. 17 as well as the main result of the present paper rely in part on results obtained and ideas introduced by Ioffe.[10, 11] Theorem 1 formulated above does not follow formally from results of refs. 10 and 11, though by using Ioffe's techniques one can obtain it with some extra work. We prefer an alternative approach, which is in part just an easier way to implement Ioffe's strategy and in part a simpler alternative technique. The former part refers to the upper bound, while the latter one refers to the lower bound, which together give the proof of Theorem 1.

The proof of the upper bound in ref. 11 is based on (i), the extension of Pfister's estimate of the probability of observing a large contour with a given skeleton (see Lemma 10.1 of ref. 13) to all temperatures below the critical, and (ii) the study of the conditional Gibbs measures conditioned by the event of absence of large contours.

An easier way to estimate the probabilities of large deviations under these conditional Gibbs measures is presented in the Appendix to ref. 17. Together with our Lemma 1, this gives the upper bound.

In what follows we present our alternative approach to the lower bound. First, we introduce some extra notations.

Given a bounded set $U \subset \mathbb{R}^2$ and $l > 0$, we will use the notation

$$lU = \{lx : x \in U\}$$

and

$$\Lambda(U) = \mathbb{Z}^2 \cap U$$

We will be considering the Gibbs measure for the system in the box $\Lambda(lU)$ with $(-)$ boundary conditions, where l is thought to be large. The set U, which gives the shape of the box, can be quite general; we will suppose that it is normalized so that its volume is 1 and that its boundary is a rectifiable curve. Under these conditions,

$$|\Lambda(lU)| = l^2 + O(l)$$

Theorem 3. Suppose that $T < T_c$. Then for every convex $V \subset U$ we have

$$\limsup_{l \to \infty} -\frac{1}{l} \log \mu_{\Lambda(lU),\, -,\, T,\, 0}(X_{\Lambda(lU)} \geqslant m) \leqslant \beta \mathscr{W}_T(\partial V)$$

for all $m < |V| m_T^* + (1 - |V|)(-m_T^*) = (2\,|V| - 1)\, m_T^*$.

The lower bound part of Theorem 1 follows from Theorem 3 applied to the case when $U = S(1)$, $V = \operatorname{Int} \gamma_T(\alpha)$, $0 < \alpha < 1$.

Theorem 3 follows at once from the following lemma.

Lemma 5. Suppose that $T < T_c$. Then for every $V \subset U$ whose boundary is a polygonal line and which satisfies

$$\operatorname{dist}_2(V, U^c) > 0$$

we have

$$\limsup_{l \to \infty} -\frac{1}{l} \log \mu_{\Lambda(lU),\, -,\, T,\, 0}(X_{\Lambda(lU)} \geqslant m) \leqslant \beta \mathscr{W}_T(\partial V)$$

where $m = |V|\, m_T^* + (1 - |V|)(-m_T^*) = (2\,|V| - 1)\, m_T^*$.

Proof. We will denote the ordered vertices of V by $x_1, \ldots, x_n$ and use the convention that $x_{n+1} = x_1$. If we denote by $\mathbf{n}_i$ the unit vector perpendicular to the edge which joins the vertices x_i and x_{i+1}, oriented outward from V, say, then

$$\mathscr{W}_T(\partial V) = \sum_{i=1}^{n} \tau_T(\mathbf{n}_i)\, \|x_{i+1} - x_i\|_2$$

Given $\varepsilon>0$ small enough [$2\varepsilon<\mathrm{dist}_2(V, U^c)$ is what we will need below], there exists $\delta>0$ so small that we have

$$\begin{aligned}
&\mu_{\Lambda(lU),-,T,0}(X_{\Lambda(lU)}\geqslant m)\\
&\qquad\geqslant\mu_{\Lambda(lU),-,T,0}(X_{\Lambda((1+\varepsilon)lV)}\geqslant m_T^*(1-\varepsilon))\\
&\qquad\quad-\mu_{\Lambda(lU),-,T,0}(X_{\Lambda(lU)\backslash\Lambda((1+\varepsilon)lV)}<-m_T^*-\delta)
\end{aligned}\tag{7.1}$$

