10 Reflection Positivity

In this chapter, we study models whose Gibbs distribution possesses a remarkable property: *reflection positivity*. Two consequences of this property, the *chessboard estimate* and the *infrared bound*, will be described in a general setting.

Before that, in order to motivate this approach, we describe the two main applications that will be discussed in this chapter. Of course, there are many other such applications, reflection positivity playing a crucial role in a large numbers of proofs in this field.

10.1 Motivation: some new results for $\mathrm{O}(N)$-type models

We remind the reader that in $\mathrm{O}(N)$ models, which were already discussed in Section 9.1, the spins take their values in the N-dimensional sphere ($N \geq 2$),

$$\Omega_0 \stackrel{\text{def}}{=} \mathbb{S}^{N-1} \subset \mathbb{R}^N ,$$

and have the formal Hamiltonian

$$-\beta \sum_{\{i,j\} \in \mathscr{E}_{\mathbb{Z}^d}} \mathbf{S}_i \cdot \mathbf{S}_j ,$$

where $\mathbf{S}_i(\omega) \stackrel{\text{def}}{=} \omega_i$ denotes the spin at $i \in \mathbb{Z}^d$ and the symbol $\cdot$ denotes the scalar product in $\mathbb{R}^N$. We denote by $\mathscr{G}(\beta)$ the set of Gibbs measures for this model at inverse temperature β. In Chapter 9, we proved that, on $\mathbb{Z}^2$, the invariance of Φ under a global rotation of the spins leads to the absence of orientational long-range order at any positive temperature. In particular, we showed that the distribution of the spin at the origin is uniform on $\mathbb{S}^{N-1}$: for all $\mu \in \mathscr{G}(\beta)$, $\langle \mathbf{S}_0 \rangle_\mu = \mathbf{0}$.

In contrast, in Section 10.5.2, we will prove that, in larger dimensions, the global symmetry under rotations is spontaneously broken at low temperature.

Theorem 10.1. *Assume that $N \geq 2$ and $d \geq 3$. There exists $0 < \beta_0 < \infty$ and $m^* = m^*(\beta) > 0$ such that, whenever $\beta > \beta_0$, for each direction $\mathbf{e} \in \mathbb{S}^{N-1}$, there exists $\mu^{\mathbf{e}} \in \mathscr{G}(\beta)$ such that*

$$\langle \mathbf{S}_0 \rangle_{\mu^{\mathbf{e}}} = m^* \mathbf{e} .$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

In our second application, in Section 10.4.3, we will consider the anisotropic XY model on $\mathbb{Z}^2$ (although the argument applies as well to higher values of d and N). This model was introduced in Remark 9.4; its spins take values in $\mathbb{S}^1$ and the formal Hamiltonian is given by

$$-\beta \sum_{\{i,j\}\in\mathscr{E}_{\mathbb{Z}^d}} \{S_i^1 S_j^1 + \alpha S_i^2 S_j^2\},$$

where $0 \le \alpha \le 1$ is the **anisotropy parameter** and we have written $\mathbf{S}_i = (S_i^1, S_i^2)$. We denote the set of Gibbs measures at inverse temperature β and anisotropy α by $\mathscr{G}(\beta,\alpha)$.

When $\alpha = 1$, this model reduces to the XY model and we have seen in Section 9.2 that there is no spontaneous magnetization in this case. The next theorem shows that, in the presence of an arbitrary weak anisotropy, there are Gibbs measures displaying spontaneous magnetization at low temperature.

Theorem 10.2. *Assume that $N = 2$ and $d = 2$. For any $0 \le \alpha < 1$, there exists $\beta_0 = \beta_0(\alpha)$ such that, for all $\beta > \beta_0$, there exist $\mu^+, \mu^- \in \mathscr{G}(\beta,\alpha)$ such that*

$$\langle \mathbf{S}_0 \cdot \mathbf{e}_1 \rangle_{\mu^+} > 0 > \langle \mathbf{S}_0 \cdot \mathbf{e}_1 \rangle_{\mu^-} .$$

Remark 10.3. Whenever $\alpha \in [0,1)$, the system possesses exactly two configurations with minimal energy: those in which the spins are either all equal to $+\mathbf{e}_1$ or all equal to $-\mathbf{e}_1$ (see Exercise 10.5). This makes it reasonable to implement a suitable version of the Peierls argument. Note, however, that the continuous nature of the spins does not allow us to apply directly the results of Pirogov–Sinai theory developed in Chapter 7, although extensions covering such situations exist [1]. ⋄

10.2 Models defined on the torus.

Positivity under reflections is naturally formulated for measures which are invariant under reflections through planes perpendicular to some coordinate axis of $\mathbb{Z}^d$. Since most of the finite systems considered previously in the book are only left invariant by a few, if any, such reflections, it turns out to be much more convenient, in this chapter, to consider finite-volume Gibbs measures with *periodic boundary conditions*.

Let us therefore denote by $\mathbb{T}_L$ the **d-dimensional torus of linear size $L > 0$**, which is obtained by identifying the opposite sides of the box $\{0,1,\dots,L\}^d$ (remember the one- and two-dimensional tori depicted on Figure 3.1). Equivalently, we can set $\mathbb{T}_L \overset{\text{def}}{=} (\mathbb{Z}/L\mathbb{Z})^d$. Note that, to lighten the notation, we will only indicate explicitly the dimensionality of the torus when the latter might not be clear from the context.

We will transfer various notions from $\mathbb{Z}^d$ to the torus. For example, we will continue using the translation by i, denoted θ_i. We denote by $\mathscr{E}_L$ the set of all edges between nearest-neighbor vertices of $\mathbb{T}_L$. (The models that fit in the framework of this chapter are not restricted to nearest-neighbor interactions, but we introduce this set for later convenience.)

As always, the single-spin space is denoted Ω_0 and the set of spin configurations on the torus is

$$\Omega_L \overset{\text{def}}{=} \Omega_0^{\mathbb{T}_L} = \{\omega = (\omega_i)_{i\in\mathbb{T}_L} : \omega_i \in \Omega_0\} .$$

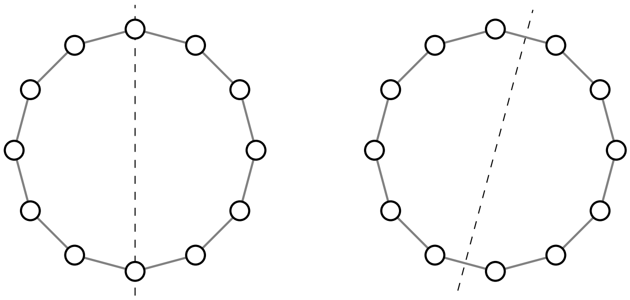

Figure 10.1: The one-dimensional torus $\mathbb{T}_{12}$, with a reflection through vertices (on the left) and through edges (on the right).

Even though the models to which we later apply the theory will have either $\Omega_0 = \mathbb{S}^{N-1}$ or $\Omega_0 = \mathbb{R}^N$, the theory has no such limitations. In fact, one of the arguments used below will require allowing Ω_0 to be far more general. Let us thus assume that Ω_0 is some topological space, on which one can define the usual Borel σ-algebra $\mathscr{B}_0$, generated by the open sets. (These notions are introduced in Section 6.10.1 and Appendix B.5.) The product σ-algebra of events on Ω_L is denoted simply $\mathscr{F}_L = \bigotimes_{i\in\mathbb{T}_L} \mathscr{B}_0$. The set of measures on $(\Omega_L, \mathscr{F}_L)$ is denoted $\mathscr{M}(\Omega_L, \mathscr{F}_L)$.

Remark 10.4. In the sequel, we will always assume L to be even. Moreover, since *all* the models considered in this chapter will be defined on $\mathbb{T}_L$, we will substantially lighten the notations by using everywhere a subscript L instead of $\mathbb{T}_L$. For example, a Gibbs distribution on Ω_L will be denoted μ_L instead of $\mu_{\mathbb{T}_L}$. ⋄

10.3 Reflections

We shall consider transformations on the torus,

$$\Theta : \mathbb{T}_L \to \mathbb{T}_L,$$

associated to *reflections* through planes that split the torus in two. (This Θ is not to be mistaken with the translation θ_i.) Before moving on to the precise definitions, the reader is invited to take a look at Figures 10.1 and 10.2, where the meaning of these reflections is made transparent.

▶ *Reflection through vertices* : Let $k \in \{1,\dots,d\}$ denote one among the d possible directions parallel to the coordinate axes and $n \in \{0,\dots,\frac{1}{2}L-1\}$. The **reflection through vertices** $\Theta : \mathbb{T}_L \to \mathbb{T}_L$ (associated to k and n), which maps $i = (i_1,\dots,i_d)$ to $\Theta(i) = (\Theta(i)_1,\dots,\Theta(i)_d)$, is defined by

$$\Theta(i)_\ell \stackrel{\text{def}}{=} \begin{cases} (2n - i_k) \mod L & \text{if } \ell = k, \\ i_\ell & \text{if } \ell \neq k. \end{cases} \tag{10.1}$$

Θ is a reflection of the torus through a plane Π which is orthogonal to the direction $\mathbf{e}_k$. The intersection between the plane and the torus is given by

$$\Pi \cap \mathbb{T}_L = \{i \in \mathbb{T}_L : i_k = n \text{ or } i_k = n + L/2\}.$$

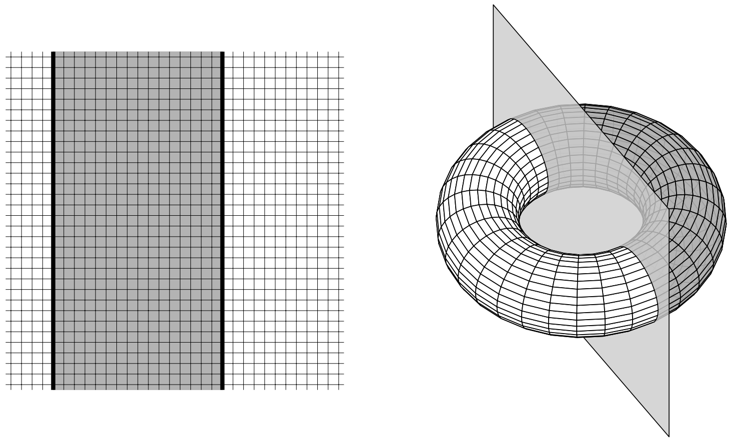

Figure 10.2: In $d = 2$, a reflection through the vertices of $\mathbb{T}_{32}$. The sets $\mathbb{T}_{32,+}$ and $\mathbb{T}_{32,-}$ are drawn in white and gray, respectively, and the intersection of the reflection plane and the torus is represented by the two thick lines. Left: planar representation. Right: Spatial representation.

This also leads to a natural decomposition of the torus into two overlapping halves: $\mathbb{T}_L = \mathbb{T}_{L,+} \cup \mathbb{T}_{L,-}$, where

$$\mathbb{T}_{L,+} = \mathbb{T}_{L,+}(\Theta) \stackrel{\text{def}}{=} \{i \in \mathbb{T}_L : n \le i_k \le n + L/2\},$$
$$\mathbb{T}_{L,-} = \mathbb{T}_{L,-}(\Theta) \stackrel{\text{def}}{=} \{i \in \mathbb{T}_L : 0 \le i_k \le n \text{ or } n + L/2 \le i_k \le L-1\}.$$

► *Reflection through edges*: The **reflection through edges** $\Theta : \mathbb{T}_L \to \mathbb{T}_L$ (associated to k and n) is defined exactly as in (10.1), but with $n \in \{\frac{1}{2}, \frac{3}{2}, \dots, \frac{L-1}{2}\}$. Now, Θ should be seen as a reflection of the torus through a plane Π with $\Pi \cap \mathbb{T}_L = \varnothing$, so that the corresponding decomposition of the torus, $\mathbb{T}_L = \mathbb{T}_{L,+} \cup \mathbb{T}_{L,-}$, is into two disjoint halves.

By definition, each transformation Θ is an involution: $\Theta^{-1} = \Theta$. Below, it will always be clear from the context whether the Θ under consideration is a reflection through vertices or edges.

10.3.1 Reflection positive measures

A reflection Θ can be made to act on spin configurations, $\Theta : \Omega_L \to \Omega_L$, by setting

$$(\Theta(\omega))_i \stackrel{\text{def}}{=} \omega_{\Theta(i)}, \quad \forall i \in \mathbb{T}_L.$$

Similarly, its action on functions $f : \Omega_L \to \mathbb{R}$ is defined by

$$\Theta(f)(\omega) \stackrel{\text{def}}{=} f(\Theta^{-1}(\omega)).$$

We denote by $\mathfrak{A}_+(\Theta)$, respectively $\mathfrak{A}_-(\Theta)$, the algebra of all bounded measurable functions f on Ω_L with support inside $\mathbb{T}_{L,+}(\Theta)$, respectively $\mathbb{T}_{L,-}(\Theta)$. The following properties will be used constantly in the sequel.

Exercise 10.1. *Check that, for any* $f, g \in \mathfrak{A}_+(\Theta)$ *and* $\lambda \in \mathbb{R}$,

$$\Theta^2(f) = f, \quad \Theta(\lambda f) = \lambda\Theta(f), \quad \Theta(f+g) = \Theta(f) + \Theta(g),$$
$$\Theta(fg) = \Theta(f)\Theta(g), \quad \Theta(e^f) = e^{\Theta(f)}.$$

Note that, in particular, the transformation Θ can be seen as an isomorphism between the algebras $\mathfrak{A}_+(\Theta)$ and $\mathfrak{A}_-(\Theta)$.

Since $\Theta : \Omega_L \to \Omega_L$ is clearly measurable, one can also define the action of Θ on a measure $\mu \in \mathscr{M}(\Omega_L, \mathscr{F}_L)$ by

$$\Theta(\mu)(A) \stackrel{\text{def}}{=} \mu(\Theta^{-1}A), \quad A \in \mathscr{F}_L.$$

Of course, this implies that, for every bounded measurable function f (remember that $\langle f\rangle_\mu \stackrel{\text{def}}{=} \int f\,\mathrm{d}\mu$),

$$\langle f\rangle_{\Theta(\mu)} = \langle \Theta(f)\rangle_\mu.$$

Definition 10.5. *Let* Θ *be a reflection. A measure* $\mu \in \mathscr{M}(\Omega_L, \mathscr{F}_L)$ *is* ***reflection positive with respect to*** $\boldsymbol{\Theta}$ *if*

1. $\langle f\,\Theta(g)\rangle_\mu = \langle g\,\Theta(f)\rangle_\mu$, *for all* $f, g \in \mathfrak{A}_+(\Theta)$;
2. $\langle f\,\Theta(f)\rangle_\mu \geq 0$, *for all* $f \in \mathfrak{A}_+(\Theta)$.

The set of measures which are reflection positive with respect to Θ *is denoted by* $\mathscr{M}_{\mathrm{RP}(\Theta)}$.

