Phase Diagram of the Two-Dimensional Ising Antiferromagnet (Computer-Assisted Proof)

R. L. Dobrushin[1], J. Kolafa[2], and S. B. Shlosman[1]

1 Institute of Information Transmission Problems, Academy of Sciences, Moscow, USSR
2 Department of Mathematical Physics, Charles University, V Holešovičkách 2,
CS-18000 Prague 8, Czechoslovakia

Abstract. We study the phase diagram of the Ising antiferromagnet on a square lattice in a neighbourhood of ground state critical points $h = \pm 4$, $T = 0$. It leads to a question about the value a_C of the critical activity of the hard-square lattice gas. Using a constructive criterion of uniqueness we prove that $a_C > 1$ and that the phase diagram of the antiferromagnet does not bulge near mentioned critical points. It is a specific feature of this work that the proof was completed with the help of a computer.

1. Introduction

Let us consider the Ising antiferromagnet on a square lattice with spins $\sigma_t = \pm 1$ attached to lattice sites $t \in \mathbb{Z}^2$ and with the Hamiltonian

$$H = \sum_{\langle st \rangle} \sigma_s \sigma_t - h \sum_t \sigma_t, \tag{1.1}$$

where the first sum is over all nearest neighbours $s, t \in \mathbb{Z}^2$, $|s - t| = 1$.

For $|h| < 2d = 4$, there are two different ground configurations $\sigma^{(1)}, \sigma^{(2)}$ in this model:

$$\sigma_t^{(1)} = \begin{cases} +1 & \text{whenever} \quad t_1 + t_2 \quad \text{is even}, \\ -1 & \text{whenever} \quad t_1 + t_2 \quad \text{is odd}, \end{cases}$$

$$\sigma_t^{(2)} = -\sigma_t^{(1)}.$$

These configurations are stable, i.e. they generate extremal Gibbs states $\langle \ \rangle_\beta^{(1)}, \langle \ \rangle_\beta^{(2)}$, corresponding to the Hamiltonian (1.1) and to an inverse temperature $\beta = T^{-1}$, which are for large β only small perturbations of corresponding non-random states concentrated on the configurations $\sigma^{(1)}, \sigma^{(2)}$, respectively. It may be shown that these two Gibbs states differ in a region
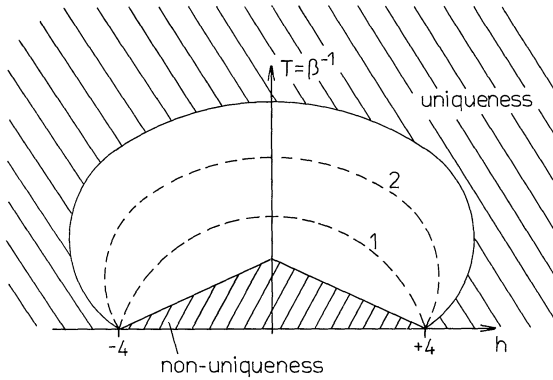
$$\beta(4 - |h|) > \bar{\mu} \tag{1.2}$$

Fig. 1. Phase diagram of the 2D Ising antiferromagnet

of parameters (h, β), with $\bar{\mu}$ a positive constant. The proof is made with the help of standard arguments based on a Peierls inequality [1]. This statement and the following ones are explained in detail in the survey [2].

Using methods described in [1] (Theorem 1) one may show that there is a unique Gibbs state of the model (1.1) in a region $\beta \leq \beta_0$, where β_0 is small enough, as well as in a region $|h| \geq h_0$, where h_0 is large enough. Finally, the ground configuration is unique for $|h| > 4$, and thus one has again a uniqueness in a region $[|h| > 4, \ \beta \geq \bar{\beta}(h)]$ with some $\bar{\beta}(h) < \infty$ [3]. Summarizing, the regions with a structure of Gibbs states rigorously known before are shown in Fig. 1.

A conjectured curve of a second order phase transition must thus pass through a non-hatched region in Fig. 1 connecting points $(-4, 0)$ and $(4, 0)$. One may imagine two essentially different behaviours of this curve as shown in Fig. 1 by dashed lines. The second curve bulges over the $|h| \geq 4$ parts of the h-axis in such a way that for some h there would be two phase transitions in temperature. The physical literature takes it for granted that the first possibility takes place. However, the evidence for it is based on considerations and calculations [4] which are not a proof in a mathematical sense. Actually, there are some other models, not very far from the one studied here, for which the second possibility takes place [2]. It has been suggested that an antiferromagnet on a bcc lattice is just of this type [5].

The main result of this work consists of the proof, a mathematically rigorous one[1], that the curve of a phase transition cannot look like the second line in Fig. 1. More precisely, we will extend the region in which a uniqueness of the Gibbs state can be proven.

Namely, we have

Theorem 1. *There exist θ, $\pi > \theta > \frac{\pi}{2}$, and $r > 0$ such that there is a unique Gibbs state of this antiferromagnet with parameters (h, T) in the domain $\{(h, T) | h - 4 = \tilde{r} \cos \tilde{\theta}, T = \tilde{r} \sin \tilde{\theta}, 0 \leq \tilde{\theta} \leq \theta, 0 \leq \tilde{r} \leq r\}$. This sector of the plane is drawn in Fig. 2 which is a magnified piece of Fig. 1.*

---

1   We leave aside the question whether efforts to prove in a mathematical sense a fact which already appears vindicated by physical considerations are justified, since this question may be discussed only from philosophical or aesthetical positions
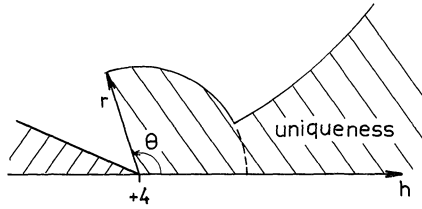
Fig. 2. Magnified piece of the phase diagram

Hence the curve of the phase transition does not bulge above the half-line $h \geq 4$ [at least near $(h=4, T=0)$]. The phase diagram near $(h=-4, T=0)$ is clearly symmetrical.