For the second term in the right-hand side we can simply use inequality (4.8) in ref. 11 [which can be applied since for small ε we have $|\Lambda(lU)\backslash\Lambda((1+\varepsilon)lV)|/l^2>\delta'>0$, for all l]. This inequality gives

$$\mu_{\Lambda(lU),-,T,0}(X_{\Lambda(lU)\backslash\Lambda((1+\varepsilon)lV)}<-m_T^*-\delta)\leqslant C_1e^{-C_2l^2}$$

Our task is therefore reduced to proving that the first term in the right-hand side of (7.1) satisfies

$$\limsup_{l\to\infty}-\frac{1}{l}\log\mu_{\Lambda(lU),-,T,0}(X_{\Lambda((1+\varepsilon)lV)}\geqslant m_T^*(1-\varepsilon))\leqslant\mathscr{W}_T(\partial V)\tag{7.2}$$

For this purpose we will use the Fortuin–Kasteleyn random cluster model (see ref. 3 for a detailed, mathematically rigorous presentation and references to the origins of the model) and the coupling between this model and the Ising model Gibbs measures, as introduced in ref. 9. We start by reviewing the basic definitions and facts, and refer the reader to the two papers just quoted for the complete proofs of the claims. We will be using definitions (2.2)–(2.5).

The Fortuin–Kasteleyn random cluster models assign probability measures to configurations of occupied and vacant bonds. Two different measures will be of relevance to us: the FK measures with free and with wired boundary conditions. Given $\Gamma\subset\subset\mathbb{Z}^2$, one introduces the FK measure with free boundary condition on $\{0,1\}^{\mathbb{B}_\Gamma}$, where the 1's are associated with occupied bonds and the 0's with vacant bonds. One proceeds as follows. We say that two sites x and y are connected in the configuration $\eta\in\{0,1\}^{\mathbb{B}_\Gamma}$, if there is a chain of sites $x=x_1, x_2,\ldots, x_{n-1}$, $x_n=y$ in Γ so that $\{x_i, x_{i+1}\}$ is occupied in the configuration η for $i=1,\ldots,n-1$. We use the notation $A\leftrightarrow B$ to denote the fact that some site in the set A is connected to some site in the set B, and if A or B is a singleton, then we replace the set by its single element in this notation. Clusters are maximal connected sets of sites. To each configuration $\eta\in\{0,1\}^{\mathbb{B}_\Gamma}$ we associate three numbers: $a(\eta)$ is the number of occupied bonds in η, $b(\eta)$ is the number of vacant bonds in η, and $c_{\text{free}}(\eta)$ is the number of distinct clusters of sites in η. For each $p \in [0, 1]$, the probability measure $\mathscr{P}_{\Lambda, \text{free}, p}$ is then introduced as

$$\mathscr{P}_{\Lambda, \text{free}, p}(\eta) \propto p^{a(\eta)}(1-p)^{b(\eta)}\, 2^{c_{\text{free}}(\eta)} \tag{7.3}$$

The FK measure with wired boundary condition on the set $\Gamma \subset\subset \mathbb{Z}^2$ is defined on the set $\{0, 1\}^{\mathbb{B}_\Gamma}$. One considers again the clusters into which each configuration partitions Γ, but special attention is given to the sites which are connected to sites of $\partial_{\text{ext}}\Gamma$. The clusters that contain sites in $\partial_{\text{ext}}\Gamma$ are called boundary clusters and the other ones are called inner clusters. $c_{\text{wired}}(\eta)$ is the number of distinct inner clusters of sites in η. For each $p \in [0, 1]$, the probability measure $\mathscr{P}_{\Gamma, \text{wired}, p}$ is then introduced as

$$\mathscr{P}_{\Gamma, \text{wired}, p}(\eta) \propto p^{a(\eta)}(1-p)^{b(\eta)}\, 2^{c_{\text{wired}}(\eta)}$$