In other words, μ is reflection positive if and only if the bilinear form $(f,g) \mapsto \langle f\,\Theta(g)\rangle_\mu$ on $\mathfrak{A}_+(\Theta)$ is symmetric and positive semi-definite. This immediately implies the validity of a Cauchy–Schwarz-type inequality, which will be the basis of the properties to be derived later:

Lemma 10.6. *Let* $\mu \in \mathscr{M}_{\mathrm{RP}(\Theta)}$. *Then, for all* $f, g \in \mathfrak{A}_+(\Theta)$,

$$\langle f\Theta(g)\rangle_\mu^2 \leq \langle f\Theta(f)\rangle_\mu \langle g\Theta(g)\rangle_\mu.$$

Proof. Let $\mu \in \mathscr{M}_{\mathrm{RP}(\Theta)}$. We have, for all $\lambda \in \mathbb{R}$,

$$0 \leq \langle(\lambda f + g)\,\Theta(\lambda f + g)\rangle_\mu = \langle f\,\Theta(f)\rangle_\mu\,\lambda^2 + 2\langle f\,\Theta(g)\rangle_\mu\,\lambda + \langle g\,\Theta(g)\rangle_\mu.$$

This implies that the latter quadratic polynomial in λ has at most one root and, therefore, the associated discriminant cannot be positive. The claim follows. □

As seen in the following exercise, the first condition in Definition 10.5 is equivalent to saying that μ is invariant under Θ; it is thus both natural and rather mild. It will always be trivially satisfied in the cases considered later.

Exercise 10.2. *Show that the first condition in Definition 10.5 holds if and only if*

$$\Theta(\mu) = \mu. \tag{10.2}$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

10.3.2 Examples of reflection positive measures

As a starting point, we consider product measures. Let ρ be a measure on $(\Omega_0, \mathscr{F}_0)$, which we will refer to as the **reference** measure, and let

$$\mu_0 \stackrel{\text{def}}{=} \bigotimes_{i \in \mathbb{T}_L} \rho \,. \tag{10.3}$$

Lemma 10.7. *μ_0 is reflection positive with respect to all reflections Θ.*

Proof. Notice that

$$\Theta(\mu_0) = \mu_0 \tag{10.4}$$

for each reflection Θ. Indeed, the measure of any rectangle $\times_{k=1}^{|\mathbb{T}_L|} B_k$ ($B_k \in \mathscr{F}_0$) is the same under μ_0 or $\Theta(\mu_0)$, since μ_0 is invariant under any relabeling of the vertices of $\mathbb{T}_L$. By Exercise 10.2, this implies that μ_0 satisfies the first condition of Definition 10.5; let us check the second one.

We first consider a reflection Θ through edges. Let $f \in \mathfrak{A}_+(\Theta)$. Since $\mathbb{T}_{L,+}(\Theta) \cap \mathbb{T}_{L,-}(\Theta) = \varnothing$, f and $\Theta(f)$ have disjoint supports and (10.4) yields

$$\langle f\,\Theta(f)\rangle_{\mu_0} = \langle f\rangle_{\mu_0} \langle \Theta(f)\rangle_{\mu_0} = (\langle f\rangle_{\mu_0})^2 \geq 0\,,$$

thus showing that $\mu_0 \in \mathscr{M}_{\mathrm{RP}(\Theta)}$.

Let us now assume that Θ is a reflection through vertices and, again, let us take $f \in \mathfrak{A}_+(\Theta)$. In this case, the supports of f and $\Theta(f)$ may intersect. Let therefore P be the set of all vertices of $\mathbb{T}_L$ belonging to the reflection plane and remember that $\mathscr{F}_P$ denotes the sigma-algebra generated by the spins attached to vertices in P. We then have

$$\begin{aligned}\langle f\Theta(f)\rangle_{\mu_0} = \big\langle \mu_0(f\,\Theta(f)\,|\,\mathscr{F}_P)\big\rangle_{\mu_0} &= \big\langle \mu_0(f\,|\,\mathscr{F}_P)\mu_0(\Theta(f)\,|\,\mathscr{F}_P)\big\rangle_{\mu_0}\\ &= \big\langle \mu_0(f\,|\,\mathscr{F}_P)^2\big\rangle_{\mu_0} \geq 0\,,\end{aligned}$$

and reflection positivity follows again. (In the second equality, we used the fact that $\mu_0(\cdot|\mathscr{F}_P)$ is again a product measure.) □

From now on, we let ρ denote some reference measure on $(\Omega_0, \mathscr{F}_0)$, which we assume to be compactly supported, with $\rho(\Omega_0) < \infty$. We define μ_0 as in (10.3). We can then define the Gibbs distribution on $(\Omega_L, \mathscr{F}_L)$, associated to a Hamiltonian $\mathscr{H}_L : \Omega_L \to \mathbb{R}$, by

$$\forall A \in \mathscr{F}_L, \quad \mu_L(A) \stackrel{\text{def}}{=} \int_{\Omega_L} \frac{e^{-\mathscr{H}_L(\omega)}}{\mathbf{Z}_L} 1_A(\omega)\,\mu_0(\mathrm{d}\omega)\,, \tag{10.5}$$

where

$$\mathbf{Z}_L = \int_{\Omega_L} e^{-\mathscr{H}_L(\omega)} \mu_0(\mathrm{d}\omega) = \langle e^{-\mathscr{H}_L}\rangle_{\mu_0}\,.$$

(Of course, for this definition to make sense, we must have $\mathbf{Z}_L < \infty$. This will always be the case below.)

Lemma 10.8. *Let μ_L be as above. Let Θ be a reflection on $\mathbb{T}_L$ and assume that the Hamiltonian can be written as*

$$-\mathscr{H}_L = A + \Theta(A) + \sum_\alpha C_\alpha \Theta(C_\alpha)\,, \tag{10.6}$$

for some functions $A, C_\alpha \in \mathfrak{A}_+(\Theta)$. Then $\mu_L \in \mathscr{M}_{\mathrm{RP}(\Theta)}$.

Proof. Using a Taylor expansion for the factor $\exp\big(\sum_\alpha C_\alpha \Theta(C_\alpha)\big)$,

$$\begin{aligned}\langle f\,\Theta(g)\rangle_{\mu_L} &= \frac{1}{\mathbf{Z}_L}\big\langle f\,\Theta(g)\,e^{A+\Theta(A)+\sum_\alpha C_\alpha\Theta(C_\alpha)}\big\rangle_{\mu_0}\\ &= \frac{1}{\mathbf{Z}_L}\sum_{n\geq 0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n}\big\langle f\,e^A C_{\alpha_1}\cdots C_{\alpha_n}\,\Theta(g\,e^A C_{\alpha_1}\cdots C_{\alpha_n})\big\rangle_{\mu_0}.\end{aligned}$$

The result now follows from Lemma 10.7, since $\mu_0 \in \mathscr{M}_{\mathrm{RP}(\Theta)}$. □

As usual, the Hamiltonian can be constructed from a potential $\Phi = \{\Phi_B\}$:

$$\mathscr{H}_L \stackrel{\text{def}}{=} \sum_{B\subset\mathbb{T}_L} \Phi_B\,,$$

where Φ_B is a measurable function with support in B. To ensure that $\mathscr{H}_L$ can be put in the form (10.6), some symmetry assumptions will be made about the functions Φ_B.

Example 10.9. Consider a translation-invariant potential $\{\Phi_B\}_{B\subset\mathbb{T}_L}$ involving interactions only between pairs (that is, satisfying $\Phi_B = 0$ whenever $|B| \neq 2$) and such that $\Phi_{\{i,j\}} = 0$ whenever $\|j-i\|_\infty > 1$. Assume Θ is a reflection through the *vertices* of $\mathbb{T}_L$ satisfying

$$\Phi_{\{i,j\}}(\omega) = \Phi_{\{\Theta(i),\Theta(j)\}}(\Theta(\omega))\,,\quad \forall\omega\,, \tag{10.7}$$

for all $\{i,j\}\subset\mathbb{T}_L$. This holds, for example, if $\Phi_{\{i,j\}}$ depends only on the distance $\|j-i\|_1$.

Let us show that $\mu_L \in \mathscr{M}_{\mathrm{RP}(\Theta)}$. Namely, let again P denote the set of vertices of $\mathbb{T}_L$ lying in the reflection plane of Θ. Notice that, since the only pairs $e=\{i,j\}$ to be considered involve points with $\|j-i\|_\infty \leq 1$, the Hamiltonian can be written as

$$\mathscr{H}_L = \sum_{e\subset\mathbb{T}_L}\Phi_e = \sum_{e\subset P}\Phi_e + \sum_{\substack{e\subset\mathbb{T}_{L,+}\\ e\not\subset P}}\Phi_e + \sum_{\substack{e\subset\mathbb{T}_{L,-}\\ e\not\subset P}}\Phi_e\,.$$

Each pair $e=\{i,j\}\subset\mathbb{T}_{L,-}$ can be paired with its reflection $\Theta(e) = \{\Theta(i),\Theta(j)\}\subset\mathbb{T}_{L,+}$. Therefore, a change of variables yields, using (10.7),

$$\sum_{\substack{e\subset\mathbb{T}_{L,-}\\ e\not\subset P}}\Phi_e(\omega) = \sum_{\substack{e\subset\mathbb{T}_{L,+}\\ e\not\subset P}}\Phi_{\Theta(e)}(\omega) = \sum_{\substack{e\subset\mathbb{T}_{L,+}\\ e\not\subset P}}\Phi_e(\Theta(\omega))\,.$$

This means that $-\mathscr{H}_L = A + \Theta(A)$, with $A\in\mathfrak{A}_+(\Theta)$ given by

$$A \stackrel{\text{def}}{=} -\tfrac{1}{2}\sum_{e\subset P}\Phi_e - \sum_{\substack{e\subset\mathbb{T}_{L,+}\\ e\not\subset P}}\Phi_e\,,$$

Lemma 10.8 now implies that $\mu_L\in\mathscr{M}_{\mathrm{RP}(\Theta)}$. ⋄

Example 10.10. Let $\Omega_0 = \mathbb{R}^\nu$ and ρ be compactly supported. We assume that, for each $1\leq m\leq\nu$ and each $1\leq k\leq d$, J_k^m is a fixed nonnegative number. We consider a Hamiltonian of the form

$$\mathscr{H}_L \stackrel{\text{def}}{=} -\sum_{\{i,j\}\in\mathscr{E}_L}\sum_{m=1}^{\nu} J_{i,j}^m\,S_i^m S_j^m\,, \tag{10.8}$$

where $J^m_{i,j} = J^m_k$ when i and j differ in their kth component and S^m_i is the mth component of $\mathbf{S}_i$. This Hamiltonian actually covers all the applications we are going to consider in this chapter.

Let Θ be a reflection through edges of the torus. Proceeding similarly to what we did in Example 10.9, it is easy to check that

$$-\mathscr{H}_L = A + \Theta(A) + \sum_{m=1}^{\nu} \sum_{\substack{i \in \mathbb{T}_{L,+}:\\ \{i,\Theta(i)\} \in \mathscr{E}_L}} C^m_i \, \Theta(C^m_i),$$

where the functions $A, C^m_i \in \mathfrak{A}_+(\Theta)$ are given by

$$A \stackrel{\text{def}}{=} \sum_{m=1}^{\nu} \sum_{\substack{\{i,j\} \in \mathscr{E}_L:\\ i,j \in \mathbb{T}_{L,+}}} J^m_{i,j} S^m_i S^m_j \quad \text{and} \quad C^m_i \stackrel{\text{def}}{=} \sqrt{J^m_{i,\Theta(i)}}\, S^m_i .$$

Lemma 10.8 implies again that $\mu_L \in \mathscr{M}_{\mathrm{RP}(\Theta)}$. ◇

Exercise 10.3. *Give an example of a translation invariant measure $\mu \in \mathscr{M}(\Omega_L, \mathscr{F}_L)$ which is not reflection positive.*

10.4 The chessboard estimate

In this section, we establish a first major consequence of reflection positivity, the chessboard estimate and provide two applications.

10.4.1 Proof of the estimate

To simplify the exposition, we shall focus on the case of reflections through edges; however, both the statement and the proof can be adapted straightforwardly to the case of reflections through vertices.

Let $B < L$ be two positive integers such that $2B$ divides L and let us define $\Lambda_B \stackrel{\text{def}}{=} \{0, \dots, B-1\}^d \subset \mathbb{T}_L$. We decompose the torus into a disjoint union of translates of Λ_B, called **blocks**. These can be indexed by $t \in \mathbb{T}_{L/B}$:

$$\mathbb{T}_L = \bigcup_{t \in \mathbb{T}_{L/B}} (\Lambda_B + Bt).$$

A function f with support inside Λ_B is said to be **Λ_B-local**. Given a Λ_B-local function f and $t \in \mathbb{T}_{L/B}$, we define a $(\Lambda_B + tB)$-local function $f^{[t]}$ by successive reflections: Let $t_0 = 0, t_1, \dots, t_k = t$ be a self-avoiding nearest-neighbor path in $\mathbb{T}_{L/B}$ and let Θ_i be the reflection through the plane going through the edges connecting $\Lambda_B + t_{i-1}B$ and $\Lambda_B + t_i B$; we set

$$f^{[t]} \stackrel{\text{def}}{=} \Theta_k \circ \Theta_{k-1} \circ \cdots \circ \Theta_1(f).$$

A glance at Figure 10.3 shows that the definition of $f^{[t]}$ does not depend on the chosen path (observe that this relies on L/B being even).

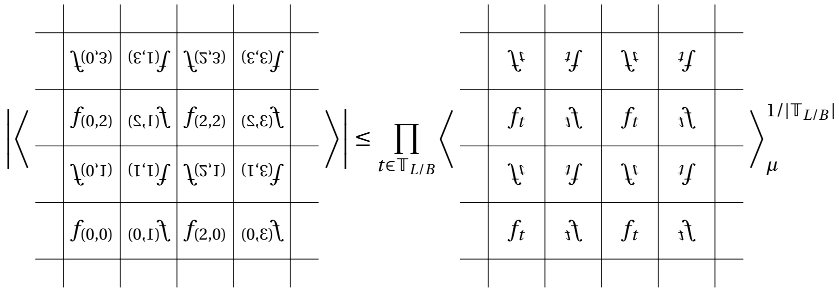

Figure 10.4: In $d = 2$, a graphical evocation of the claim of the chessboard estimate.

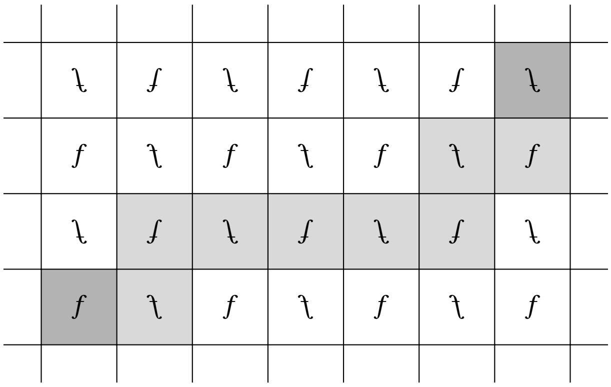

Figure 10.3: A graphical evocation of the definition of $f^{[t]}$; it is obtained by applying reflections to the original function f (located in the bottom left block Λ_B), until reaching the block indexed by t (top right on the picture). The definition of $f^{[t]}$ is independent of the chosen path (shaded cells).