However, the main goal of our work is to advertize the method used in the proof rather than the result itself. Our method is based on a constructive criterion of uniqueness presented by two authors of this article in [6]. Let us first recall that in [6] a system of conditions C_V (for Gibbsian conditional distributions in a volume V), $V \subset \mathbb{Z}^2$, $|V| < \infty$ was constructed. These conditions have the property that if an interaction and β are such that a condition C_V is true for a finite volume V, then there is a unique Gibbs state for these interaction and β, as well as for interactions and β's in its neighbourhood. The condition C_V for $|V|=1$ is essentially the same as that presented in [1] (the mean-field approximation). The set of all conditions C_V exhausts all the region of interactions in which an analyticity (in a natural sense) takes place (and even strictly contains it!) [7]. When a volume V enlarges, it becomes more complicated to verify the condition C_V, so that it may be necessary to use a computer. On the other hand, a region of interactions fulfilling the condition C_V enlarges if the volume V increases.

In this work we have tested an efficiency of a criterion given in [6]. It turned out that the proof of no-bulging near $(h=4, T=0)$ follows from the condition C_V for a rectangular volume $V = 3 \times 4$. It is natural to expect that a sector of uniqueness (see Fig. 2) will enlarge if the volume V increases and, in the limit $V \to \infty$, will touch the phase transition curve near $(4, 0)$.

Our paper is organized as follows. In Sect. 2 we discuss a connection between the Ising antiferromagnet and the hard-square lattice gas [the latter turns out to be a ground state of the former at the point $(h=4, T=0)$]. In Sect. 3 we rewrite the uniqueness criterion of [6] for the hard-square gas and deduce our main statement. Section 4 is devoted to some details of checking the conditions C_V $(V = 2 \times 2$ and $V = 3 \times 4)$ for the hard-square gas.

2. The Hard-Square Lattice Gas

It is a specific feature of the Ising antiferromagnet that it possesses an infinite number of ground configurations at the points $(h = \pm 2d = \pm 4, T=0)$. One would thus take for the ground states not every ground configuration itself but rather some special measures on these configurations. Such an approach is developed in detail in [2] where a ground state is defined as a state which is consistent with a ground specification. The ground specification, in its turn, is defined as a consistent

set of conditional distributions which are concentrated, for any boundary condition, on configurations which minimize a functional of a relative energy.

In the following we shall consider a neighbourhood of the point $(h=4, T=0)$. It can be directly shown [2] that all configurations σ for which

$$|s-t|>1 \quad \text{whenever} \quad \sigma_s=\sigma_t=-1 \qquad (2.1)$$

are ground configurations at this point (we take $|s|=|s_1|+|s_2|$ for $s \in \mathbb{Z}^2$); in other words, spin values (-1) cannot be neighbouring. We shall call compatible all configurations with this property. The model with such configurations is known as the model of hard squares: let to any $\sigma_t=-1$ correspond an open square with edges of length $\sqrt{2}$ and with vertices in points $(t_1 \pm 1, t_2)$ and $(t_1, t_2 \pm 1)$. Then the condition (2.1) just means that squares do not intersect.

Look now for measures arising naturally on these ground configurations of the Ising antiferromagnet at the point $(4,0)$. To do this, let us consider conditional distributions $q_A^{h;\beta}(\cdot \mid \cdot)$ of the Ising antiferromagnet in a volume Λ as functions of parameters h and β, and let us consider the limit $\beta \to \infty$, $h \to 4$ along the direction

$$h=4+\mu\beta^{-1}. \qquad (2.2)$$

We obtain a specification $q_A^\mu(\cdot \mid \cdot)$ for the system of hard squares with a chemical potential μ which corresponds to the interaction Φ^μ:

$$\Phi_A^\mu(\sigma_A)=\begin{cases} +\infty, & A=\{s,t\}, \quad |s-t|=1, \quad \sigma_s=\sigma_t=-1, \\ -\mu\sigma_t, & A=\{t\}, \\ 0 & \text{otherwise}. \end{cases} \qquad (2.3)$$

A value $\sigma_t=-1$ corresponds to a particle at the site t, while $\sigma_t=+1$ to a vacancy. [Usually the hard-square gas is described by assigning the value 1 for a particle and 0 for a vacancy. In these variables the chemical potential $\tilde{\mu}$ has a value (-2μ).] The observation about the possibility of describing ground states of the Ising antiferromagnet in terms of the hard-square gas goes back to [4] (see also [2]).

It has been shown [1] that for $\mu \leq -\bar{\mu}$, where $\bar{\mu}$ is a certain positive constant [the same as in (1.2)] there are, for the interaction Φ^μ, at least two periodic Gibbs states which go to configurations $\sigma^{(1)}, \sigma^{(2)}$ as $\mu \to -\infty$, and that there is a unique Gibbs state whenever $\mu \geq \bar{\mu}$. By the methods used in [1] one may obtain an estimate

$$\bar{\mu} \leq \tfrac{1}{2}\ln 3. \qquad (2.4)$$

It is natural to conjecture that there exists a value μ_{cr}, $-\bar{\mu} \leq \mu_{\mathrm{cr}} \leq \bar{\mu}$, which separates regions of uniqueness and non-uniqueness. In addition, it is natural to expect that for $\mu > \mu_{\mathrm{cr}}$ a Gibbs state of the Ising antiferromagnet with h and β satisfying (2.2) ought to be unique, while for $\mu < \mu_{\mathrm{cr}}$ and $\beta > \beta(\mu)$ non-unique. It would mean that the value μ_{cr} determines an angle at which the line of the phase transition in the Ising antiferromagnet intersects the h-axis.