The FK probability measure with wired boundary condition on Γ is related to the Gibbs measures $\mu_{\Gamma, -, T, 0}$ and $\mu_{\Gamma, +, T, 0}$, with p and T being related by

$$p = 1 - e^{-\beta}$$

On the other hand, the FK probability measure with free boundary condition on Γ is related to the Gibbs measure $\mu_{\Gamma, 0, T, 0}$, which corresponds to taking as boundary condition outside of Γ spins which take the value 0, or equivalently, supposing that the lattice is just the finite set Γ and that there are no sites outside this set. Two basic relations are:

1. For all $x \in \Gamma$

$$\mathscr{P}_{\Gamma, \text{wired}, p}(x \leftrightarrow \partial_{\text{ext}}\Gamma) = \langle \sigma(x) \rangle_{\Gamma, +, T, 0} \tag{7.4}$$

2. For all $x, y \in \Gamma$

$$\mathscr{P}_{\Gamma, a, p}(x \leftrightarrow y) = \langle \sigma(x)\, \sigma(y) \rangle_{\Gamma, b, T, 0} \tag{7.5}$$

where either a is wired and b is $(-)$ or $(+)$, or else a is free and b is 0.

Further relations between the FK measures and the corresponding Gibbs measures can be obtained via the coupling introduced in ref. 9, which we describe next. First we consider $\mathscr{P}_{\Gamma, \text{free}, p}(\eta)$, and for each cluster of sites we choose either the value $+1$ or the value -1, independent of the previous random choice of the configuration and independently from cluster to cluster. If to all the sites in each cluster we assign spins which take the common value chosen for this cluster in this procedure, we obtain a random configuration on $\{-1, +1\}^{\Gamma}$. The law of this configuration happens to be exactly $\mu_{\Gamma,0,T,0}$. We will use $\mathbb{P}_{\Gamma,0,T}$ to denote this coupled probability measure on $\{0,1\}^{\mathbb{B}_\Gamma}\times\{-1,+1\}^{\Gamma}$. Similarly, if we start with $\mathscr{P}_{\Gamma,\text{wired},p}(\eta)$ and assign values for the spins at sites which belong to inner clusters in precisely the same way as above, but set all the spins at sites which belong to boundary clusters as -1 (resp., $+1$), then we obtain a random configuration on $\{-1,+1\}^{\Gamma}$ with law $\mu_{\Gamma,-,T,0}$ (resp., $\mu_{\Gamma,+,T,0}$). We will use $\mathbb{P}_{\Gamma,-,T}$ (resp. $\mathbb{P}_{\Gamma,+,T}$) to denote this coupled probability measure on $\{0,1\}^{\mathbb{B}_\Gamma}\times\{-1,+1\}^{\Gamma}$.

Finally, we review the duality relations used in ref. 5. The dual lattice is $\mathbb{Z}^2_*=\mathbb{Z}^2+(1/2,1/2)$, and the set of dual bonds is

$$\mathbb{B}^*=\{\{x,y\}: x,y\in\mathbb{Z}^2_* \text{ and } \|x-y\|_1=1\}$$

There is a natural one-to-one mapping between $\mathbb{B}$ and $\mathbb{B}^*$, given by $\{x,y\}^*=\{u,v\}$ in case the straight-line segments joining x to y and joining u to v intersect each other in their middle point. We will need to consider a slight generalization of the measures $\mathscr{P}_{\Gamma,\text{free},p}$; given $\mathscr{E}\subset\mathbb{B}_\Gamma$, we define the probability measure $\mathscr{P}_{(\Gamma,\mathscr{E}),\text{free},p}$ on $\{0,1\}^{\mathscr{E}}$ by the same formula (7.3); in particular, $\mathscr{P}_{(\Gamma,\mathbb{B}_\Gamma),\text{free},p}=\mathscr{P}_{\Gamma,\text{free},p}$. Given $\Gamma\subset\subset\mathbb{Z}^2$, we define

$$\Gamma^*=\{u\in\mathbb{Z}^2_*: \{u,v\}=\{x,y\}^* \text{ for some } v\in\mathbb{Z}^2_* \text{ and } \{x,y\}\in\bar{\mathbb{B}}_\Gamma\}$$