Let us say that $\mu \in \mathscr{M}(\Omega_L, \mathscr{F}_L)$ is ***B*-periodic** if it is invariant under translations by B along any coordinate axis: $\mu = \mu \circ \theta_{B\mathbf{e}_k}$ for all $k \in \{1, \dots, d\}$.

Theorem 10.11 (Chessboard estimate)**.** *Let $\mu \in \mathscr{M}(\Omega_L, \mathscr{F}_L)$ be B-periodic and such that $\mu \in \mathscr{M}_{\mathrm{RP}(\Theta)}$ for all reflections Θ between neighboring blocks (that is, pairs $\Lambda_B + tB$, $\Lambda_B + t'B$, where t and t' are nearest neighbors of $\mathbb{T}_{L/B}$). Then (see Figure 10.4), for any family $(f_t)_{t \in \mathbb{T}_{L/B}}$ Λ_B-local functions, which are either all bounded or all nonnegative,*

$$\Big|\Big\langle \prod_{t \in \mathbb{T}_{L/B}} f_t^{[t]} \Big\rangle_\mu\Big| \le \prod_{t \in \mathbb{T}_{L/B}} \Big[\Big\langle \prod_{s \in \mathbb{T}_{L/B}} f_t^{[s]} \Big\rangle_\mu\Big]^{1/|\mathbb{T}_{L/B}|}. \tag{10.9}$$

Proof of Theorem 10.11: We can assume that the functions f_t are bounded. Indeed, if they are unbounded (but nonnegative), we can apply the result to the bounded functions $f_t \wedge K$ ($K \in \mathbb{N}$) and use monotone convergence to take the limit $K \uparrow \infty$.

The proof is done by induction on the dimension.

The case $d = 1$: In the one-dimensional case, the boxes are simply intervals indexed by $t \in \{0,1,\dots,2N-1\}$, where $N = L/(2B) \in \mathbb{N}$. Observe that, in this case, given a Λ_B-local function f, each function $f^{[t]}$ coincides either with a translate of f or with a translate of the Λ_B-local function defined by

$$\bar{f}(\omega_0,\omega_1,\dots,\omega_{B-1},\omega_B,\dots,\omega_{L-1}) \stackrel{\text{def}}{=} f(\omega_{B-1},\omega_{B-2},\dots,\omega_0,\omega_{L-1},\dots,\omega_B)\,.$$

Consider the following multilinear functional on the $2N$-tuples of Λ_B-local functions:

$$F(f_0,\dots,f_{2N-1}) \stackrel{\text{def}}{=} \Big\langle \prod_{t=0}^{2N-1} f_t^{[t]} \Big\rangle_\mu .$$

Reformulated in terms of F, the chessboard estimate (10.9) that we want to establish can be expressed as

$$\big|F(f_0,\dots,f_{2N-1})\big| \le \prod_{t=0}^{2N-1} F(f_t,\dots,f_t)^{1/2N}. \tag{10.10}$$

Observe that each of the expectations in the right-hand side of (10.10) is nonnegative. Indeed, we can write

$$F(f_t,\dots,f_t) = \Big\langle \big(\prod_{t=0}^{N-1} f_t^{[t]}\big)\Theta\big(\prod_{t=0}^{N-1} f_t^{[t]}\big)\Big\rangle_\mu ,$$

where Θ is the reflection through the edge between the blocks $N-1$ and N (and $2N-1$ and 0). This implies, in particular, that (10.10) trivially holds whenever $F(f_0,\dots,f_{2N-1}) = 0$. We will thus assume from now on that $(f_0,\dots,f_{2N-1})$ is fixed and that

$$F(f_0,\dots,f_{2N-1}) \neq 0\,. \tag{10.11}$$

We start with two fundamental properties of F.

Lemma 10.12. *For all Λ_B-local functions $f_0,\dots,f_{2N-1}$,*

$$F(f_0,f_1,\dots,f_{2N-1}) = F(\bar{f}_{2N-1},\bar{f}_0,\bar{f}_1,\dots,\bar{f}_{2N-2}) \tag{10.12}$$

and

$$\begin{aligned} &F(f_0,\dots,f_{N-1},f_N,\dots,f_{2N-1})^2 \\ &\qquad \le F(f_0,\dots,f_{N-1},f_{N-1},\dots,f_0)F(f_{2N-1},\dots,f_N,f_N,\dots,f_{2N-1})\,. \end{aligned} \tag{10.13}$$

Exercise 10.4. *Show that, in general, $F(f_0,\dots,f_{2N-1}) \neq F(\bar{f}_0,\dots,\bar{f}_{2N-1})$.*

Proof. The first identity is a simple consequence of the B-periodicity of μ and of the definition of F. To prove the second one, let again Θ denote the reflection through the edge between the blocks $N-1$ and N (and $2N-1$ and 0). Observe that, for each $N \le t \le 2N-1$, $f_t^{[t]} = \Theta(f_{2N-1-t'}^{[t']})$, where $t' = 2N-1-t \in \{0,\dots,N-1\}$. Therefore, by Lemma 10.6,

$$\begin{aligned} F(f_0,\dots,f_{N-1},f_N,\dots,f_{2N-1})^2 &= \Big\langle \big(\prod_{t=0}^{N-1} f_t^{[t]}\big)\Theta\big(\prod_{t=0}^{N-1} f_{2N-1-t}^{[t]}\big)\Big\rangle_\mu^2 \\ &\le F(f_0,\dots,f_{N-1},f_{N-1},\dots,f_0)F(f_{2N-1},\dots,f_N,f_N,\dots,f_{2N-1})\,. \quad \square \end{aligned}$$

When $2N$ is a power of 2, repeated use of the above lemma leads directly to (10.10)*. For simplicity, assume first that $2N = 4$. In this case,* (10.13) *yields*

$$F(f_0, f_1, f_2, f_3)^4 \le F(f_0, f_1, f_1, f_0)^2 F(f_3, f_2, f_2, f_3)^2 .$$

Now, by (10.12)*, $F(f_0, f_1, f_1, f_0) = F(\bar f_0, \bar f_0, \bar f_1, \bar f_1)$, $F(f_3, f_2, f_2, f_3) = F(\bar f_2, \bar f_2, \bar f_3, \bar f_3)$. But, using again* (10.13) *twice,*

$$F(\bar f_0, \bar f_0, \bar f_1, \bar f_1)^2 F(\bar f_2, \bar f_2, \bar f_3, \bar f_3)^2 \le \prod_{t=0}^{3} F(\bar f_t, \bar f_t, \bar f_t, \bar f_t) .$$

This implies (10.10)*, since $F(\bar f_t, \bar f_t, \bar f_t, \bar f_t) = F(f_t, f_t, f_t, f_t)$ by* (10.12)*. Clearly, if $2N = 2^M$, the same argument can be used repeatedly. The proof of* (10.10) *for general values of N relies on a variant of this argument, as we explain below.* ◇

Let us consider the auxiliary functional

$$G(f_0, \dots, f_{2N-1}) \stackrel{\text{def}}{=} \frac{|F(f_0, \dots, f_{2N-1})|}{\prod_{t=0}^{2N-1} F(f_t, \dots, f_t)^{1/2N}} ,$$

which is well defined thanks to the following property.

Lemma 10.13. *For each $t \in \{0, 1, \dots, 2N-1\}$, $F(f_t, \dots, f_t) > 0$.*

Proof. For the sake of readability, we treat explicitly only the case $2N = 6$. The extension to general values of $2N$ is straightforward, as explained below. Let $K_N \stackrel{\text{def}}{=} (\max_t \|f_t\|_\infty)^{2N}$. Applying (10.13), we get

$$\begin{aligned}|F(f_0, f_1, f_2, f_3, f_4, f_5)| &\le F(f_0, f_1, f_2, f_2, f_1, f_0)^{1/2} F(f_5, f_4, f_3, f_3, f_4, f_5)^{1/2} \\ &\le K_3^{1/2} F(f_0, f_1, f_2, f_2, f_1, f_0)^{1/2} .\end{aligned}$$

We now apply (10.12) in order to push the two copies of f_0 in the first two slots:

$$K_3^{1/2} F(f_0, f_1, f_2, f_2, f_1, f_0)^{1/2} = K_3^{1/2} F(\bar f_0, \bar f_0, \bar f_1, \bar f_2, \bar f_2, \bar f_1)^{1/2} .$$

Using (10.13) once more, we obtain

$$K_3^{1/2} F(\bar f_0, \bar f_0, \bar f_1, \bar f_2, \bar f_2, \bar f_1)^{1/2} \le K_3^{3/4} F(\bar f_0, \bar f_0, \bar f_1, \bar f_1, \bar f_0, \bar f_0)^{1/4} .$$

Again, (10.12) allows us to push the four copies of $\bar f_0$ in the first four slots:

$$K_3^{3/4} F(\bar f_0, \bar f_0, \bar f_1, \bar f_1, \bar f_0, \bar f_0)^{1/4} = K_3^{3/4} F(\bar f_0, \bar f_0, \bar f_0, \bar f_0, \bar f_1, \bar f_1)^{1/4} .$$

Applying (10.13) one last time yields

$$K_3^{3/4} F(\bar f_0, \bar f_0, \bar f_0, \bar f_0, \bar f_1, \bar f_1)^{1/4} \le K_3^{7/8} F(\bar f_0, \bar f_0, \bar f_0, \bar f_0, \bar f_0, \bar f_0)^{1/8}$$

and thus, since $F(f_0, f_0, f_0, f_0, f_0, f_0) = F(\bar f_0, \bar f_0, \bar f_0, \bar f_0, \bar f_0, \bar f_0)$ by (10.12),

$$F(f_0, f_0, f_0, f_0, f_0, f_0) \ge K_3^{-7} |F(f_0, f_1, f_2, f_3, f_4, f_5)|^8 > 0 .$$

(We used our assumption (10.11).) General values of $2N$ are treated in exactly the same way, applying (10.12) and (10.13) alternatively until all $2N$ slots of F are filled

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

by copies of f_0. Since the number of such copies doubles at each stage, the number of required iterations is given by the smallest integer M such that $2^M \ge 2N$, which yields

$$F(f_0,\dots,f_0) \ge K_N^{1-2^M} |F(f_0,\dots,f_{2N-1})|^{2^M} > 0\,.$$

The same argument applies to other values of t using (10.12). □

By construction, G verifies the same properties as those satisfied by F in (10.12) and (10.13). Moreover, $G(f_t,\dots,f_t) = 1$ for all t. In terms of G, we will obtain (10.10) by showing that $G(f_0,\dots,f_{2N-1}) \le 1$, which is equivalent to saying that G reaches its maximum value on $2N$-tuples of functions (from the collection $\{f_0, f_1,\dots,f_{2N-1}\}$) which are composed of a single function f_t.

Let $(g_0,\dots,g_{2N-1})$ be such that

(i) $g_i \in \{f_0,\dots,f_{2N-1}\}$ for each $i \in \{0,\dots,2N-1\}$;

(ii) $(g_0,\dots,g_{2N-1})$ maximizes G;

(iii) $(g_0,\dots,g_{2N-1})$ is minimal, in the sense that it contains the longest contiguous substring of the form $f_i,\dots,f_i$ for some $i \in \{0,\dots,2N-1\}$. Here g_{2N-1} and g_0 are considered contiguous (because of property 10.12).

Let k be the length of the substring in (iii). Thanks to (10.12), we can assume that the latter occurs at the beginning of the string $(g_0,\dots,g_{2N-1})$, that is, that $g_0 = g_1 = \cdots = g_{k-1} = f_i$ (or $\bar f_i$, with bars on each of the $2N$ entries). We shall now check that $k = 2N$, which will conclude the proof of the one-dimensional case.

Suppose that $k < 2N$. We have

$$\begin{aligned} G(g_0,\dots,g_{2N-1})^2 &\le G(g_0,\dots,g_{N-1},g_{N-1},\dots,g_0)G(g_{2N-1},\dots,g_N,g_N,\dots,g_{2N-1})\\ &\le G(g_0,\dots,g_{N-1},g_{N-1},\dots,g_0)G(g_0,\dots,g_{2N-1})\,, \end{aligned}$$

since $(g_0,\dots,g_{2N-1})$ maximizes G. Therefore ($G(g_0,\dots,g_{2N-1}) > 0$ by (10.11)),

$$G(g_0,\dots,g_{2N-1}) \le G(g_0,\dots,g_{N-1},g_{N-1},\dots,g_0)\,,$$

which means that $(g_0,\dots,g_{N-1},g_{N-1},\dots,g_0)$ is also a maximizer of G. But this is impossible, since the string $(g_0,\dots,g_{N-1},g_{N-1},\dots,g_0)$ possesses a substring $f_i,\dots,f_i$ of length $\min\{2N,2k\} > k$, which would violate our minimality assumption (iii).

The case $d \ge 2$: We now assume that the chessboard estimate (10.9) has been established for all dimensions $d' \in \{1,\dots,d\}$ and show that it also holds in dimension $d+1$. Although this induction step is rather straightforward, it involves a few subtleties which we discuss after the proof, in Remark 10.14.