In Sect. 4 we state results of computations which prove, by checking a condition C_V, that every interaction which is close enough to the interaction of

hard squares generates a unique Gibbs state whenever $\mu \geq 0$. In this way we prove the statement about the Ising antiferromagnet formulated in Theorem 1. (By saying that interactions are close, we mean that their specifications are close.) In addition, we enlarge a region (2.4) in which uniqueness for the hard-square model was proven earlier.

Numerical methods (the corner transfer matrix method [4]) give the following value for the critical chemical potential:

$$\mu_{cr} = -\tfrac{1}{2}\ln(3.7962 \pm 0.0001).$$

However, this estimate has not been verified on a mathematical level of rigour.

We hope that our estimate of the uniqueness region can be significantly improved by checking the condition C_V for larger volumes V.

3. Condition of Uniqueness

In this section we state our criterion of uniqueness in a form suitable for the hard-square model. A general formulation of the criterion is given in [6].

Let us recall some notations. Let $(X, \bar{\varrho})$ be a metric space with a finite number of points and ξ, η two probability measures on X. A joint distribution of ξ, η is a probability measure P on $X \times X$ such that for every subset $Y \subset X$:

$$\begin{aligned} P(Y \times X) &= \xi(Y), \\ P(X \times Y) &= \eta(Y). \end{aligned} \tag{3.1}$$

We denote by $\mathscr{P}(\xi, \eta)$ the set of all joint distributions of measures ξ, η. The Kantorovich (Kantorovich-Rubinstein-Ornstein-Vasserstein) distance $R(\xi, \eta)$ is defined by the formula

$$R(\xi, \eta) = \min_{P \in \mathscr{P}(\xi, \eta)} \sum_{x, y \in X} \bar{\varrho}(x, y) P(x, y). \tag{3.2}$$

Let now $V \subset \mathbb{Z}^2$ be an arbitrary finite volume and $t_0 \in \partial V = \{t \in \mathbb{Z}^2 \setminus V \mid \text{dist}(t, V) = 1\}$. We introduce a metric ϱ on the set of all configurations $\Omega_V = \{\sigma_V \mid \sigma_V : V \to \{-1, +1\}\}$ by

$$\varrho(\sigma_V^{(1)}, \sigma_V^{(2)}) = \tfrac{1}{2} \sum_{t \in V} |\sigma_V^{(1)}(t) - \sigma_V^{(2)}(t)|. \tag{3.3}$$

Let $\tilde{\sigma} \in \Omega_{\partial V \setminus t_0}$ be a configuration; we define the configurations $\tilde{\sigma}^{\pm} \in \Omega_{\partial V}$ by the formula

$$\tilde{\sigma}^{\pm}(t) = \begin{cases} \tilde{\sigma}(t) & \text{for } t \neq t_0, \\ \pm 1 & \text{for } t = t_0. \end{cases} \tag{3.4}$$

Finally, let Φ be a pair nearest neighbour interaction [i.e. a translation invariant system of functions $\Phi_{\{s\}}(\sigma)$, $\Phi_{\{s,t\}}(\sigma, \tau)$, $s, t \in \mathbb{Z}^d$, $|s - t| = 1$, $\sigma, \tau = \pm 1$ with values in $\mathbb{R}^1 \cup \{\infty\}$], and $q_V^{\Phi}(\cdot \mid \tilde{\sigma}^{\pm})$ the corresponding conditional Gibbs distribution in a volume V with the boundary condition $\tilde{\sigma}^{\pm}$. We set

$$k_{t_0}^V = \max_{\tilde{\sigma}} R(q_V^{\Phi}(\cdot \mid \tilde{\sigma}^+), q_V^{\Phi}(\cdot \mid \tilde{\sigma}^-)). \tag{3.5}$$

Theorem 2. *Let $V \subset \mathbb{Z}^2$ be a finite volume such that an interaction Φ satisfies the following condition C_V:*

$$\sum_{t \in \partial V} k_t^V < |V|. \tag{3.6}$$

Then there is a unique Gibbs state for the interaction Φ.

The proof is given in [6].

We say that pair nearest neighbour interactions Φ' and Φ'' are ε-close if

$$|\exp(-\Phi'_{\{s\}}(\sigma)) - \exp(-\Phi''_{\{s\}}(\sigma))| < \varepsilon$$

and $\hspace{10cm}$ (3.7)

$$|\exp(-\Phi'_{\{s,t\}}(\sigma, \tau)) - \exp(-\Phi''_{\{s,t\}}(\sigma, \tau))| < \varepsilon$$

for all $\sigma, \tau = \pm 1$ and all points $s \in \mathbb{Z}^d$ and all pairs $\{s, t\}$ of the nearest neighbours.

Corollary 1. *Let an interaction Φ satisfy the condition C_V with some $V \subset \mathbb{Z}^2$. Then there exists $\varepsilon_V > 0$ such that the statement of a uniqueness in Theorem 2 is valid for every interaction Φ', which is ε_V-close to Φ.*

Indeed, taking into account the definition (3.7), a value ε_V may be chosen in such a way that the condition C_V is fulfilled also for Φ'.

Corollary 2. *Suppose that the interaction Φ^μ of the hard-square lattice system satisfies the condition C_V (where V may depend on μ) whenever $\mu \geq 0$. Then*

(a) there exists $\mu_0 > 0$ such that the uniqueness takes place in the hard-square model whenever $\mu > -\mu_0$,

(b) the phase diagram of the Ising antiferromagnet looks as described in Theorem 1.