In other words, Γ^* is the smallest subset Γ' of $\mathbb{Z}^2_*$ such that $(\bar{\mathbb{B}}_\Gamma)^*\subset\mathbb{B}_{\Gamma'}$, but observe that these two sets do not need to be identical. Each configuration $\eta\in\{0,1\}^{\mathbb{B}_\Gamma}$ induces a configuration

$$\eta^*\in\{0,1\}^{\mathbb{B}_{\Gamma^*}}$$

by declaring the dual bonds to be occupied or vacant when the corresponding bond in $\bar{\mathbb{B}}_\Gamma$ is, respectively, vacant or occupied. In this fashion the measure $\mathscr{P}_{\Gamma,\text{wired},p}$ induces a probability measure on the set of configurations on the dual set of bonds. This measure turns out to be precisely $\mathscr{P}_{(\Gamma^*,(\mathbb{B}_\Gamma)^*),\text{free},p^*}$, where

$$p^*=\frac{2-2p}{2-p}$$

If we write $p=1-e^{-\beta}$ and $p^*=1-e^{-\beta^*}$, then this relation is equivalent to the Krammer–Wannier duality relation

$$e^{-\beta^*}=\tanh\left(\frac{\beta}{2}\right)$$

which defines $\beta^*=\beta^*(\beta)$ as a strictly decreasing function with the unique fixed point $\beta_c=1/T_c$.

Now we are ready to go back to the proof of (7.2). We introduce the event

$$\mathscr{G} = \{\partial_{\text{int}} \Lambda((1+\varepsilon)\, lV) \leftrightarrow \partial_{\text{ext}} \Lambda((1+2\varepsilon)\, lV)\}^c$$

Clearly

$$\begin{aligned}\mu_{\Lambda(lU),\,-,\,T,\,0}&(X_{\Lambda((1+\varepsilon)\,lV)} \geqslant m_T^*(1-\varepsilon))\\ &\geqslant \mathbb{P}_{\Lambda(lU),\,-,\,T}(X_{\Lambda((1+\varepsilon)\,lV)} \geqslant m_T^*(1-\varepsilon) \mid \mathscr{G})\, \mathbb{P}_{\Lambda(lU),\,-,\,T}(\mathscr{G})\end{aligned}$$

Therefore we will be finished once we show that

$$\mathbb{P}_{\Lambda(lU),\,-,\,T}(X_{\Lambda((1+\varepsilon)\,lV)} \geqslant m_T^*(1-\varepsilon) \mid \mathscr{G}) \geqslant 1/3 \tag{7.6}$$

for large l, and that

$$\limsup_{l\to\infty} -\frac{1}{l} \log \mathbb{P}_{\Lambda(lU),\,-,\,T}(\mathscr{G}) \leqslant \mathscr{W}_T(\partial V) \tag{7.7}$$

To show (7.6), we let

$$\mathscr{C} = \{x \in \Lambda((1+2\varepsilon)\, lV) \colon x \leftrightarrow \partial_{\text{ext}} \Lambda((1+2\varepsilon)\, lV)\}$$

and we let $\{\mathscr{G}_\alpha\}$ be the partition of $\mathscr{G}$ according to what $\mathscr{C}$ is. For each α let

$$\Gamma(\alpha) = \Lambda((1+2\varepsilon)\, lV) \backslash \mathscr{C}_\alpha$$

where $\mathscr{C}_\alpha$ is the set $\mathscr{C}$ for configurations in $\mathscr{G}_\alpha$. Clearly, for each α

$$\Lambda((1+\varepsilon)\, lV) \subset \Gamma(\alpha) \tag{7.8}$$

For this reason we obtain, from the way the coupling $\mathbb{P}_{\Lambda(lU),\,-,\,T}$ is constructed, that for each α,

$$\begin{aligned}\mathbb{P}_{\Lambda(lU),\,-,\,T}&(X_{\Lambda((1+\varepsilon)\,lV)} \geqslant m_T^*(1-\varepsilon) \mid \mathscr{G}_\alpha)\\ &= \mu_{\Gamma(\alpha),\,0,\,T,\,0}(X_{\Lambda((1+\varepsilon)\,lV)} \geqslant m_T^*(1-\varepsilon))\end{aligned} \tag{7.9}$$