We temporarily denote the d-dimensional torus by $\mathbb{T}_L^d$ and consider $\mathbb{T}_L^{d+1}$ as L adjacent copies of $\mathbb{T}_L^d$:

$$\mathbb{T}_L^{d+1} = \mathbb{T}_L^1 \times \mathbb{T}_L^d\,.$$

We can thus write $u \in \mathbb{T}^{d+1}_{L/B}$ as $u = (i,t)$, with $i \in \mathbb{T}^1_{L/B}$ and $t \in \mathbb{T}^d_{L/B}$, and use the shorthand notation $f^{[u]} = f^{[(i,t)]} \equiv f^{[i,t]}$. Therefore, applying (10.9) with $d' = 1$,

$$\begin{aligned}\Big|\Big\langle \prod_{u\in\mathbb{T}^{d+1}_{L/B}} f_u^{[u]}\Big\rangle_\mu\Big| &= \Big|\Big\langle \prod_{i\in\mathbb{T}^1_{L/B}} \Big\{\prod_{t\in\mathbb{T}^d_{L/B}} f_{(i,t)}^{[t]}\Big\}^{[i]}\Big\rangle_\mu\Big| \\ &\leq \prod_{i\in\mathbb{T}^1_{L/B}} \Big[\Big\langle \prod_{j\in\mathbb{T}^1_{L/B}} \Big\{\prod_{t\in\mathbb{T}^d_{L/B}} f_{(i,t)}^{[t]}\Big\}^{[j]}\Big\rangle_\mu\Big]^{1/|\mathbb{T}^1_{L/B}|} \qquad (10.14) \\ &= \prod_{i\in\mathbb{T}^1_{L/B}} \Big[\Big\langle \prod_{t\in\mathbb{T}^d_{L/B}} \Big\{\prod_{j\in\mathbb{T}^1_{L/B}} f_{(i,t)}^{[j]}\Big\}^{[t]}\Big\rangle_\mu\Big]^{1/|\mathbb{T}^1_{L/B}|}. \qquad (10.15)\end{aligned}$$

The expectation in the right-hand side can be bounded using (10.9) once more, this time with $d' = d$: for each $i \in \mathbb{T}^1_{L/B}$,

$$\begin{aligned}\Big\langle \prod_{t\in\mathbb{T}^d_{L/B}} \Big\{\prod_{j\in\mathbb{T}^1_{L/B}} f_{(i,t)}^{[j]}\Big\}^{[t]}\Big\rangle_\mu &\leq \prod_{t\in\mathbb{T}^d_{L/B}} \Big[\Big\langle \prod_{s\in\mathbb{T}^d_{L/B}} \Big\{\prod_{j\in\mathbb{T}^1_{L/B}} f_{(i,t)}^{[j]}\Big\}^{[s]}\Big\rangle_\mu\Big]^{1/|\mathbb{T}^d_{L/B}|} \qquad (10.16) \\ &= \prod_{t\in\mathbb{T}^d_{L/B}} \Big[\Big\langle \prod_{v\in\mathbb{T}^{d+1}_{L/B}} f_{(i,t)}^{[v]}\Big\rangle_\mu\Big]^{1/|\mathbb{T}^d_{L/B}|}.\end{aligned}$$

Inserting the latter bound into (10.15),

$$\begin{aligned}\Big|\Big\langle \prod_{u\in\mathbb{T}^{d}_{L/B}} f_u^{[u]}\Big\rangle_\mu\Big| &\leq \prod_{i\in\mathbb{T}^1_{L/B}} \prod_{t\in\mathbb{T}^d_{L/B}} \Big[\Big\langle \prod_{v\in\mathbb{T}^{d+1}_{L/B}} f_{(i,t)}^{[v]}\Big\rangle_\mu\Big]^{1/|\mathbb{T}^{d+1}_{L/B}|} \\ &= \prod_{u\in\mathbb{T}^{d+1}_{L/B}} \Big[\Big\langle \prod_{v\in\mathbb{T}^{d+1}_{L/B}} f_u^{[v]}\Big\rangle_\mu\Big]^{1/|\mathbb{T}^{d+1}_{L/B}|}.\end{aligned}$$

This completes the proof of Theorem 10.11. □

Remark 10.14. Let us make a comment about what was done in the last part of the proof. The verification of certain claims made below is left as an exercise to the reader.

With $\mathbb{T}^{d+1}_L = \mathbb{T}^1_L \times \mathbb{T}^d_L$, we are naturally led to identify Ω_L, the set of configurations on $\mathbb{T}^{d+1}_L$, with the set of configurations on $\mathbb{T}^1_L$ defined by

$$\widetilde{\Omega}_L \stackrel{\text{def}}{=} \big\{\widetilde{\omega} = (\widetilde{\omega}_i)_{i\in\mathbb{T}^1_L} : \widetilde{\omega}_i \in \widetilde{\Omega}_0\big\},$$

where we introduced the new single-spin space

$$\widetilde{\Omega}_0 \stackrel{\text{def}}{=} \bigtimes_{j\in\mathbb{T}^d_L} \Omega_0 .$$

Let us denote this identification by $\phi : \Omega_L \to \widetilde{\Omega}_L$. Each $f : \Omega_L \to \mathbb{R}$ can be identified with $\widetilde{f} : \widetilde{\Omega}_L \to \mathbb{R}$, by $\widetilde{f}(\widetilde{\omega}) \stackrel{\text{def}}{=} f(\phi^{-1}(\widetilde{\omega}))$. The single-spin space $\widetilde{\Omega}_0$ can of course be equipped with its natural σ-algebra of Borel sets, leading to the product σ-algebra $\widetilde{\mathscr{F}}_L$ on $\widetilde{\Omega}_L$. The measure μ on $(\Omega_L, \mathscr{F}_L)$ can be identified with the measure $\widetilde{\mu}$ on $(\widetilde{\Omega}_L, \widetilde{\mathscr{F}}_L)$ defined by $\widetilde{\mu} \stackrel{\text{def}}{=} \mu \circ \phi^{-1}$. We then have

$$\langle f\rangle_\mu = \langle \widetilde{f}\rangle_{\widetilde{\mu}},$$

for every bounded measurable function f and, clearly, $\widetilde{\mu}$ is reflection positive with respect to all reflections of $\mathbb{T}_L^1$. This is what guarantees that the one-dimensional chessboard estimate can be used to prove (10.14). A similar argument justifies the second use of the chessboard estimate in (10.16). ⋄

Remark 10.15. We will actually make use of a version of Theorem 10.11 in which the cubic block Λ_B (and its translates) is replaced by a rectangular box $\times_{i=1}^d \{0,\dots,B_i\}$ (and its translates) such that $2B_i$ divides L for all i. Of course, the conditions of periodicity and reflection positivity have to be correspondingly modified, but the proof applies essentially verbatim. ⋄

10.4.2 Application: the Ising model in a large magnetic field

In this section, we show a use of the chessboard estimate in the simplest possible setting. A more involved application is described in the following sections.

We have studied the Ising model in a large magnetic field in Section 5.7.1. In particular, we obtained in (5.34) a convergent cluster expansion for the pressure of the model, in terms of $z = e^{-2\beta h}$. When $h > 0$, $\frac{\partial \psi_\beta}{\partial h} = \langle \sigma_0 \rangle^+_{\beta,h}$ and, therefore,

$$\mu^+_{\beta,h}(\sigma_0 = -1) = \tfrac{1}{2}\big(1 - \tfrac{\partial \psi_\beta}{\partial h}\big).$$

The expansion (5.34) thus implies that $\mu^+_{\beta,h}(\sigma_0 = -1) = e^{-2h-4d\beta} + O(e^{-4h})$ for $h > 0$ large enough. Here, we show how a simple application of the chessboard estimate leads to an upper bound for this probability (on the torus) valid for all $h, \beta \geq 0$.

For convenience, we write the Hamiltonian of the d-dimensional Ising model on $\mathbb{T}_L$ as

$$\mathscr{H}_{L;\beta,h}(\omega) \stackrel{\text{def}}{=} -\beta \sum_{\{i,j\}\in\mathscr{E}_L} (\omega_i\omega_j - 1) - h \sum_{i\in\mathbb{T}_L} \omega_i\,. \tag{10.17}$$

Let $\mu_{L;\beta,h}$ be the corresponding Gibbs distribution.

Proposition 10.16. *For all $h \geq 0$, uniformly in L (even) and $\beta \geq 0$,*

$$\mu_{L;\beta,h}(\sigma_0 = -1) \leq e^{-2h}. \tag{10.18}$$

Proof. The first observation is that $\mathscr{H}_{L;\beta,h}$ can be put in the form (10.8) (up to an irrelevant constant), from which we conclude that $\mu_{L;\beta,h}$ is reflection positive with respect to all reflections through edges.

Using 1×1 blocks (which we naturally identify with the vertices of $\mathbb{T}_L$) and setting $f_0 \stackrel{\text{def}}{=} \mathbf{1}_{\{\sigma_0=-1\}}$ and $f_t \stackrel{\text{def}}{=} 1$ for all $t \in \mathbb{T}_L \setminus \{0\}$, the chessboard estimates yields

$$\langle \mathbf{1}_{\{\sigma_0=-1\}} \rangle_{L;\beta,h} \leq \Big\langle \prod_{s\in\mathbb{T}_L} \mathbf{1}_{\{\sigma_s=-1\}} \Big\rangle^{1/|\mathbb{T}_L|}_{L;\beta,h}. \tag{10.19}$$

(Just observe that all the factors corresponding to $t \neq 0$ in the product in (10.9) are equal to 1.) This can be rewritten as

$$\mu_{L;\beta,h}(\sigma_0 = -1) \leq \mu_{L;\beta,h}(\eta^-)^{1/|\mathbb{T}_L|} = \Big\{ \frac{e^{-\mathscr{H}_{L;\beta,h}(\eta^-)}}{\mathbf{Z}_{L;\beta,h}} \Big\}^{1/|\mathbb{T}_L|},$$

where $\eta^-_j = -1$ for all $j \in \mathbb{T}_L$. On the one hand, $\mathscr{H}_{L;\beta,h}(\eta^-) = h|\mathbb{T}_L|$. On the other hand, we obtain a lower bound on the partition function by keeping only the configuration $\eta^+ \equiv 1$: $\mathbf{Z}_{L;\beta,h} \geq e^{-\mathscr{H}_{L;\beta,h}(\eta^+)} = e^{+h|\mathbb{T}_L|}$. This proves (10.18). □

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

In probabilistic terms, (10.19) *shows how the chessboard estimate allows us to bound the probability of a* local *event, namely* $\{\sigma_0 = -1\}$*, by the probability of the same event, but "spread out throughout the system":* $\bigcap_{s\in\mathbb{T}_L}\{\sigma_s = -1\}$*. This* global *event is much easier to estimate.* ⋄

10.4.3 Application: the two-dimensional anisotropic XY model

We now consider the two-dimensional anisotropic XY model, in which the spins take values in $\Omega_0 = \mathbb{S}^1$ and whose Hamiltonian on $\mathbb{T}_L$ is defined by

$$\mathscr{H}_{L;\beta,\alpha} \stackrel{\text{def}}{=} -\beta \sum_{\{i,j\}\in\mathscr{E}_L} \bigl\{S_i^1 S_j^1 + \alpha S_i^2 S_j^2\bigr\}, \tag{10.20}$$

where $0 \le \alpha \le 1$ is the **anisotropy parameter** and we have written $\mathbf{S}_i = (S_i^1, S_i^2)$ for the spin at i. We denote by $\mu_{L;\beta,\alpha}$ the corresponding Gibbs distribution on Ω_L (see (10.5)), with reference measure ρ on Ω_0 given by the normalized Lebesgue measure (that is, such that $\rho(\Omega_0) = 1$).

To quantify global ordering, we will again use the magnetization density :

$$\mathbf{m}_L \stackrel{\text{def}}{=} \frac{1}{|\mathbb{T}_L|} \sum_{i\in\mathbb{T}_L} \mathbf{S}_i\,,$$

which now takes values in the unit disk $\bigl\{u \in \mathbb{R}^2 : \|u\|_2 \le 1\bigr\}$. By translation invariance and symmetry,

$$\langle \mathbf{m}_L \rangle_{L;\beta,\alpha} = \langle \mathbf{S}_0 \rangle_{L;\beta,\alpha} = \mathbf{0}\,.$$

Nevertheless, we will see that, the distribution of $\mathbf{m}_L$ is far from uniform at low temperatures, when $\alpha < 1$. This, in turn, will lead to the proof of Theorem 10.2.

First, as the following exercise shows, when $\alpha < 1$, this model possesses exactly two ground states: one in which all spins take the value $\mathbf{e}_1$ and one in which this value is $-\mathbf{e}_1$.

Exercise 10.5. *Let* $\mathbf{S}_i = (S_i^1, S_i^2)$ *and* $\mathbf{S}_j = (S_j^1, S_j^2)$ *be two unit vectors in* $\mathbb{R}^2$*. Show that, when* $0 \le \alpha < 1$*, the function*

$$f(\mathbf{S}_i, \mathbf{S}_j) \stackrel{\text{def}}{=} -S_i^1 S_j^1 - \alpha S_i^2 S_j^2\,,$$

is minimal when either $\mathbf{S}_i = \mathbf{S}_j = \mathbf{e}_1$ *or* $\mathbf{S}_i = \mathbf{S}_j = -\mathbf{e}_1$*.*

In view of this, it is reasonable to expect that, at sufficiently low temperature, typical configurations should be given by local perturbations of these two ground states, even in the thermodynamic limit. This is confirmed by the following result.

Theorem 10.17. *For each* $0 \le \alpha < 1$ *and each* $\epsilon > 0$*, there exists* $\beta_0 = \beta_0(\alpha, \epsilon)$ *such that, for all* $\beta > \beta_0$*,*

$$\bigl\langle (\mathbf{m}_L \cdot \mathbf{e}_1)^2 \bigr\rangle_{L;\beta,\alpha} \ge 1 - \epsilon\,, \qquad \textit{and therefore} \qquad \bigl\langle (\mathbf{m}_L \cdot \mathbf{e}_2)^2 \bigr\rangle_{L;\beta,\alpha} \le \epsilon\,,$$

uniformly in L (multiple of 4*).*

This result is in sharp contrast with the case $\alpha = 1$ (see Exercise 9.1); it will be a consequence of the orientational long-range order that occurs at low enough temperatures. Observe that

$$\big\langle (\mathbf{m}_L \cdot \mathbf{e}_1)^2 \big\rangle_{L;\beta,\alpha} = \frac{1}{|\mathbb{T}_L|^2} \sum_{i,j\in\mathbb{T}_L} \langle S_i^1 S_j^1 \rangle_{L;\beta,\alpha} .$$

We will use reflection positivity to prove the following result, of which Theorem 10.17 is a direct consequence.

Proposition 10.18. *For each* $0 \le \alpha < 1$ *and each* $\epsilon > 0$*, there exists* $\beta_0 = \beta_0(\alpha,\epsilon)$ *such that, for all* $\beta > \beta_0$*,*

$$\langle S_i^1 S_j^1 \rangle_{L;\beta,\alpha} \ge 1 - \epsilon, \qquad \forall i,j \in \mathbb{T}_L, \tag{10.21}$$

uniformly in L (multiple of 4).

Proof. First, we easily check that $\mathscr{H}_{L;\beta,\alpha}$ can be put in the form (10.8), from which we conclude that $\mu_{L;\beta,\alpha}$ is reflection positive with respect to all reflections through edges of the torus.

Second, since $\mu_{L;\beta,\alpha}$ is invariant under all translations of the torus, we only need to prove that

$$\langle S_0^1 S_j^1 \rangle_{L;\beta,\alpha} \ge 1 - \epsilon(\beta), \tag{10.22}$$

uniformly in L and in $j \in \mathbb{T}_L^2$, with $\epsilon(\beta) \to 0$ when $\beta \to \infty$.

Now, in view of the discussion before Theorem 10.17, we expect that $|S_i^1|$ should be close to 1 for most spins in the torus and that the sign of S_i^1 should be the same at most vertices. To quantify this, let us fix some $\delta \in (0,1)$. If (i) $|S_0^1| \ge \delta$, (ii) $|S_j^1| \ge \delta$ and (iii) $S_0^1 S_j^1 > 0$, then $S_0^1 S_j^1 \ge \delta^2$. Therefore, we can write

$$\langle S_0^1 S_j^1 \rangle_{L;\beta,\alpha} \ge \delta^2 - \mu_{L;\beta,\alpha}(|S_0^1| < \delta) - \mu_{L;\beta,\alpha}(|S_j^1| < \delta) - \mu_{L;\beta,\alpha}(S_0^1 S_j^1 \le 0). \tag{10.23}$$

Since $\mu_{L;\beta,\alpha}(|S_j^1| < \delta) = \mu_{L;\beta,\alpha}(|S_0^1| < \delta)$ by translation invariance, the claim of Proposition 10.18 follows immediately from Lemmas 10.19 and 10.20 below: choose $\delta^2 = 1 - \frac{1}{4}\epsilon$ and let β be sufficiently large to ensure that the last three terms in (10.23) are smaller than $\epsilon/4$. □

Lemma 10.19. *For any* $0 \le \alpha < 1$, $0 < \delta < 1$ *and* $\epsilon > 0$*, there exists* $\beta_0' = \beta_0'(\epsilon,\alpha,\delta)$ *such that, for all* $\beta > \beta_0'$*,*

$$\mu_{L;\beta,\alpha}(|S_0^1| < \delta) \le \epsilon,$$

uniformly in L (even).