Proof of Corollary 2. Let us notice that the distance $R(q_V^1, q_V^2)$ is continuous function of values of measures $q_V^i(\sigma_V^{(1)}), ..., q_V^i(\sigma_V^{(2^{|V|})})$, $i = 1, 2$, $\sigma_V^{(j)} \in \Omega_V$. Hence we obtain (a) from Corollary 1. It is easy to show that $q_V^{h, \beta} \to q_V^\mu$ along the line (2.2) uniformly with respect to μ on any finite interval. So in order to prove (b) it is enough to show that an interval of inverse temperatures $[\beta_\mu, +\infty]$ exists for which the uniqueness for the Ising antiferromagnet with a magnetic field $h = 4 + \mu \beta^{-1}$ follows from the uniqueness for the hard squares with the chemical potential μ, and this interval does not contract to the point $+\infty$ as $\mu \to +\infty$. However, it can be easily shown that for every $\varepsilon > 0$ there exists $\mu(\varepsilon) < \infty$ such that for all $\mu > \mu(\varepsilon)$ the interactions $\beta \Phi^h$ (where Φ^h is an interaction of the Ising antiferromagnet for $h = 4 + \mu \beta^{-1}$) and Φ^μ are ε-close for all β! But

$$R(q_{\{t\}}^\mu(\cdot | \bar{\sigma}_1), q_{\{t\}}^\mu(\cdot | \bar{\sigma}_2)) \to 0$$

for $\mu \to \infty$ uniformly in $\bar{\sigma}_1, \bar{\sigma}_2$. Thus, taking into account Corollary 1, we can set $\beta_\mu = 0$ for μ large enough. In other words, if we consider the sector of a uniqueness for a small angle θ, we may take for its radius $r = +\infty$!

Remark. In this way we have obtained yet another proof of a uniqueness for the Ising antiferromagnet in a region $h > 4$: the uniqueness takes place whenever $\beta \geq \bar{\bar{\beta}}(h) \equiv C(h - 4)^{-1}$ for some $C < \infty$.

4. A Check of the Uniqueness Conditions with the Help of a Computer

According to the theory described above we should evaluate the Kantorovich distance between distributions $q_V^\mu(\cdot|\tilde{\sigma}^+), q_V^\mu(\cdot|\tilde{\sigma}^-)$. It has an important property that for any joint distribution $P \in \mathscr{P}(q_V^\mu(\cdot|\tilde{\sigma}^+), q_V^\mu(\cdot|\tilde{\sigma}^-))$ we have an estimate [see (3.2), (3.3)],

$$R(q_V^\mu(\cdot|\tilde{\sigma}^+), q_V^\mu(\cdot|\tilde{\sigma}^-)) \leqq d(P) \equiv \sum_{\sigma_V^1, \sigma_V^2 \in \Omega_V} P(\sigma_V^1, \sigma_V^2)\varrho(\sigma_V^1, \sigma_V^2). \tag{4.1}$$

Thus we need not look for an "optimal" joint distribution. It is enough just to guess a "good" one.

Our task is thus to find a finite volume V for each $\mu \in [0, +\infty]$ and a bound $K_{t_0}^V$ of $k_{t_0}^V$ for every $t_0 \in \partial V$, so that the condition C_V [see (3.6)]

$$\sum_{t \in \partial V} K_t^V < |V|$$

takes place with [see (3.5), (4.1)]

$$d(P) \leqq K_{t_0}^V \tag{4.2}$$

for all boundary conditions $\tilde{\sigma} \in \Omega_{\partial V \setminus t_0}$.

4.1. The Estimate P'' of the Joint Distribution

A joint distribution P may be considered as a matrix $P(\sigma_V^1, \sigma_V^2)$ labelled in rows by configurations σ_V^1, compatible with the boundary condition $\tilde{\sigma}^+$, and in columns by configurations σ_V^2, compatible with $\tilde{\sigma}^-$. Let us denote the corresponding sets by $\Omega_V(\tilde{\sigma}^\pm)$. Let us emphasize here that the configurations $\tilde{\sigma}^+$ or $\tilde{\sigma}^-$ itself need not be compatible. If we were interested in the hard-square model only we could confine ourselves to compatible boundary conditions $\tilde{\sigma}^\pm$. However, if we want to study a uniqueness of the antiferromagnet, we have to investigate the hard-square model in a way which is stable with respect to small perturbations, and we have thus to take into account all boundary conditions.

The most simple way to define a joint distribution is to set

$$P(\sigma_V^1, \sigma_V^2) = q_V^\mu(\sigma_V^1|\tilde{\sigma}^+) \; q_V^\mu(\sigma_V^2|\tilde{\sigma}^-). \tag{4.3}$$

However, this way is clearly not the most effective one. To see it, let us consider a square block of the matrix P which is labelled by $(\sigma_V^1, \sigma_V^2) \in A \equiv \Omega_V(\tilde{\sigma}^-) \times \Omega_V(\tilde{\sigma}^-)$, i.e. by pairs (σ_V^1, σ_V^2), which are compatible with both $\tilde{\sigma}^-$ and $\tilde{\sigma}^+$ (see Fig. 3). It can be easily shown that the matrix

$$P'(\sigma_V^1, \sigma_V^2) = \begin{cases} q_V^\mu(\sigma_V^1|\tilde{\sigma}^+), & \sigma_V^1 = \sigma_V^2, & (\sigma_V^1, \sigma_V^2) \in A, \\ 0, & \sigma_V^1 \neq \sigma_V^2, & (\sigma_V^1, \sigma_V^2) \in A, \\ P(\sigma_V^1, \sigma_V^2) & \text{otherwise} \end{cases} \tag{4.4}$$