The same argument used to prove Theorem 2 in ref. 5 shows that, because (7.8) is satisfied, we have

$$\mu_{\Gamma(\alpha),\,0,\,T,\,0}(X_{\Lambda((1+\varepsilon)\,lV)} \geqslant m_T^*(1-\varepsilon)) \to \tfrac{1}{2} \tag{7.10}$$

uniformly over α, as $l \to \infty$. The estimate (7.6) follows from (7.9) and (7.10).

Now we turn to the proof of (7.7). First observe that the event $\mathscr{G}$ can be seen as stating the presence of an occupied dual circuit inside $[\Lambda((1+2\varepsilon)\, lV)\backslash\Lambda((1+\varepsilon)\, lV)]^*$ which surrounds $\Lambda((1+\varepsilon)\, lV)$. We will estimate from below the probability of the presence of such an occupied dual circuit, by constructing it from pieces which lie close to the corresponding edges of $(1+\varepsilon)\, lV$. For this purpose define V_i as the quadrilateral with vertices $(1+\varepsilon)\, x_i$, $(1+\varepsilon)\, x_{i+1}$, $(1+2\varepsilon)\, x_i$, and $(1+2\varepsilon)\, x_{i+1}$. Note that we can take n sites $x_1(l), ..., x_n(l)$ in $\mathbb{Z}^2_*$ with the properties that:

1. The Euclidean distance between each $x_i(l)$ and the point $(1+(3/2)\,\varepsilon)\, lx_i \in \mathbb{R}^2$ is bounded above by 1.
2. $x_i \in [\Lambda(lV_{i-1})]^* \cap [\Lambda(lV_i)]^*$.

Let $\mathscr{G}_i$ be the event that there is a path of occupied dual bonds which lies in $[\Lambda(lV_i)]^*$ and connects $x_i(l)$ to $x_{i+1}(l)$. Using the FKG inequalities for the random cluster model, we obtain now

$$\begin{aligned}\mathbb{P}_{\Lambda(lU),\,-,\,T}(\mathscr{G}) &\geqslant \mathscr{P}_{([\Lambda(lU)]^*,\,[\mathbb{B}_{\Lambda(lU)}]^*),\,\text{free},\,p^*}\left(\bigcap_{i=1}^{n} \mathscr{G}_i\right)\\ &\geqslant \prod_{i=1}^{n} \mathscr{P}_{([\Lambda(lU)]^*,\,[\mathbb{B}_{\Lambda(lU)}]^*),\,\text{free},\,p^*}(\mathscr{G}_i)\\ &\geqslant \prod_{i=1}^{n} \mathscr{P}_{[\Lambda(lV_i)]^*,\,\text{free},\,p^*}(\mathscr{G}_i)\end{aligned}$$

To estimate each term in the right-hand side we use (7.5) and write

$$\mathscr{P}_{[\Lambda(lV_i)]^*,\,\text{free},\,p^*}(\mathscr{G}_i) = \langle \sigma(x_i(l))\, \sigma(x_{i+1}(l))\rangle_{[\Lambda(lV_i)]^*,\,0,\,T^*,\,0}$$

The same method used to prove Lemma 6.3 in ref. 13 gives the following result, which is a simple generalization of that well-known lemma:

$$\lim_{l\to\infty} -\frac{1}{\|x_i(l)-x_{i+1}(l)\|_2} \times \log\langle \sigma(x_i(l))\, \sigma(x_{i+1}(l))\rangle_{[\Lambda(lV_i)]^*,\,0,\,T^*,\,0} = \beta\tau_T(\mathbf{n}_i)$$

Combining the last three displayed statements, we have

$$\limsup_{l\to\infty} -\frac{1}{l}\log \mathbb{P}(\mathscr{G}) \leqslant (1+(3/2)\,\varepsilon)\,\beta\mathscr{W}_T(\partial V)$$

from which (7.7) follows from the arbitrariness of ε. This completes the proof of Lemma 5.