Lemma 10.20. *For any* $0 \le \alpha < 1$, $0 < \delta < 1$ *and* $\epsilon > 0$*, there exists* $\beta_0'' = \beta_0''(\epsilon,\alpha)$ *such that, for all* $\beta > \beta_0''$*,*

$$\mu_{L;\beta,\alpha}(S_0^1 S_j^1 \le 0) \le \epsilon,$$

uniformly in $j \in \mathbb{T}_L$ *and L (multiple of 4).*

Proof of Lemma 10.19. We proceed as in the proof of Proposition 10.16. Applying the chessboard estimate, Theorem 10.11, with $d = 2$, $B = 1$, $f_0 \stackrel{\text{def}}{=} \mathbf{1}_{\{|S_0^1| < \delta\}}$ and $f_t \stackrel{\text{def}}{=} 1$ for $t \in \mathbb{T}_L \setminus \{0\}$, we obtain

$$\mu_{L;\beta,\alpha}(|S_0^1| < \delta) \le \mu_{L;\beta,\alpha}(|S_i^1| < \delta,\ \forall i \in \mathbb{T}_L)^{1/|\mathbb{T}_L|}. \tag{10.24}$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

We write

$$\mu_{L;\beta,\alpha}(|S_i^1|<\delta,\ \forall i\in\mathbb{T}_L)=\frac{\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\mathbf{1}_{\{|S_i^1|<\delta,\ \forall i\in\mathbb{T}_L\}}\big\rangle_{\mu_0}}{\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\rangle_{\mu_0}}, \tag{10.25}$$

where we remind the reader that $\mu_0(\mathrm{d}\omega)=\bigotimes_{i\in\mathbb{T}_L}\rho(\mathrm{d}\omega_i)$, with ρ the uniform probability measure on $\mathbb{S}^1$.

We first bound the numerator in (10.25) from above. When $|S_i^1|<\delta$ for all $i\in\mathbb{T}_L$, a simple computation shows that

$$\mathscr{H}_{L;\beta,\alpha}\geq-\beta\sum_{\{i,j\}\in\mathscr{E}_L}\big(\delta^2+\alpha(1-\delta^2)\big)=-2\beta\big(\delta^2+\alpha(1-\delta^2)\big)|\mathbb{T}_L|.$$

(We used the fact that $|\mathscr{E}_L|=2|\mathbb{T}_L|$ in $d=2$.) Consequently, since μ_0 is normalized by assumption,

$$\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\mathbf{1}_{\{|S_i^1|<\delta,\ \forall i\in\mathbb{T}_L\}}\big\rangle_{\mu_0}\leq e^{2\beta(\delta^2+\alpha(1-\delta^2))|\mathbb{T}_L|}. \tag{10.26}$$

To obtain a lower bound on the denominator, let $0<\tilde\delta<1$ and write

$$\begin{aligned}\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\rangle_{\mu_0}&\geq\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\mathbf{1}_{\{S_i^1\geq\tilde\delta,\ \forall i\in\mathbb{T}_L\}}\big\rangle_{\mu_0}\\&=\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\ \big|\ S_i^1\geq\tilde\delta,\ \forall i\in\mathbb{T}_L\big\rangle_{\mu_0}\mu_0(S_i^1\geq\tilde\delta,\ \forall i\in\mathbb{T}_L)\\&=\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\big\rangle_{\tilde\mu_0}\mu_0(S_i^1\geq\tilde\delta,\ \forall i\in\mathbb{T}_L),\end{aligned} \tag{10.27}$$

where we have introduced the probability measure $\tilde\mu_0(\cdot)\stackrel{\text{def}}{=}\mu(\cdot\,|\,S_i^1\geq\tilde\delta,\ \forall i\in\mathbb{T}_L)$. On the one hand, observe that $\langle S_i^2\rangle_{\tilde\mu_0}=0$, by symmetry, and $\langle S_i^1\rangle_{\tilde\mu_0}\geq\tilde\delta$. Therefore,

$$\langle\mathscr{H}_{L;\beta,\alpha}\rangle_{\tilde\mu_0}=-\beta\sum_{\{i,j\}\in\mathscr{E}_L}\langle S_i^1\rangle_{\tilde\mu_0}\langle S_j^1\rangle_{\tilde\mu_0}\leq-2\beta\tilde\delta^2|\mathbb{T}_L|.$$

So, an application of Jensen's inequality yields

$$\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\rangle_{\tilde\mu_0}\geq e^{-\langle\mathscr{H}_{L;\beta,\alpha}\rangle_{\tilde\mu_0}}\geq e^{2\beta\tilde\delta^2|\mathbb{T}_L|}. \tag{10.28}$$

On the other hand,

$$\mu_0(S_i^1\geq\tilde\delta,\ \forall i\in\mathbb{T}_L)=\big(\tfrac{1}{\pi}\arccos(\tilde\delta)\big)^{|\mathbb{T}_L|}=e^{-b(\tilde\delta)|\mathbb{T}_L|},$$

where $b(\tilde\delta)\stackrel{\text{def}}{=}-\log(\frac{1}{\pi}\arccos(\tilde\delta))>0$. Inserting this and (10.28) into (10.27) yields

$$\langle e^{-\mathscr{H}_{L;\beta,\alpha}}\rangle_{\mu_0}\geq\exp\big\{\big(2\beta\tilde\delta^2-b(\tilde\delta)\big)|\mathbb{T}_L|\big\}. \tag{10.29}$$

Let us then choose $\tilde\delta$ such that $\tilde\delta^2=\frac12(1+\delta^2+\alpha(1-\delta^2))\in(0,1)$. By (10.26) and (10.29),

$$\begin{aligned}\mu_{L;\beta,\alpha}(|S_i^1|<\delta,\ \forall i\in\mathbb{T}_L)&\leq\exp\big[-\beta\big\{(1-\delta^2)(1-\alpha)-b(\tilde\delta)/\beta\big\}|\mathbb{T}_L|\big]\\&\leq\exp\big[-\tfrac12(1-\delta^2)(1-\alpha)\beta|\mathbb{T}_L|\big],\end{aligned}$$

for all $\beta\geq\beta_1(\alpha,\delta)\stackrel{\text{def}}{=}b(\tilde\delta)/((1-\delta)^2(1-\alpha))$. By (10.24), this ensures that, for any $\alpha,\delta<1$ and any $\beta\geq\beta_1(\alpha,\delta)$,

$$\mu_{L;\beta,\alpha}(|S_0^1|<\delta)\leq\exp\big[-\tfrac12(1-\delta^2)(1-\alpha)\beta\big].$$

The right-hand side can be made as small as desired by taking β large enough. □

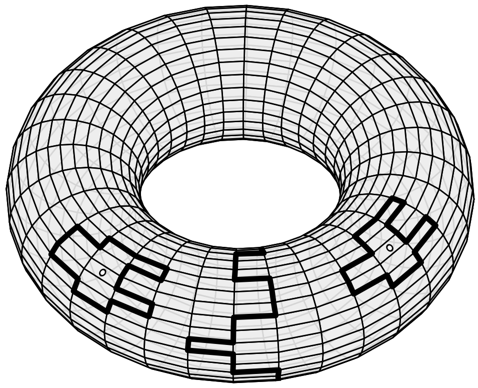

Figure 10.5: The three types of contours on a torus, separating the vertices 0 and j (indicated by the two dots). The mesh corresponds to the dual lattice here, with the vertices in the middle of the faces.

Proof of Lemma 10.20. This proof relies on a variant of Peierls' argument, as exposed in Section 3.7.2. We assume that the reader is familiar with this material.

Let, for each $i \in \mathbb{T}_L$, $I_i^+ \stackrel{\text{def}}{=} \mathbf{1}_{\{S_i^1 \geq 0\}}$ and $I_i^- \stackrel{\text{def}}{=} \mathbf{1}_{\{S_i^1 \leq 0\}}$. We have, by symmetry,

$$\mu_{L;\beta,\alpha}(S_0^1 S_j^1 \leq 0) = 2\langle I_0^+ I_j^- \rangle_{L;\beta,\alpha}.$$

Since, almost surely, $I_i^+ + I_i^- = 1$ for all $i \in \mathbb{T}_L$ (namely, $\{S_i^1 = 0\}$ has measure zero under the reference measure and therefore also under $\mu_{L;\beta,\alpha}$), we can write

$$\langle I_0^+ I_j^- \rangle_{L;\beta,\alpha} = \Big\langle I_0^+ I_j^- \prod_{i \in \mathbb{T}_L \setminus \{0,j\}} (I_i^+ + I_i^-) \Big\rangle_{L;\beta,\alpha} = \sum_{\substack{\eta \in \{-1,1\}^{\mathbb{T}_L} \\ \eta_0 = 1, \eta_j = -1}} \Big\langle \prod_{i \in \mathbb{T}_L} I_i^{\eta_i} \Big\rangle_{L;\beta,\alpha}.$$

To each configuration η appearing in the sum, we associate the corresponding set of contours $\Gamma(\eta)$, exactly as in Section 3.7.2 (including the deformation rule). Note, however, that it would not be possible to reconstruct a configuration ω only from the geometry of its contours: the latter only determine the configuration up to a global spin flip. In order to avoid this problem, we consider contours that are not purely geometrical objects, but also include the information of the values of the spins on both "sides". When u, v denote neighbors separated by γ, we will make the convention that $\eta_u = +1$ and $\eta_v = -1$.

The configurations η appearing in the sum above are such that $\eta_0 \neq \eta_j$. Therefore, there exists (at least) one contour γ separating 0 and j, in the sense that it satisfies one of the three following conditions (see Figure 10.5): (i) γ surrounds 0 but not j, (ii) γ surrounds j but not 0, (iii) γ is winding around the torus (of course, in this case, there must be at least one other such contour). We can thus write

$$\Big\langle I_0^+ I_j^- \Big\rangle_{L;\beta,\alpha} \leq \sum_{\gamma} \sum_{\eta : \Gamma(\eta) \ni \gamma} \Big\langle \prod_{i \in \mathbb{T}_L} I_i^{\eta_i} \Big\rangle_{L;\beta,\alpha}, \tag{10.30}$$

where the first sum is taken over all contours separating 0 and j.

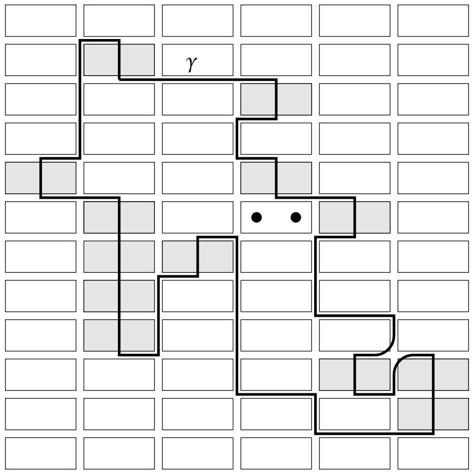

Figure 10.6: The partition of $\mathbb{T}_L$ (here, $L = 12$) in 2×1 blocks, using translates of $\{(0,0),(1,0)\}$ (represented by the two dots). For a contour γ, the set $\mathscr{E}_\gamma^{\mathrm{h},0}$ (whose corresponding blocks were highlighted) represents the blocks of that partition which are crossed by γ in their middle.

Given a contour γ, we denote by $\mathscr{E}_\gamma$ the set of all edges of $\mathscr{E}_L$ which are crossed by γ. Notice that

$$\Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha} = \Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma} I_u^+ I_v^- \prod_{k\in\mathbb{T}_L} (I_k^+ + I_k^-) \Big\rangle_{L;\beta,\alpha}$$
$$\geq \sum_{\eta:\Gamma(\eta)\ni\gamma} \Big\langle \prod_{i\in\mathbb{T}_L} I_i^{\eta_i} \Big\rangle_{L;\beta,\alpha}. \tag{10.31}$$

The last inequality is due to the fact that forcing $\eta_u \neq \eta_v$ for each $\{u,v\}\in\mathscr{E}_\gamma$ is not sufficient to guarantee that $\gamma\in\Gamma(\eta)$ (remember, in particular, the deformation rule used in the definition of contours). Putting all this together,

$$\mu_{L;\beta,\alpha}(S_0^1 S_j^1 \leq 0) \leq 2\sum_\gamma \Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha}. \tag{10.32}$$

The chessboard estimate will be used to show that the presence of a contour is strongly suppressed when $\alpha < 1$ and β is taken sufficiently large:

$$\Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha} \leq e^{-c(\alpha)\beta|\gamma|}, \tag{10.33}$$

where $c(\alpha) \stackrel{\text{def}}{=} (1-\alpha)/16 > 0$.

In order to use the chessboard estimate, we consider four distinct partitions of the torus into blocks. Consider first the partition of $\mathbb{T}_L$ into blocks of sizes 2×1, translates of $\{(0,0),(1,0)\}$ by all vectors of the form $2m\mathbf{e}_1 + n\mathbf{e}_2$ (m,n are integers) (see Figure 10.6). This partition can be identified with the set $\mathscr{E}_L^{\mathrm{h},0}\subset\mathscr{E}_L$ of horizontal nearest-neighbor edges with both endpoints in the same block of the partition. We will use $\{u,v\}\in\mathscr{E}_L^{\mathrm{h},0}$ to index the $|\mathbb{T}_L|/2$ blocks of this partition.

Similarly, one defines the partition made of 2×1 blocks that are translates of $\{(1,0),(2,0)\}$; the corresponding set of horizontal edges is written $\mathscr{E}_L^{\mathrm{h},1}\subset\mathscr{E}_L$.

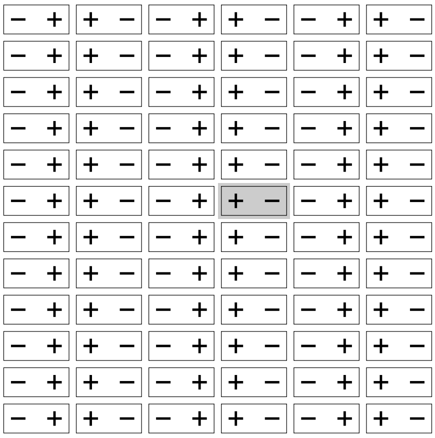

Figure 10.7: The configuration $\eta^{\mathrm{h},0}$ (on the same torus as in Figure 10.6). The shaded block, containing the origin, is the support of ϕ^{+-}, which is then spread out throughout the torus, by successive reflections through the edges separating the blocks.