is again a joint distribution of $q_V^\mu(\cdot|\tilde{\sigma}^+), q_V^\mu(\cdot|\tilde{\sigma}^-)$ and, of course, $d(P') \leqq d(P)$. We have used here the following elementary observation: if a matrix $P(\sigma_V^1, \sigma_V^2)$ of a
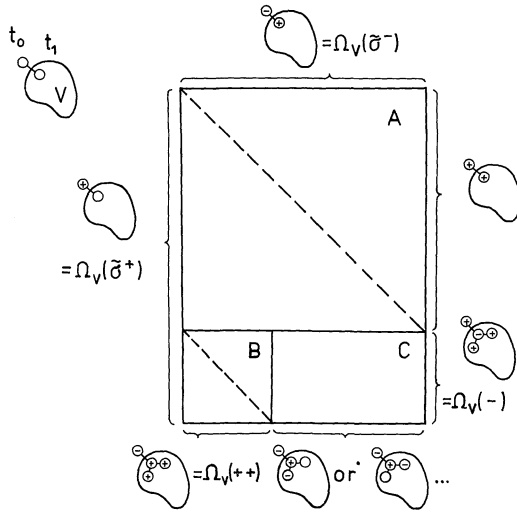
Fig. 3. Blocks A, B, and C of the joint distribution matrix

joint distribution may be represented as a sum

$$P(\sigma_V^1, \sigma_V^2) = Q_1(\sigma_V^1, \sigma_V^2) + Q_2(\sigma_V^1, \sigma_V^2)$$

with $Q_i(\sigma_V^1, \sigma_V^2) \geqq 0$ and if a matrix $\tilde{Q}_1$ satisfies the following conditions:

$$\tilde{Q}_1(\sigma_V^1, \sigma_V^2) \geqq 0,$$

$$\sum_{\sigma_V^1} \tilde{Q}_1(\sigma_V^1, \sigma_V^2) = \sum_{\sigma_V^1} Q_1(\sigma_V^1, \sigma_V^2),$$

and

$$\sum_{\sigma_V^2} \tilde{Q}_1(\sigma_V^1, \sigma_V^2) = \sum_{\sigma_V^2} Q_1(\sigma_V^1, \sigma_V^2),$$

then the matrix $P' = \tilde{Q}_1 + Q_2$ is a joint distribution as well. Constructing the matrix (4.4) we have used this observation and the fact that the block A (more precisely, the block labelled by pairs from A) is a symmetric matrix.

Now we shall introduce a block B of a joint distribution matrix. Let $t_1 \in V$ be a nearest neighbour of $t_0 \in \partial V$ (there is only one such t_1 if the volume V is rectangular). We say that $(\sigma_V^1, \sigma_V^2) \in B$ if

$$\sigma_V^1(t_1) = -1,$$

$$\sigma_V^2(t_1) = +1 \quad \text{and} \quad \sigma_V^2(t) = +1 \quad \text{whenever} \quad t \in V, \quad |t - t_1| = 1.$$

Let us denote the former set of configurations by $\Omega_V(-)$ and the latter by $\Omega_V(++)$ (see Fig. 3). In the same way as it was done with the block A we may change the symmetric square block B of the matrix P' into "diagonal" submatrix with a
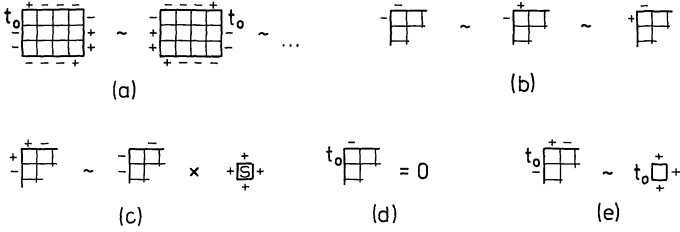
Fig. 4. Equivalence of boundary conditions

diagonal given by $\sigma_V^1(t) = \sigma_V^2(t)$ for $t \in V \backslash t_1$. Thus we obtain the matrix

$$P''(\sigma_V^1, \sigma_V^2)$$

$$= \begin{cases} q_V^\mu(\sigma_V^1|\tilde{\sigma}^+) \displaystyle\sum_{\sigma_V \in \Omega_V(++)} q_V^\mu(\sigma_V|\tilde{\sigma}^-) \\ \quad = q_V^\mu(\sigma_V^2|\tilde{\sigma}^-) \displaystyle\sum_{\sigma_V \in \Omega_V(-)} q_V^\mu(\sigma_V|\tilde{\sigma}^+), \quad \sigma_V^1(t) = \sigma_V^2(t), \quad t \in V \backslash t_1, \quad (\sigma_V^1, \sigma_V^2) \in B, \\ 0, \qquad\qquad \sigma_V^1(t) \neq \sigma_V^2(t) \quad \text{for at least one} \quad t \in V \backslash t_1, \quad (\sigma_V^1, \sigma_V^2) \in B, \\ P'(\sigma_V^1, \sigma_V^2) \quad \text{otherwise}. \end{cases}$$

Again $d(P'') \leqq d(P')$. Finally, we denote by C the block of pairs (σ_V^1, σ_V^2) which are not in A or B.

Remark. Clearly, an "optimal" matrix can be chosen from $\mathscr{P}(q_V^\mu(\cdot|\tilde{\sigma}^+), q_V^\mu(\cdot|\tilde{\sigma}^-))$ in such a way that it contains the same block A as P' or P''. However, an "optimal" matrix which contains the block B of P'' need not exist. But this is not very relevant for us since we are looking for an upper bound.

4.2. Eliminating of Some Boundary Conditions

As we have mentioned above we have to find a "good" joint distribution for all sites $t_0 \in \partial V$ and all boundary conditions $\tilde{\sigma} \in \Omega_{\partial V \backslash t_0}$. Since their number is considerable ($|\partial V|2^{|\partial V \backslash t_0|}$), let us first eliminate some of them for a case of a rectangular volume V using the following rules (see also Fig. 4).

(a) We use a space symmetry.

(b) If $t_1, t_2 \in \partial V \backslash t_0$ are neighbours of a corner-site $s \in V$, then all three boundary conditions with $\tilde{\sigma}(t_i) = -1$ for $i = 1$ or $i = 2$ or both, are equivalent [they give the same sets $\Omega_V(\tilde{\sigma}^{\pm})$].