8. THE PROOF OF THE MAIN RESULT

As mentioned above, the proof of Theorem 2 goes along the same lines as the proof of the Theorem 1 of ref. 17. The only statement that requires some additional arguments is the statement b.4 about the appearance of the $(-)$-phase in the vicinity of the corners of the box $\Lambda(B/h)$ for $B > B_0$, which statement was avoided in ref. 17 by considering the Wulff-shaped boxes $\Delta(B/h)$ instead of the square boxes $\Lambda(B/h)$. The second difference—the statement in part b.2 about the shortness of all other contours except the one living in the annulus between the curves $\gamma_T(\alpha_B)\, B(1-\varepsilon)/h$ and $\gamma_T(\alpha_B)\, B(1+\varepsilon)/h$—requires no extra work. Indeed, the stability result of ref. 7, used in ref. 17 to derive part b.2 of Theorem 1 of that paper, deals exactly with the situation when a family of curves surrounds a total phase volume larger than or equal to 1, and whose Wulff functionals add up to an amount $\beta w_T + \kappa$ that is not much larger than the minimum possible, βw_T. The stability statement claims then that there must be a curve in this family which, modulus a translation, is close in Hausdorff distance to a Wulff curve which surrounds a volume 1, while the total length of the remaining curves is of the order of κ. In ref. 17 only the first half of the statement was used, which implied the statement about the largest contour; the second half gives the statement about the contours remaining.

In the proof below of part b.4 we will use statements a.2 and b.2 of our Theorem 2. The idea of the proof is first to show that the support $S_{f,x}$ of the shifted function $\theta_{x(h)} f$ with overwhelming probability is situated outside the large contour of part b.2, which surrounds the central part of our box. We then show that with the same high probability this support $S_{f,x}$ is encircled by a $(*)$-circuit of $(-)$-spins of relatively small size, and that allows us to apply part a.2 of the theorem to prove the convergence to the $(-)$-phase.

To implement the program we start by introducing explicitly the value d of the lim inf from the relation (3.12):

$$d = \liminf_{h \searrow 0} \operatorname{dist}_2(hx(h)/B, \gamma_T(\alpha_B) \cup \partial S(1))$$

In our use of part b.2 of Theorem 2 the value of the parameter ε will be supposed to satisfy

$$\varepsilon < \frac{d}{3}$$

Let us introduce the set $\mathscr{F}$ of configurations σ in $\Lambda(B/h)$ for which the size of any contour Γ of σ, such that $\operatorname{Int}(\Gamma) \cap S_{f,x} \neq \varnothing$, does not exceed ε/h. By part b.2 we can state that

$$\mu_{\Lambda(B/h),\, -,\, T,\, h}(\mathscr{F}) > 1 - Ce^{-\delta/h}$$

for some C and δ. Hence for any configuration $\sigma \in \mathscr{F}$ there exists a (*)-circuit of $(-)$-spins which surrounds the set $S_{f,x}$ and lies inside the box $\Lambda(2\varepsilon/h)+x(h)$, provided h is small enough. Let us partition $\mathscr{F}$ according to the outermost such circuit lying in $\Lambda(2\varepsilon/h)+x(h)$, and use the notation $\{\mathscr{F}_\alpha\}$ to denote this partition. Using a self-explanatory notation for conditional expectations, from the Markov property and the FKG-Holley inequalities, we obtain for each α

$$\begin{aligned}\langle \theta_{x(h)} f \mid \mathscr{F}_\alpha \rangle_{\Lambda(B/h),\,-,\,T,\,h} &\leqslant \langle \theta_{x(h)} f \rangle_{\Lambda(2\varepsilon/h)+x(h),\,-,\,T,\,h} \\ &= \langle f \rangle_{\Lambda(2\varepsilon/h),\,-,\,T,\,h} \qquad (8.1)\end{aligned}$$