Finally, we define two partitions made of 1×2 blocks, which are, respectively, translates of the block $\{(0,0),(0,1)\}$ or of the block $\{(0,1),(0,2)\}$. The corresponding sets of vertical edges are denoted $\mathscr{E}_L^{\mathrm{v},0}$ and $\mathscr{E}_L^{\mathrm{v},1}$ respectively.

This leads us to split the edges crossing γ into four families, according to the element of the partition to which they belong: $\mathscr{E}_\gamma = \mathscr{E}_\gamma^{\mathrm{h},0} \cup \mathscr{E}_\gamma^{\mathrm{h},1} \cup \mathscr{E}_\gamma^{\mathrm{v},0} \cup \mathscr{E}_\gamma^{\mathrm{v},1}$. Applying twice the Cauchy–Schwarz inequality,

$$\Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha} \le \prod_{\substack{a\in\{\mathrm{h},\mathrm{v}\}\\ \#\in\{0,1\}}} \Big(\Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma^{a,\#}} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha} \Big)^{1/4}. \tag{10.34}$$

The four factors in the right-hand side can be treated in the same way. To be specific, we consider the factor with $a = \mathrm{h}$ and $\# = 0$. Notice that, for each $\{u,v\} \in \mathscr{E}_\gamma^{\mathrm{h},0}$, the function $I_u^+ I_v^-$ can be obtained by successive reflections through edges (between the blocks of $\mathscr{E}_L^{\mathrm{h},0}$) of one of the two following $\{(0,0),(1,0)\}$-local functions: $\phi^{+-} \overset{\mathrm{def}}{=} I_{(0,0)}^+ I_{(1,0)}^-$ and $\phi^{-+} \overset{\mathrm{def}}{=} I_{(0,0)}^- I_{(1,0)}^+$; we denote the corresponding function $f_{\{u,v\}}$. For each $\{u,v\} \in \mathscr{E}_L^{\mathrm{h},0} \setminus \mathscr{E}_\gamma^{\mathrm{h},0}$, we take $f_{\{u,v\}} \overset{\mathrm{def}}{=} 1$. The chessboard estimate (which we use for non-square blocks here, see Remark 10.15) yields

$$\Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma^{\mathrm{h},0}} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha} \le \prod_{\{u,v\}\in\mathscr{E}_\gamma^{\mathrm{h},0}} \Big\langle \prod_{\{u',v'\}\in\mathscr{E}_\gamma^{\mathrm{h},0}} f_{\{u,v\}}{}^{[\{u',v'\}]} \Big\rangle^{1/(|\mathbb{T}_L|/2)}. \tag{10.35}$$

Now, by translation invariance, for all $\{u,v\} \in \mathscr{E}_\gamma^{\mathrm{h},0}$,

$$\Big\langle \prod_{\{u',v'\}\in\mathscr{E}_\gamma^{\mathrm{h},0}} f_{\{u,v\}}{}^{[\{u',v'\}]} \Big\rangle = \Big\langle \prod_{\{u',v'\}\in\mathscr{E}_\gamma^{\mathrm{h},0}} (\phi^{+-})^{[\{u',v'\}]} \Big\rangle_{L;\beta,\alpha} = \Big\langle \prod_{i\in\mathbb{T}_L} I_i^{\eta_i^{\mathrm{h},0}} \Big\rangle_{L;\beta,\alpha}, \tag{10.36}$$

where $\eta^{\mathrm{h},0} \in \{\pm1\}^{\mathbb{T}_L}$ is depicted in Figure 10.7.

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

To evaluate the last expectation in (10.36), we proceed similarly to what we did in the proof of Lemma 10.19:

$$\Big\langle \prod_{i\in\mathbb{T}_L} I_i^{\eta_i^{\mathrm{h},0}} \Big\rangle_{\mu_0} = \frac{\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}} \prod_{i\in\mathbb{T}_L} I_i^{\eta_i^{\mathrm{h},0}} \big\rangle_{\mu_0}}{\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}} \big\rangle_{\mu_0}}\,.$$

Let us bound from below the energy of any configuration for which $\prod_{i\in\mathbb{T}_L} I_i^{\eta_i^{\mathrm{h},0}} = 1$. Each edge between two vertices at which $\eta^{\mathrm{h},0}$ takes the same value contributes at least $-\beta$ to the energy and those edges account for $\frac{3}{4}$ of all the edges of the torus. However, it is easy to check that, for spins located at the endpoints of the remaining edges, the minimal energy is obtained when both their first components vanish; this yields a minimal contribution of $-\beta\alpha$. We conclude that the energy of the relevant configurations is always at least $-2(\frac{3}{4}+\frac{1}{4}\alpha)\beta|\mathbb{T}_L|$ and, therefore,

$$\big\langle e^{-\mathscr{H}_{L;\beta,\alpha}} \prod_{i\in\mathbb{T}_L} I_i^{\eta_i^{\mathrm{h},0}} \big\rangle_{\mu_0} \le \exp\big\{\tfrac{1}{2}\beta(3+\alpha)|\mathbb{T}_L|\big\}\,.$$

Combining this with the lower bound (10.29), we obtain, choosing $\tilde{\delta}^2 = (7+\alpha)/8$,

$$\begin{aligned}\Big\langle \prod_{i\in\mathbb{T}_L} I_i^{\eta_i^{\mathrm{h},0}} \Big\rangle_{L;\beta,\alpha} &\le \exp\Big\{-2\beta\big(\tilde{\delta}^2 - \frac{b(\tilde{\delta})}{2\beta} - \tfrac{1}{4}(3+\alpha)\big)|\mathbb{T}_L|\Big\}\\ &= \exp\Big\{-\tfrac{1}{4}\beta\big(1-\alpha-\frac{4b(\tilde{\delta})}{\beta}\big)|\mathbb{T}_L|\Big\}\\ &\le \exp\big\{-\tfrac{1}{8}(1-\alpha)\beta|\mathbb{T}_L|\big\}\,,\end{aligned}$$

for all $\beta \ge 8b(\tilde{\delta})/(1-\alpha)$. Inserting this into (10.35), we obtain

$$\Big\langle \prod_{\{u,v\}\in\mathscr{E}_\gamma^{\mathrm{h},0}} I_u^+ I_v^- \Big\rangle_{L;\beta,\alpha} \le \exp\big\{-\tfrac{1}{4}(1-\alpha)\beta|\mathscr{E}_\gamma^{\mathrm{h},0}|\big\}\,.$$

Doing this for the other three partitions and using (10.34) and the fact that $|\mathscr{E}_\gamma^{\mathrm{h},0}| + |\mathscr{E}_\gamma^{\mathrm{h},1}| + |\mathscr{E}_\gamma^{\mathrm{v},0}| + |\mathscr{E}_\gamma^{\mathrm{v},1}| = |\gamma|$, (10.33) follows. Using this estimate, (10.32) becomes

$$\mu_{L;\beta,\alpha}(S_0^1 S_j^1 \le 0) \le 2\sum_\gamma e^{-c(\alpha)\beta|\gamma|}\,.$$

We can now conclude the proof following the energy-entropy argument used when implementing Peierls' argument in Section 3.7.2. There is only one minor difference: in the sum over γ, there are also contours that wind around the torus, a situation we did not have to consider in Chapter 3. However, since such contours have length at least L,

$$\sum_{\gamma,\ \mathrm{winding}} e^{-c(\alpha)\beta|\gamma|} = \sum_{k\ge L} e^{-c(\alpha)\beta k}\#\{\gamma:\ \mathrm{winding}, |\gamma|=k\} \le \sum_{k\ge L} e^{-c(\alpha)\beta k}(L^2 8^k)\,.$$

Taking β sufficiently large, this last sum is bounded uniformly in L and can be made as small as desired. The conclusion thus follows exactly as in Chapter 3. □

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

Remark 10.21. As the reader can check, the arguments above apply more generally. In particular, they extend readily to the anisotropic $O(N)$ model, in which the spins $\mathbf{S}_i = (S_i^1, \dots, S_i^N) \in \mathbb{S}^{N-1}$ and the Hamiltonian is given by

$$\mathscr{H}_{L;\beta,\alpha} = -\beta \sum_{\{i,j\}\in\mathscr{E}_L} \{S_i^1 S_j^1 + \alpha\big(S_i^2 S_j^2 + \cdots + S_i^N S_j^N\big)\}.$$

Also, the extension to $d \ge 3$ is rather straightforward (only the implementation of Peierls' argument is affected and can be dealt with as in Exercise 3.20). ⋄

We can finally conclude the proof of Theorem 10.2. We will rely on the main result of Section 6.11.

Proof of Theorem 10.2. Let $m_L^1 \stackrel{\text{def}}{=} \mathbf{m}_L \cdot \mathbf{e}_1$ and

$$\begin{aligned}\psi(h) &\stackrel{\text{def}}{=} \lim_{L\to\infty} \frac{1}{|\mathbb{T}_L|} \log\Big\langle \exp\Big\{h \sum_{j\in\mathbb{T}_L} S_j^1\Big\}\Big\rangle_{L;\beta,\alpha} \\ &= \lim_{L\to\infty} \frac{1}{|\mathbb{T}_L|} \log\big\langle e^{h m_L^1 |\mathbb{T}_L|}\big\rangle_{L;\beta,\alpha}.\end{aligned}$$

Existence of this limit and its convexity in h follow from Lemma 6.89 (used with $g \stackrel{\text{def}}{=} S_0^1$ and with periodic boundary conditions, for which that result also holds). We have seen that m_L^1 remains bounded away from zero with high probability, uniformly in L, when β is large. We are going to show that this implies that ψ is not differentiable at $h = 0$.

Let $0 \le \alpha < 1$, $\epsilon > 0$ and $\beta > \beta_0(\alpha,\epsilon)$, where $\beta_0(\alpha,\epsilon)$ was introduced in Theorem 10.17. Let also $0 < \delta < 1$ be such that $\delta^2 < 1-\epsilon$. To start, observe that, for any $h \ge 0$,

$$\big\langle e^{h m_L^1 |\mathbb{T}_L|}\big\rangle_{L;\beta,\alpha} \ge \big\langle e^{h m_L^1 |\mathbb{T}_L|} \mathbf{1}_{\{m_L^1 \ge \delta\}}\big\rangle_{L;\beta,\alpha} \ge e^{\delta h |\mathbb{T}_L|} \mu_{L;\beta,\alpha}(m_L^1 \ge \delta). \tag{10.37}$$

But Theorem 10.17 implies that, uniformly in L (multiple of 4),

$$1-\epsilon \le \big\langle (m_L^1)^2\big\rangle_{L;\beta,\alpha} \le \delta^2 + \mu_{L;\beta,\alpha}\big((m_L^1)^2 \ge \delta^2\big).$$

Therefore, again uniformly in L (multiple of 4),

$$\mu_{L;\beta,\alpha}(m_L^1 \ge \delta) = \tfrac{1}{2}\mu_{L;\beta,\alpha}\big((m_L^1)^2 \ge \delta^2\big) \ge \tfrac{1}{2}(1-\epsilon-\delta^2) > 0.$$

Inserting this estimate in (10.37), we conclude that $(\psi(h)-\psi(0))/h = \psi(h)/h \ge \delta$, for all $h > 0$. Letting $h \downarrow 0$ yields

$$\frac{\partial \psi}{\partial h^+}\Big|_{h=0} \ge \delta > 0.$$

Since $\psi(-h) = \psi(h)$, this implies that ψ is not differentiable at $h = 0$. Proposition 6.91 then guarantees the existence of two Gibbs measures $\mu^+ \neq \mu^-$ such that

$$\langle S_0^1\rangle_{\mu^+} = \frac{\partial \psi}{\partial h^+}\Big|_{h=0} > 0 > \frac{\partial \psi}{\partial h^-}\Big|_{h=0} = \langle S_0^1\rangle_{\mu^-},$$

thereby completing the proof of Theorem 10.2. □

10.5 The infrared bound

We now turn to the second major consequence of reflection positivity, the *infrared bound*, which provides one of the few known approaches to proving spontaneous breaking of a continuous symmetry. In this section we go back to the d-dimensional torus $\mathbb{T}_L$, $d \geq 1$.

In order to motivate the infrared bound, we start by showing how it appears as the central tool to prove Theorem 10.1, in Section 10.5.2. The proof of the infrared bound and of the related Gaussian domination is provided in Section 10.5.3.

10.5.1 Models to be considered

The infrared bound holds for a wide class of models, but still requires a little more structure than just reflection positivity. Namely, we will assume that the spins are ν-dimensional vectors,

$$\Omega_0 \stackrel{\text{def}}{=} \mathbb{R}^\nu ,$$

and that the Hamiltonian is given by

$$\mathscr{H}_{L;\beta} \stackrel{\text{def}}{=} \beta \sum_{\{i,j\}\in\mathscr{E}_L} \|\mathbf{S}_i - \mathbf{S}_j\|_2^2 , \tag{10.38}$$

with $\beta \geq 0$. As before, we assume that the reference measure ρ on Ω_0 (equipped with the Borel subsets of $\mathbb{R}^\nu$) is supported on a compact subset of $\mathbb{R}^\nu$ and write $\mu_0 \stackrel{\text{def}}{=} \bigotimes_{i\in\mathbb{T}_L} \rho$. The Gibbs distribution $\mu_{L;\beta}$ on $(\Omega_L, \mathscr{F}_L)$, associated to $\mathscr{H}_{L;\beta}$, is then defined exactly as in (10.5).

The choice of the reference measure ρ leads to various interesting models encountered in previous chapters.

Example 10.22. Choose $\nu = N$ and let ρ be the Lebesgue measure on the sphere $\mathbb{S}^{N-1} \subset \mathbb{R}^N$. Since $\|\mathbf{S}_i\|_2 = 1$ for all $i \in \mathbb{T}_L$, almost surely, the Hamiltonian can be rewritten as

$$\mathscr{H}_{L;\beta} = 2\beta|\mathscr{E}_L| - 2\beta \sum_{\{i,j\}\in\mathscr{E}_L} \mathbf{S}_i \cdot \mathbf{S}_j .$$

We recognize (up to an irrelevant constant $2\beta|\mathscr{E}_L|$) the Hamiltonian of the **O(*N*) model**. ⋄

Example 10.23. Choose $\nu = q - 1$ and let ρ be the uniform distribution concentrated on the vertices of the regular ν-simplex (see Figure 10.8). The vertices of this simplex lie on $\mathbb{S}^{\nu-1}$ and the scalar product of any two vectors from the origin to two distinct vertices of the simplex is always the same. Note that this is just the ***q*-state Potts model** in disguise. Indeed, a configuration can almost surely be identified with a configuration $\omega' \in \{1,\dots,q\}^{\mathbb{T}_L}$, where $1,\dots,q$ is a numbering of the vertices of the simplex. Then, up to an irrelevant constant, we see that the Hamiltonian becomes $-\beta_\nu \sum_{\{i,j\}\in\mathscr{E}_L} \delta_{\omega'_i,\omega'_j}$, for some $\beta_\nu \geq 0$ proportional to β. ⋄

10.5.2 Application: Orientational long-range order in the $O(N)$ model

In order to motivate the infrared bound, we start with one of its major applications: the proof that, when $d \geq 3$, there is orientational long-range order at low temperatures for models with continuous spins of the type described above.