(c) Let us denote by $\tilde{\sigma}_{-1}$ some boundary condition with t_1, t_2 and $\tilde{\sigma}(t_i)$ as above and with $t_1', t_2' \in \partial V \backslash t_0$ such that $|t_j - t_j'| = 1$, $j = 1, 2$ (thus $|t_1' - t_2'| = 4$), $\tilde{\sigma}_{-1}(t_j') = -1$. Moreover, let $\tilde{\sigma}_{+1}$ be such that $\tilde{\sigma}_{+1}(t_j') = -1$, $\tilde{\sigma}_{+1}(t_j) = +1$. Then $\Omega_V(\tilde{\sigma}_{+1}^{\pm}) = \Omega_V(\tilde{\sigma}_{-1}^{\pm}) \times \{\sigma(s) = +1, \sigma(s) = -1\}$ and knowing a joint distribution for $\tilde{\sigma}_{-1}$ we can easily obtain a joint distribution for $\tilde{\sigma}_{+1}$ with the same distance.

(d) If t_0 and $t \in \partial V \backslash t_0$ are neighbours of a corner-site $s \in V$, then a configuration $\tilde{\sigma} \in \Omega_{\partial V \backslash t_0}$ for which $\tilde{\sigma}(t) = -1$ gives the zero distance since $\Omega_V(\tilde{\sigma}^+) = \Omega_V(\tilde{\sigma}^-)$ and thus P can be taken as a diagonal square matrix.

Fig. 5. Distances for four nonequivalent boundary conditions in the $V = 2 \times 2$ case

Fig. 6. Example of the joint distribution matrix. $a = e^{-2\mu}$, $1/N = (1 + a)(1 + 2a)$

(e) Let t'_j, t_j be as in (c) but t_0 stays instead of t_2 and $\tilde{\sigma}_{+1}(t_1) = +1$. Then

$$R(q_V^\mu(\cdot \mid \tilde{\sigma}_{+1}^+), q_V^\mu(\cdot \mid \tilde{\sigma}_{+1}^-)) = \frac{1}{e^{2\mu} + 1},$$

which is the same as for $V = 1 \times 1 = \{s\}$.

4.3. Computation for $V = 2 \times 2$

As an illustration (and also to get a first bound of the uniqueness region) we shall consider a square $V = 2 \times 2$. We have four (nonequivalent in the sense described in Sect. 4.2) boundary conditions $\tilde{\sigma} \in \Omega_{\partial V \setminus t_0}$ in this case, see Fig. 5. The largest dimension of the block C is 3×2 in the case of the fourth boundary condition on Fig. 5 and the computation of $d(P'')$ for each of the four cases can be easily made "by hand." (Figure 6 shows an example of the matrix P'' for one boundary condition.) The result is

$$\max_{\tilde{\sigma}} d(P'') = \frac{a(1 + 2a)}{1 + 3a + a^2}, \tag{4.6}$$

where $a = e^{-2\mu}$. According to the condition C_V we need $K_t^V < \dfrac{|V|}{|\partial V|} = \dfrac{1}{2}$ for all $t \in \partial V$. Then (4.2) holds for

$$e^{-2\mu} = a \leq 0.76759 < \frac{1 + \sqrt{13}}{6} \equiv \exp(-2\mu_{2 \times 2}). \tag{4.7}$$

This is a better estimate of the uniqueness region than (2.4). (If we considered the volume $V = 1 \times 2$ we should obtain only $e^{-2\mu} < \frac{1}{2}$.)

4.4. Computation for $V = 3 \times 4$

However, our aim is to reach a bound $e^{-2\mu} \leqq 1$. Thus we have to consider larger volumes V (it turned out that $V = 3 \times 4$ is enough) and a finer guess of a "good" joint distribution. In principle we could continue on the procedure of changing square blocks of the matrix into diagonal submatrices [6], but we have chosen here to apply another, a more simple approach which is based on elementary theory of transport problems (so called eliminating of 2-cycles). Nevertheless, the volume of computations enlarges so that it does not pay to perform them "by hand" and we finally algorithmize the problem for a computer.

To this end, we write the distance in the form

$$d(P) = d_A(P) + d_B(P) + d_C(P), \tag{4.8}$$

where

$$d_A(P) = \sum_{(\sigma_V^1, \sigma_V^2) \in A} P(\sigma_V^1, \sigma_V^2) \varrho(\sigma_V^1, \sigma_V^2) \tag{4.9}$$

and $d_B(P), d_C(P)$ are defined analogically. [Note that, by definition, $d_A(P'') = 0$.]

Now we describe an inductive algorithm which modifies the block C of the joint distribution matrix (we use the matrix P'' as an initial estimate) so that the partial distance d_C decreases.

Algorithm. *Let any initial matrix P be given.*

Step 1. One chooses randomly a pair of configurations $(\sigma_V^1, \sigma_V^2) \in C$ *so that* $P(\sigma_V^1, \sigma_V^2) \neq 0$.

Step 2. One chooses randomly a pair $(\tau_V^1, \tau_V^2) \in C$, $\tau_V^1 \neq \sigma_V^1$, $\tau_V^2 \neq \sigma_V^2$.