Now we can invoke part a.2 of our theorem, which tells us that the r.h.s. of (8.1) goes to $\langle f \rangle_{-,T}$ as $h \searrow 0$, provided $2\varepsilon < B_0$. That provides us with the upper bound for the limiting value of the expectation $\langle \theta_{x(h)} f \rangle_{\Lambda(B/h),\,-,\,T,\,h}$. The complementary inequality

$$\langle \theta_{x(h)} f \rangle_{\Lambda(B/h),\,-,\,T,\,h} \geqslant \langle f \rangle_{-,T}$$

follows immediately from the FKG-Holley inequalities.

ACKNOWLEDGEMENTS

We are grateful to a referee for pointing out to us the Blaschke selection theorem as a tool to obtain the compactness argument.

The work of R.H.S was partially supported by the NSF through grants DMS 9100725 and DMS 9400644 and that of S.B.S. by the NSF through grant DMS 9208029 and by the Russian Fund for Fundamental Research through grant 930101470.

REFERENCES

1. N. Akutsu and Y. Akutsu, Relationship between the anisotropic interface tension, the scaled interface width and the equilibrium shape in two dimensions, *J. Phys. A: Math. Gen.* **19**:2813–2820 (1986).
2. K. Alexander, J. T. Chayes, and L. Chayes, The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation, *Commun. Math. Phys.* **131**:1–50 (1990).
3. M. Aizenman, J. T. Chayes, L. Chayes, and C. N. Newman, Discontinuity of the magnetization in one-dimensional $1/|x-u|^2$ Ising and Potts models, *J. Stat. Phys.* **50**:1–40 (1988).
4. T. Bonnesen and W. Fenchel, *Theory of Convex Bodies* (BCS Associates, 1987).
5. J. Chayes, L. Chayes and R. H. Schonmann, Exponential decay of connectivities in the two-dimensional Ising model, *J. Stat. Phys.* **49**:433–445 (1987).
6. F. Cesi, G. Guadagni, F. Martinelli, and R. H. Schonmann, On the 2D dynamical Ising model in the phase coexistence region near the critical point, *J. Stat. Phys.*, to appear.

7. R. L. Dobrushin, R. Kotecký, and S. Shlosman, *Wulff Construction: A Global Shape from Local Interaction* (AMS Translations Series, 1992).
8. R. L. Dobrushin and S. Shlosman, Thermodynamic inequalities for the surface tension and the geometry of the Wulff construction, in *Ideas and Methods in Mathematical Analysis, Stochastics and Applications*, S. Albeverio, ed. (Cambridge University Press, 1992), Vol. 2, pp. 371–403.
9. R. G. Edwards and A. D. Sokal, Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm, *Phys. Rev. D* **38**:2009–2012 (1988).
10. D. Ioffe, Large deviations for the 2D Ising model: A lower bound without cluster expansions, *J. Stat. Phys.* **74**:411–432 (1994).
11. D. Ioffe, Exact large deviation bounds up to T_c for the Ising model in two dimensions, *Prob. Th. Rel. Fields* **102**:313–330 (1995).
12. D. G. Martirosyan, Theorems on strips in the classical Ising ferromagnetic model, *Sov. J. Contemp. Math. Anal.* **22**:59–83 (1987).
13. C. E. Pfister, Large deviations and phase separation in the two-dimensional Ising model, *Helv. Phys. Acta* **64**:953–1054 (1991).
14. A. Pisztora, Surface order large deviations of the Ising model below T_c, *Prob. Th. Rel. Fields*, to appear.
15. R. H. Schonmann, Slow droplet-driven relaxation of stochastic Ising models in the vicinity of the phase coexistence region, *Commun. Math. Phys.* **161**:1–49 (1994).
16. S. Shlosman, The droplet in the tube: A case of phase transition in the canonical ensemble, *Commun. Math. Phys.* **125**:81–90 (1989).
17. R. H. Schonmann and S. Shlosman, Complete analyticity for 2D Ising completed, *Commun. Math. Phys.* **170**:453–482 (1995).