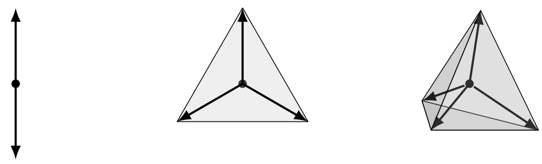

Figure 10.8: The simplex representation for the 2-, 3- and 4-state Potts model.

We have seen, when proving the Mermin–Wagner Theorem in Chapter 9, that the absence of orientational long-range order in the two-dimensional $\mathsf{O}(N)$ model was due to the fact that *spin waves*, by which we meant spin configurations varying slowly over macroscopic regions, were created in the system, at arbitrarily low cost (remember Figure 9.3). If we want to establish orientational long-range order, we have to exclude the existence of such excitations. In order to do that, it is very convenient to consider the *Fourier representation* of the variables $(\mathbf{S}_j)_{j\in\mathbb{T}_L}$.

Consider the **reciprocal torus**, defined by

$$\mathbb{T}_L^\star \stackrel{\text{def}}{=} \left\{\frac{2\pi}{L}(n_1,\dots,n_d) : 0 \le n_i < L\right\}.$$

Note that $|\mathbb{T}_L^\star| = |\mathbb{T}_L|$. The **Fourier transform of** $(\mathbf{S}_j)_{j\in\mathbb{T}_L}$ is $(\widehat{\mathbf{S}}_p)_{p\in\mathbb{T}_L^\star}$, defined by

$$\widehat{\mathbf{S}}_p \stackrel{\text{def}}{=} \frac{1}{|\mathbb{T}_L|^{1/2}} \sum_{j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot j}\mathbf{S}_j, \qquad p \in \mathbb{T}_L^\star.$$

Let us recall two important properties,

1. First, the original variables can be reconstructed from their Fourier transform, by the inversion formula:

$$\mathbf{S}_j = \frac{1}{|\mathbb{T}_L^\star|^{1/2}} \sum_{p\in\mathbb{T}_L^\star} e^{-\mathrm{i}p\cdot j}\widehat{\mathbf{S}}_p, \qquad j \in \mathbb{T}_L.$$

   Each index p is called a **mode** and corresponds to an oscillatory term $e^{-\mathrm{i}p\cdot j}$. This sum should be interpreted as the contributions of the different Fourier modes to the field variable $\mathbf{S}_j$. The importance of mode p is measured by $\|\widehat{\mathbf{S}}_p\|_2$. On the one hand, modes with small values of p describe slow variations of $\mathbf{S}_j$, meaning variations detectable only on macroscopic regions, at the scale of the torus. In particular, the mode $p = 0$ corresponds to the non-oscillating ("infinite wavelength") component of $\mathbf{S}_j$ and is proportional to the magnetization of the system (see below). On the other hand, modes with large p represent rapid oscillations present in $\mathbf{S}_j$.

2. Second, **Plancherel's Theorem** states that

$$\sum_{p\in\mathbb{T}_L^\star} \|\widehat{\mathbf{S}}_p\|_2^2 = \sum_{j\in\mathbb{T}_L} \|\mathbf{S}_j\|_2^2. \tag{10.39}$$

Exercise 10.6. *Prove the above two properties.*

As mentioned above, the magnetization density $\mathbf{m}_L = \frac{1}{|\mathbb{T}_L|}\sum_{i\in\mathbb{T}_L}\mathbf{S}_i$ is simply related to the $p=0$ mode by

$$\mathbf{m}_L = \frac{1}{|\mathbb{T}_L|^{1/2}}\widehat{\mathbf{S}}_0 .$$

Therefore, the importance of the $p=0$ mode characterizes the presence or absence of orientational long-range order in the system. For example, we have seen in Exercise 9.1 that the contribution of the $p=0$ mode becomes negligible in the thermodynamic limit for the two-dimensional XY model; this was due to the appearance of spin waves. Therefore, to prove that orientational long-range order *does* occur, one must show that the $p=0$ mode has a non-zero contribution even in the thermodynamic limit.

In order to do this, we add a new restriction to the class of models we consider. Namely, we will assume in the rest of this section that the reference measure ρ is such that, almost surely,

$$\|\mathbf{S}_j\|_2 = 1 \quad \forall j\in\mathbb{T}_L .$$

This is of course the case in the $\mathsf{O}(N)$ and Potts models. With this assumption, (10.39) implies that $\sum_{p\in\mathbb{T}_L^\star}\|\widehat{\mathbf{S}}_p\|_2^2 = |\mathbb{T}_L|$, which yields

$$\|\widehat{\mathbf{S}}_0\|_2^2 = |\mathbb{T}_L| - \sum_{\substack{p\in\mathbb{T}_L^\star\\ p\neq 0}}\|\widehat{\mathbf{S}}_p\|_2^2 .$$

Moreover, by translation invariance,

$$\langle\|\hat{\mathbf{S}}_p\|_2^2\rangle_{L;\beta} = \frac{1}{|\mathbb{T}_L|}\sum_{i,j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot(j-i)}\langle\mathbf{S}_i\cdot\mathbf{S}_j\rangle_{L;\beta} = \sum_{j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot j}\langle\mathbf{S}_0\cdot\mathbf{S}_j\rangle_{L;\beta} .$$

Gathering these identities, we conclude that

$$\langle\|\mathbf{m}_L\|_2^2\rangle_{L;\beta} = \frac{1}{|\mathbb{T}_L|}\langle\|\widehat{\mathbf{S}}_0\|_2^2\rangle_{L;\beta} = 1 - \Big\{\frac{1}{|\mathbb{T}_L|}\sum_{\substack{p\in\mathbb{T}_L^\star\\ p\neq 0}}\sum_{j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot j}\langle\mathbf{S}_0\cdot\mathbf{S}_j\rangle_{L;\beta}\Big\} . \tag{10.40}$$

To obtain a lower bound on $\langle\|\mathbf{m}_L\|_2^2\rangle_{L;\beta}$, we thus need to find an upper bound on the double sum appearing on the right-hand side of the previous display. This is precisely at this stage that the infrared bound becomes crucial; its proof will be provided in Section 10.5.3.

Theorem 10.24 (Infrared bound)**.** *Let $\mu_{L;\beta}$ be the Gibbs distribution associated to a Hamiltonian of the form* (10.38)*. Then, for any $p\in\mathbb{T}_L^\star\setminus\{0\}$,*

$$\sum_{j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot j}\langle\mathbf{S}_0\cdot\mathbf{S}_j\rangle_{L;\beta} \le \frac{\nu}{4\beta d}\Big\{1-\frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}^{-1} .$$

Using the infrared bound in (10.40), we get

$$\big\langle\|\mathbf{m}_L\|_2^2\big\rangle_{L;\beta} \ge 1 - \frac{\nu}{4\beta d}\frac{1}{|\mathbb{T}_L|}\sum_{\substack{p\in\mathbb{T}_L^\star\\ p\neq 0}}\Big\{1-\frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}^{-1} .$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

The reader can recognize a Riemann sum in the right-hand side, which implies that

$$\begin{aligned}\beta_0 &\stackrel{\text{def}}{=} \frac{\nu}{4d(2\pi)^d} \lim_{L\to\infty} \frac{(2\pi)^d}{|\mathbb{T}_L|} \sum_{\substack{p\in\mathbb{T}_L^\star\\ p\neq 0}} \Big\{1-\frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}^{-1}\\ &= \frac{\nu}{4d(2\pi)^d}\int_{[-\pi,\pi]^d}\Big\{1-\frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}^{-1}\mathrm{d}p. \qquad (10.41)\end{aligned}$$

Notice that this integral is improper, precisely because of the singularity of the integrand at $p=0$. Therefore,

$$\liminf_{L\to\infty}\langle \|\mathbf{m}_L\|_2^2\rangle_{L;\beta} \geq 1-\frac{\beta_0}{\beta}.$$

This proves that there is orientational long-range order for all $\beta>\beta_0$. The only remaining task is to make sure that β_0 is indeed finite.

It turns out (see Theorem B.72) that β_0 is finite if and only if the symmetric simple random walk on $\mathbb{Z}^d$ is transient. As shown in Corollary B.73 (by directly studying the integral above), this occurs if and only if $d\geq 3$.

We have thus proved the following result.

Theorem 10.25. *Assume that $d\geq 3$. Let $\mu_{L,\beta}$ be defined with respect to a reference measure μ_0 under which $\|\mathbf{S}_i\|_2=1$, almost surely, for all $i\in\mathbb{T}_L$. Let*

$$\beta_0 \stackrel{\text{def}}{=} \frac{\nu}{4d(2\pi)^d}\int_{[-\pi,\pi]^d}\{1-\tfrac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\}^{-1}\mathrm{d}p.$$

Then $\beta_0<\infty$ and, for any $\beta>\beta_0$,

$$\liminf_{L\to\infty}\big\langle \|\mathbf{m}_L\|_2^2\big\rangle_{L;\beta} \geq 1-\frac{\beta_0}{\beta}. \qquad (10.42)$$

Remark 10.26. Theorem 10.25 implies the existence of orientational long-range order for the q-state Potts model on $\mathbb{Z}^d$, $d\geq 3$. The latter, however, displays orientational long-range order also in dimension 2, even though this cannot be inferred from the infrared bound. (This can be proved, for example, via a Peierls argument as in Section 3.7.2 or using the chessboard estimate, using a variant of the proof of Lemma 10.19.) The crucial difference, of course, is that the symmetry group is discrete in this case. ◇

With the help of Theorem 10.25, we can now prove existence of a continuum of distinct Gibbs states in such models, as stated in Theorem 10.1.

Proof of Theorem 10.1. First, fix some unit vector $\mathbf{e}\in\mathbb{S}^{N-1}$. For simplicity and with no loss of generality, we can take $\mathbf{e}=\mathbf{e}_1$ (indeed, $\mu_{L;\beta}$ is invariant under any global rotation of the spins). Then, define

$$\psi(h) \stackrel{\text{def}}{=} \lim_{L\to\infty}\frac{1}{|\mathbb{T}_L|}\log\Big\langle \exp\Big\{h\sum_{j\in\mathbb{T}_L}\mathbf{S}_j\cdot\mathbf{e}_1\Big\}\Big\rangle_{L;\beta}.$$

We again use Theorem 10.25 to show that ψ is not differentiable at $h=0$, following the pattern used to prove Theorem 10.2, and conclude using Proposition 6.91.

The only difference here is when showing that $\mu_{L;\beta}(\mathbf{m}_L\cdot\mathbf{e}_1\geq\delta)$ is bounded away from zero. To use the lower bound we have on $\langle\|\mathbf{m}_L\|_2\rangle_{L;\beta}$, we can use the following comparisons:

$$\mu_{L;\beta}(\mathbf{m}_L\cdot\mathbf{e}_1\geq\delta)\geq\tfrac{1}{2d}\mu_{L;\beta}(\|\mathbf{m}_L\|_\infty\geq\delta)\geq\tfrac{1}{2d}\mu_{L;\beta}(\|\mathbf{m}_L\|_2\geq\delta\sqrt{d})$$

When δ is sufficiently small, a lower bound on the latter, uniform in L, can be obtained as before. □

10.5.3 Gaussian domination and the infrared bound

The infrared bound relies on the following proposition, whose proof will use reflection positivity. Let $h=(h_i)_{i\in\mathbb{T}_L}\in(\mathbb{R}^\nu)^{\mathbb{T}_L}$ and

$$\mathbf{Z}_{L;\beta}(h)\stackrel{\text{def}}{=}\Big\langle\exp\Big\{-\beta\sum_{\{i,j\}\in\mathscr{E}_L}\|\mathbf{S}_i-\mathbf{S}_j+h_i-h_j\|_2^2\Big\}\Big\rangle_{\mu_0}.$$

Notice that $\mathbf{Z}_{L;\beta}(h)$ is well defined (since we are assuming ρ to have compact support) and that $\mathbf{Z}_{L;\beta}(0)$ coincides with the partition function $\mathbf{Z}_{L;\beta}$ associated to $\mathscr{H}_{L;\beta}$.

Proposition 10.27 (Gaussian domination)**.** *For all $h=(h_i)_{i\in\mathbb{T}_L}$,*

$$\mathbf{Z}_{L;\beta}(h)\leq\mathbf{Z}_{L;\beta}(0)\,. \tag{10.43}$$

Proposition 10.27 will be a consequence of the following lemma, which is a version of the Cauchy–Schwarz-type inequality of Lemma 10.6.

Lemma 10.28. *Let $\mu\in\mathscr{M}_{\mathrm{RP}(\Theta)}$ and $A,B,C_\alpha,D_\alpha\in\mathfrak{A}_+(\Theta)$. Then*

$$\big\{\big\langle e^{A+\Theta(B)+\sum_\alpha C_\alpha\Theta(D_\alpha)}\big\rangle_\mu\big\}^2\leq\big\langle e^{A+\Theta(A)+\sum_\alpha C_\alpha\Theta(C_\alpha)}\big\rangle_\mu\big\langle e^{B+\Theta(B)+\sum_\alpha D_\alpha\Theta(D_\alpha)}\big\rangle_\mu.$$

Proof. Expanding the exponential,

$$\big\langle e^{A+\Theta(B)+\sum_\alpha C_\alpha\Theta(D_\alpha)}\big\rangle_\mu=\sum_{n\geq0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n}\big\langle e^AC_{\alpha_1}\cdots C_{\alpha_n}\Theta(e^BD_{\alpha_1}\cdots D_{\alpha_n})\big\rangle_\mu. \tag{10.44}$$

By Lemma 10.6,

$$\begin{aligned}\big\langle e^AC_{\alpha_1}\cdots C_{\alpha_n}&\Theta(e^BD_{\alpha_1}\cdots D_{\alpha_n})\big\rangle_\mu\\&\leq\big\langle e^AC_{\alpha_1}\cdots C_{\alpha_n}\Theta(e^AC_{\alpha_1}\cdots C_{\alpha_n})\big\rangle_\mu^{1/2}\\&\qquad\times\big\langle e^BD_{\alpha_1}\cdots D_{\alpha_n}\Theta(e^BD_{\alpha_1}\cdots D_{\alpha_n})\big\rangle_\mu^{1/2}.\end{aligned}$$

The classical Cauchy–Schwarz inequality, $\sum_k|a_kb_k|\leq(\sum_k a_k^2)^{1/2}(\sum_k b_k^2)^{1/2}$, yields

$$\begin{aligned}\sum_{\alpha_1,\dots,\alpha_n}&\langle e^AC_{\alpha_1}\cdots C_{\alpha_n}\Theta(e^BD_{\alpha_1}\cdots D_{\alpha_n})\rangle_\mu\\&\leq\Big[\sum_{\alpha_1,\dots,\alpha_n}\big\langle e^AC_{\alpha_1}\cdots C_{\alpha_n}\Theta(e^AC_{\alpha_1}\cdots C_{\alpha_n})\big\rangle_\mu\Big]^{1/2}\\&\qquad\times\Big[\sum_{\alpha_1,\dots,\alpha_n}\big[\big\langle e^BD_{\alpha_1}\cdots D_{\alpha_n}\Theta(e^BD_{\alpha_1}\cdots D_{\alpha_n})\big\rangle_\mu\Big]^{1/2}.\end{aligned}$$