Step 3. One tries to change the matrix

$$
\begin{pmatrix}
 & \vdots & & \vdots & \\
\cdots & P(\sigma_V^1, \sigma_V^2) & \cdots & P(\sigma_V^1, \tau_V^2) & \cdots \\
 & \vdots & & \vdots & \\
\cdots & P(\tau_V^1, \sigma_V^2) & \cdots & P(\tau_V^1, \tau_V^2) & \cdots \\
 & \vdots & & \vdots &
\end{pmatrix}
$$

into   (4.10)

$$
\begin{pmatrix}
 & \vdots & & \vdots & \\
\cdots & P(\sigma_V^1, \sigma_V^2) + x & \cdots & P(\sigma_V^1, \tau_V^2) - x & \cdots \\
 & \vdots & & \vdots & \\
\cdots & P(\tau_V^1, \sigma_V^2) - x & \cdots & P(\tau_V^1, \tau_V^2) + x & \cdots \\
 & \vdots & & \vdots &
\end{pmatrix}
$$

with x being extremal such that all new matrix elements $P(\cdot, \cdot)$ *remain non-negative (and then at least one of them becomes zero) and such that the new distance is as small as possible.*

If no change is possible (only $x = 0$ *is allowed) the procedure continues by Step 1.*

If the change gives a lower distance it is accepted and we again go to Step 1.

If *the change with any (extremal and nonzero) x does not alter the distance, i.e. if*

$$\varrho(\sigma_V^1, \sigma_V^2) + \varrho(\tau_V^1, \tau_V^2) = \varrho(\sigma_V^1, \tau_V^2) + \varrho(\tau_V^1, \sigma_V^2), \tag{4.11}$$

it is accepted, too, and the pair (σ', σ'') *for which* $P(\sigma', \sigma'')$ *has been diminished to zero is denoted by* (σ_V^1, σ_V^2) *and the procedure continues by Step 2. (If both extremal x are nonzero then that value is used which was first hit while testing, i.e. random in effect.)*

After the ith change we obtain a matrix P^i and we can test if the partial distance $d_C(P^i)$ falls below some given bound.

It is a typical feature of a structure of the sets $\Omega_V(\tilde{\sigma}^{\pm})$ that the block C usually contains a lot of "squares" satisfying (4.11). After several steps of Algorithm the matrix P contains zeros and the third "if" causes an intensive shuffling of the matrix content. Hence those "squares" for which (4.11) is not valid, but which contain a zero matrix element so that the d_C-diminishing change (4.10) is not possible, can gain a nonzero element so that the d_C-diminishing change is again possible. This part of the algorithm is thus expected to speed up a convergence. Of course, nothing is said about a convergence to the absolute minimum of the distance.

It is naturally impossible to perform this procedure for a continuum of μ's in an interval $\mu \in [0, \mu_{2 \times 2}]$. So we used the following trick. Let $[\mu', \mu'']$ be any subinterval of an investigated region of chemical potentials. Let

$$\mu' \equiv \mu_0 < \mu_1 < \dots < \mu_n \equiv \mu'' \tag{4.12}$$

be an appropriate partition of the interval $[\mu', \mu'']$. Let

$$
\begin{aligned}
&P_{[\mu', \mu'']}(\sigma_V^1, \sigma_V^2) \\
&= \min_{0 \le j \le n} \frac{\exp(-\mu_{j+1}|\sigma_V^1|) \exp(-\mu_{j+1}|\sigma_V^2|)}{\sum\limits_{\sigma_V \in \Omega_V(\tilde{\sigma}^+)} \exp(-\mu_j|\sigma_V|) \sum\limits_{\sigma_V \in \Omega_V(\tilde{\sigma}^-)} \exp(-\mu_j|\sigma_V|)},
\end{aligned}
\tag{4.13}
$$

where $(\sigma_V^1, \sigma_V^2) \in C$ and $|\sigma_V| = \sum\limits_{t \in V} \sigma_V(t)$, be a matrix (not necessarily a joint distribution!) and

$$
\begin{aligned}
&\Delta P_{[\mu', \mu'']}(\sigma_V^1, \sigma_V^2) \\
&= \max_{0 \le j \le n} \frac{\exp(-\mu_j|\sigma_V^1|) \exp(-\mu_j|\sigma_V^2|)}{\sum\limits_{\sigma_V \in \Omega_V(\tilde{\sigma}^+)} \exp(-\mu_{j+1}|\sigma_V|) \sum\limits_{\sigma_V \in \Omega_V(\tilde{\sigma}^-)} \exp(-\mu_{j+1}|\sigma_V|)} \\
&\quad - P_{[\mu', \mu'']}(\sigma_V^1, \sigma_V^2).
\end{aligned}
\tag{4.14}
$$

Thus the elements of the matrix P'' in the block C are bounded as follows

$$
\begin{aligned}
P_{[\mu', \mu'']}(\sigma_V^1, \sigma_V^2) &\le P''(\sigma_V^1, \sigma_V^2) \\
&\le P_{[\mu', \mu'']}(\sigma_V^1, \sigma_V^2) + \Delta P_{[\mu', \mu'']}(\sigma_V^1, \sigma_V^2),
\end{aligned}
\tag{4.15}
$$

whenever $\mu \in [\mu', \mu'']$. Let now $P^i_{[\mu', \mu'']}$ be a matrix obtained by Algorithm after the ith change from the matrix $P_{[\mu', \mu'']}$. (Note that the blocks A and B remain unchanged.) Then the matrix

$$\tilde{P}^i_\mu \equiv P^i_{[\mu', \mu'']} + (P'' - P_{[\mu', \mu'']}) \tag{4.16}$$

Table 1. Bounds $K_{t_0}^V$ of the Kantorovich distance for a volume $V = 3 \times 4$. The last column contains a number of nonequivalent boundary conditions

Neighbour of $t_0 \in \partial V$	Symmetry	$K_{t_0}^V$	
Centre of shorter side	$2 \times$	0.81599	214
Corner	$8 \times$	0.865	275
Interior of longer side	$4 \times$	0.862	350
Sum		11.99998	839

is a joint distribution and using (4.15) to estimate the term in parentheses, we obtain finally

$$d_C(\tilde{P}_\mu^i) \leqq d_C(P_{[\mu',\mu'']}^i) + d_C(\Delta P_{[\mu',\mu'']})$$ (4.17)

uniformly for $\mu \in [\mu', \mu'']$.

An estimate of $d_B(P'')$ was made by a similar trick.

4.5. Computer Proof

Algorithm and the formulae (4.13)–(4.17) are now our main tools to complete the proof. They were implemented, for a volume $V = 3 \times 4$, on a computer. Table 1 shows three nonequivalent sites $t_0 \in \partial V$ and corresponding bounds $K_{t_0}^V$ which were used in the program. They satisfy the condition C_V:

$$\sum_{t \in \partial V} K_t^V = 11.99998 < 12 = |V|.$$

The interval of interest $e^{-2\mu} \in [1.00001, 0.76759] \equiv [a_I, a_{II}]$ was split into m ($3 \leqq m \leqq 8$) nonequidistant subintervals. In each subinterval (denoted by $[\mu', \mu'']$) an estimate of d_C (4.17) and d_B was computed. Here also a nonequidistant partition (4.12) of $[\mu', \mu'']$ into $n = 12$ subintervals was used, and this is the partition entering the estimates (4.13)–(4.14). Let us remark that the distance increases for activities near a_I, and thus we need more accurate estimates in this region to check the condition C_V. That is why we thickened the partition points in this region using the formula

$$a_k = a_I + \frac{k^2 - 1}{(mn + 1)^2 - 1}(a_{II} - a_I), \quad 1 \leqq k \leqq mn + 1.$$

Now

$$[a_I, a_{II}] = \bigcup [e^{-2\mu'}, e^{-2\mu''}] \equiv \bigcup_{i=1}^m [a_{1+(i-1)n}, a_{1+in}]$$

and

$$[e^{-2\mu'}, e^{-2\mu''}] = \bigcup_{j=1+(i-1)n}^{in} [a_j, a_{j+1}].$$

Actually, this choice is more a guess than an optimal one. The choice of m depends on a boundary condition. For a majority of them the value $m = 3$ was sufficient, however, in several important cases of boundary conditions with only few particles we had to use more subintervals.

While running the program, a joint distribution matrix P was found for all (nonequivalent) sites t_0 and boundary conditions and all subintervals $[\mu', \mu'']$ so that in every case the distance $d(P)$ falls below the bound $K_{t_0}^V$. Thus a uniqueness for $e^{-2\mu} \in [1.00001, 0.76759]$ and, including (4.7), for $\mu \in [0, +\infty]$ is proven.

Finally, we add some remarks about the practical realization. The algorithm was programmed in FORTRAN. We have used the "double precision" arithmetic (8-byte representation of real numbers with accuracy about 16 decimal digits) to calculate powers of activities, sums, fractions etc.; rounding errors are unessential at these steps. In order to avoid a cumulation of rounding errors during an execution of Algorithm, and also to speed up the run, we used an integer arithmetic in this part of the program; a conversion from real was managed so that the conversion errors do not violate the inequalities (4.13)–(4.14). Before computing the resulting distance a test of row and column checksums and of a positivity of the matrix P was performed.

The source program contains 348 FORTRAN statements and has been compiled by an optimalized H-FORTRAN on the EC 1040. The runs of the proof took 1.5 h of CPU time and required 108 kbyte of memory. A copy of the program is available upon request from the second author.

Finally, we should note that a similar computation for $\mu = 0$ and a volume $V = 3 \times 3$ leads only to an upper (not necessarily optimal) bound of the sum

$$\sum_{t \in \partial V} K_t^V = 9.139 \nless |V|.$$

5. Discussion

The algorithm for checking uniqueness conditions which we have used in this work is not the best one. Firstly, it does not necessarily yield the minimum of the Kantorovich distance since it only eliminates 2-cycles in the transport problem. Moreover, it takes only partially into account some specific features of the problem, e.g. the fact that both the measures $q_V^\mu(\cdot|\tilde{\sigma}^+)$ and $q_V^\mu(\cdot|\tilde{\sigma}^-)$ differ essentially only near the site t_0 (cf. the algorithm of [6]). Nevertheless, for a relatively small volume $V = 3 \times 4$ the given algorithm is quite satisfactory. In general a choice of a suitable algorithm must depend on the concrete problem and on an accessible computing power. We believe that an application of our methods with the help of more powerful computers can essentially enlarge regions in which a uniqueness can be proven. A generalization of our results to different three-dimensional lattices seems to be especially interesting.

References

1. Dobrushin, R.L.: Unicity of the Gibbs random field and phase transition problem. Funct. Anal. Appl. **2**, 44 (1968)
2. Dobrushin, R.L., Shlosman, S.B.: The problem of translation invariance in statistical physics. Sov. Sci. Rev. Ser. C Vol. **5** (1985)
3. Dobrushin, R.L., Pecherski, E.A.: Uniqueness conditions for finitely dependent random fields. In: Random fields. Fritz, J., Lebowitz, J.L., Szasz, D. (eds.). Amsterdam: North-Holland
4. Baxter, R.J., Enting, I.G., Tsang, S.K.: Hard-square lattice gas. J. Stat. Phys. **22**, 465–489 (1980)
5. Ràcz, Z.: Phase boundary of Ising antiferromagnets near $H = H_c$ and $T = 0$; Results from hard-core lattice gas calculations. Phys. Rev. B **21**, 4012–4016 (1980)
6. Dobrushin, R.L., Shlosman, S.B.: Constructive unicity criterion. In: Statistical mechanics and dynamical systems. Fritz, J., Jaffe, A., Szasz, D. (eds.). New York: Birkhäuser 1985
7. Dobrushin, R.L., Shlosman, S.B.: Completely analytical Gibbs fields. In: Statistical mechanics and dynamical systems. Fritz, J., Jaffe, A., Szasz, D. (eds.). New York: Birkhäuser 1985