Inserting this in (10.44), using again the Cauchy–Schwarz inequality (this time, to the sum over n) and resumming the series, we get

$$\begin{aligned}\big\langle e^{A+\Theta(B)+\sum_\alpha C_\alpha\Theta(D_\alpha)}\big\rangle_\mu \\ &\le \Big[\sum_{n\ge0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n}\big\langle e^A C_{\alpha_1}\cdots C_{\alpha_n}\Theta(e^A C_{\alpha_1}\cdots C_{\alpha_n})\big\rangle_\mu\Big]^{1/2} \\ &\qquad\times\Big[\sum_{n\ge0}\frac{1}{n!}\sum_{\alpha_1,\dots,\alpha_n}\big\langle e^B D_{\alpha_1}\cdots D_{\alpha_n}\Theta(e^B D_{\alpha_1}\cdots D_{\alpha_n})\big\rangle_\mu\Big]^{1/2} \\ &= \Big[\big\langle e^{A+\Theta(A)+\sum_\alpha C_\alpha\Theta(C_\alpha)}\big\rangle_\mu\big\langle e^{B+\Theta(B)+\sum_\alpha D_\alpha\Theta(D_\alpha)}\big\rangle_\mu\Big]^{1/2}.\end{aligned}$$

□

Proof of Proposition 10.27. First, notice that $\mathbf{Z}_{L;\beta}(h) = \mathbf{Z}_{L;\beta}(h')$ whenever there exists $c \in \mathbb{R}$ such that $h_i - h'_i = c$ for all $i \in \mathbb{T}_L$. There is thus no loss of generality in assuming that $h_0 = 0$, which we do from now on. Next, observe that $\mathbf{Z}_{L;\beta}(h)$ tends to 0 as any $\|h_i\|_2 \to \infty$, $i \ne 0$. In particular, there exists C such that $\sum_i \|h_i\|_2^2 \le C$ for all h that maximize $\mathbf{Z}_{L;\beta}(h)$. Among the latter, let us denote by $h^\star$ a maximizer that minimizes the quantity $N(h) \stackrel{\text{def}}{=} \#\big\{\{i,j\} \in \mathscr{E}_L : h_i \ne h_j\big\}$. We claim that $N(h^\star) = 0$. Since $h^\star_0 = 0$, this will then imply that $h^\star_i = 0$ for all $i \in \mathbb{T}_L$, which will conclude the proof.

Let us therefore suppose to the contrary that $N(h^\star) > 0$. In that case, we can find $\{i,j\} \in \mathscr{E}_L$ such that $h^\star_i \ne h^\star_j$. Let Π be the reflection plane going through the middle of the edge $\{i,j\}$ and let Θ denote the reflection through Π. Below, we use $\{i',j'\}$ to denote the edges that cross Π, with $i' \in \mathbb{T}_{L,+}$ and $j' = \Theta(i') \in \mathbb{T}_{L,-}$. Since $\|\omega_{i'} - \omega_{j'} + h_{i'} - h_{j'}\|_2^2 = \|\omega_{i'} + h_{i'}\|_2^2 + \|\omega_{j'} + h_{j'}\|_2^2 - 2(\omega_{i'} + h_{i'})\cdot(\omega_{j'} + h_{j'})$, we can write

$$-\beta\sum_{\{i,j\}\in\mathscr{E}_L}\|\omega_i - \omega_j + h_i - h_j\|_2^2 = A + \Theta(B) + \sum_{i'} C_{i'}\cdot\Theta(D_{i'}),$$

where $A, B, C_i, D_i \in \mathfrak{A}_+(\Theta)$, and

$$\begin{aligned} A &\stackrel{\text{def}}{=} -\beta\sum_{\substack{\{i,j\}\in\mathscr{E}_L:\\ i,j\in\mathbb{T}_{L,+}(\Theta)}}\|\omega_i - \omega_j + h_i - h_j\|_2^2 - \beta\sum_{i'}\|\omega_{i'} + h_{i'}\|_2^2, \\ \Theta(B) &\stackrel{\text{def}}{=} -\beta\sum_{\substack{\{i,j\}\in\mathscr{E}_L:\\ i,j\in\mathbb{T}_{L,-}(\Theta)}}\|\omega_i - \omega_j + h_i - h_j\|_2^2 - \beta\sum_{j'}\|\omega_{j'} + h_{j'}\|_2^2, \\ C_{i'} &\stackrel{\text{def}}{=} \sqrt{2\beta}(\omega_{i'} + h_{i'}), \qquad \Theta(D_{i'}) \stackrel{\text{def}}{=} \sqrt{2\beta}(\omega_{j'} + h_{j'}).\end{aligned}$$

(Remember that Θ acts on ω, not on h; this implies that in general, $A \ne B$ and $C_{i'} \ne D_{i'}$.) One can thus use Lemma 10.28 to obtain

$$\mathbf{Z}_{L;\beta}(h^\star)^2 \le \mathbf{Z}_{L;\beta}(h^{\star,+})\mathbf{Z}_{L;\beta}(h^{\star,-}),$$

where

$$h_i^{\star,+} = \begin{cases} h_i^\star & \forall i \in \mathbb{T}_{L,+}(\Theta), \\ h^\star_{\Theta(i)} & \forall i \in \mathbb{T}_{L,-}(\Theta), \end{cases} \qquad h_i^{\star,-} = \begin{cases} h_i^\star & \forall i \in \mathbb{T}_{L,-}(\Theta), \\ h^\star_{\Theta(i)} & \forall i \in \mathbb{T}_{L,+}(\Theta). \end{cases}$$

Our choice of Θ guarantees that $\min\big\{N(h^{\star,+}), N(h^{\star,-})\big\} < N(h^\star)$. To be specific, let us assume that $N(h^{\star,+}) < N(h^\star)$. Then, since $h^\star$ is a maximizer,

$$\mathbf{Z}_{L;\beta}(h^\star)^2 \le \mathbf{Z}_{L;\beta}(h^{\star,+})\mathbf{Z}_{L;\beta}(h^{\star,-}) \le \mathbf{Z}_{L;\beta}(h^{\star,+})\mathbf{Z}_{L;\beta}(h^\star),$$

that is, $\mathbf{Z}_{L;\beta}(h^{\star,+}) \geq \mathbf{Z}_{L;\beta}(h^\star)$. This implies that $h^{\star,+}$ is also a maximizer which satisfies $N(h^{\star,+}) < N(h^\star)$. This contradicts our choice of $h^\star$, and therefore implies that $N(h^\star) = 0$. □

The following exercise provides some motivation for the terminology "Gaussian domination". One can define the discrete Laplacian of $h = (h_i)_{i\in\mathbb{T}_L}$, Δh, by

$$(\Delta h)_i \stackrel{\text{def}}{=} \sum_{j\sim i} (h_j - h_i)\,, \qquad i \in \mathbb{Z}^d\,.$$

The Discrete Green identities of Lemma 8.7 can also be used here; they take slightly simpler forms due to the absence of boundary terms on the torus.

Exercise 10.7. *Show that* (10.43) *can be rewritten as*

$$\Big\langle \exp\Big\{2\beta \sum_{i\in\mathbb{T}_L} (\Delta h)_i \cdot (\mathbf{S}_i - \mathbf{S}_0)\Big\}\Big\rangle_{L;\beta} \leq \exp\Big\{-\beta \sum_{i\in\mathbb{T}_L} (\Delta h)_i \cdot h_i\Big\}. \tag{10.45}$$

Let $\nu_{L;\beta}$ be the Gibbs distribution corresponding to the reference measure given by the Lebesgue measure: $\rho(\mathrm{d}\omega_i) = \mathrm{d}\omega_i$. Show that

$$\Big\langle \exp\Big\{2\beta \sum_{i\in\mathbb{T}_L} (\Delta h)_i \cdot (\mathbf{S}_i - \mathbf{S}_0)\Big\}\Big\rangle_{\nu_{L;\beta}} = \exp\Big\{-\beta \sum_{i\in\mathbb{T}_L} (\Delta h)_i \cdot h_i\Big\},$$

so that the bound (10.45) *is saturated by the Gaussian measure $\nu_{L;\beta}$.*

We can now turn to the proof of the infrared bound.

Proof of Theorem 10.24: We know from Proposition 10.27 that $\mathbf{Z}_{L;\beta}(h)$ is maximal at $h \equiv 0$. Consequently, at fixed h,

$$\frac{\partial}{\partial\lambda}\mathbf{Z}_{L;\beta}(\lambda h)\big|_{\lambda=0} = 0 \quad \text{and} \quad \frac{\partial^2}{\partial\lambda^2}\mathbf{Z}_{L;\beta}(\lambda h)\big|_{\lambda=0} \leq 0\,. \tag{10.46}$$

The first claim in (10.46) does not provide any nontrivial information, but the second one is instrumental in the proof. Elementary computations show that

$$\begin{aligned}
&\frac{\partial^2}{\partial\lambda^2}\mathbf{Z}_{L;\beta}(\lambda h)\big|_{\lambda=0} \\
&\qquad = 4\beta^2 \Big\langle \Big| \sum_{\{i,j\}\in\mathscr{E}_L} (\mathbf{S}_i - \mathbf{S}_j)\cdot(h_i - h_j)\Big|^2 \exp\{-\beta \sum_{\{i,j\}\in\mathscr{E}_L} \|\mathbf{S}_i - \mathbf{S}_j\|_2^2\}\Big\rangle_{\mu_0} \\
&\qquad\qquad - 2\beta \sum_{\{i,j\}\in\mathscr{E}_L} \|h_i - h_j\|_2^2 \Big\langle \exp\{-\beta \sum_{\{i,j\}\in\mathscr{E}_L} \|\mathbf{S}_i - \mathbf{S}_j\|_2^2\}\Big\rangle_{\mu_0}.
\end{aligned}$$

The inequality in (10.46) is thus equivalent to

$$\Big\langle \Big| \sum_{\{i,j\}\in\mathscr{E}_L} (\mathbf{S}_i - \mathbf{S}_j)\cdot(h_i - h_j)\Big|^2 \Big\rangle_{L;\beta} \leq \frac{1}{2\beta} \sum_{\{i,j\}\in\mathscr{E}_L} \|h_i - h_j\|_2^2\,. \tag{10.47}$$

The latter holds for any $h \in (\mathbb{R}^\nu)^{\mathbb{T}_L}$, but it is easily seen that it also extends to any $h \in (\mathbb{C}^\nu)^{\mathbb{T}_L}$ (just treat separately the real and imaginary parts). Let us fix $p \in \mathbb{T}_L^\star \setminus \{0\}$, $\ell \in \{1, \dots, \nu\}$ and make the following specific choice:

$$\forall j \in \mathbb{T}_L\,, \qquad \alpha_j \stackrel{\text{def}}{=} e^{\mathrm{i}\,p\cdot j}\,, \qquad h_j \stackrel{\text{def}}{=} \alpha_j \mathbf{e}_\ell\,.$$

The Green identity (8.14) yields ($\bar{\alpha}$ denoting the complex conjugate of α)

$$\sum_{\{i,j\}\in\mathscr{E}_L} \|h_i - h_j\|_2^2 = \sum_{\{i,j\}\in\mathscr{E}_L} (\nabla\bar{\alpha})_{ij}(\nabla\alpha)_{ij} = \sum_{i\in\mathbb{T}_L} \bar{\alpha}_i(-\Delta\alpha)_i$$
$$= 2d|\mathbb{T}_L|\Big\{1 - \frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\},$$

since, for any $i \in \mathbb{T}_L$,

$$(-\Delta\alpha)_i = \sum_{j\sim i}(\alpha_i - \alpha_j) = e^{\mathrm{i}p\cdot i}\sum_{j\sim i}(1 - e^{\mathrm{i}p\cdot(j-i)}) = 2de^{\mathrm{i}p\cdot i}\Big\{1 - \frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}.$$

Similarly, denoting by $S_i^\ell \stackrel{\text{def}}{=} \mathbf{S}_i\cdot\mathbf{e}_\ell$ the ℓth component of $\mathbf{S}_i$,

$$\sum_{\{i,j\}\in\mathscr{E}_L} (\mathbf{S}_i - \mathbf{S}_j)\cdot(h_i - h_j) = \sum_{\{i,j\}\in\mathscr{E}_L} (\nabla S^\ell)_{ij}(\nabla\alpha)_{ij} = \sum_{i\in\mathbb{T}_L} S_i^\ell(-\Delta\alpha)_i$$
$$= 2d\Big\{1 - \frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}\sum_{i\in\mathbb{T}_L} S_i^\ell e^{\mathrm{i}p\cdot i},$$

and therefore

$$\Big\langle \Big|\sum_{\{i,j\}\in\mathscr{E}_L}(\mathbf{S}_i - \mathbf{S}_j)\cdot(h_i - h_j)\Big|^2\Big\rangle_{L;\beta} = 4d^2\Big|1 - \frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big|^2\Big\langle\Big|\sum_{i\in\mathbb{T}_L} S_i^\ell e^{\mathrm{i}p\cdot i}\Big|^2\Big\rangle_{L;\beta}.$$

The inequality (10.47) thus implies that

$$\Big\langle\Big|\sum_{i\in\mathbb{T}_L} S_j^\ell e^{\mathrm{i}p\cdot i}\Big|^2\Big\rangle_{L;\beta} \le \frac{|\mathbb{T}_L|}{4d\beta}\Big\{1 - \frac{1}{2d}\sum_{j\sim 0}\cos(p\cdot j)\Big\}^{-1}.$$

Since, by translation invariance of $\mu_{L;\beta}$,

$$\Big\langle\Big|\sum_{i\in\mathbb{T}_L} S_j^\ell e^{\mathrm{i}p\cdot i}\Big|^2\Big\rangle_{L;\beta} = \sum_{i,j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot(j-i)}\langle S_i^\ell S_j^\ell\rangle_{L;\beta} = |\mathbb{T}_L|\sum_{j\in\mathbb{T}_L} e^{\mathrm{i}p\cdot j}\langle S_0^\ell S_j^\ell\rangle_{L;\beta},$$

the conclusion follows by summing over $\ell \in \{1,\dots,\nu\}$ to recover the inner product. □

10.6 Bibliographical remarks

There exist several nice reviews on reflection positivity, which can serve as complements to what is discussed in this chapter and provide additional examples of applications. These include the reviews by Shosman [305] and Biskup [22] and the books by Sinai [312, Chapter 3], Prum [282, Chapter 7] and Georgii [134, Part IV]. The present chapter was largely inspired by the presentation in [22].

Reflection positivity. Reflection positivity was first introduced in the context of constructive quantum field theory, where it plays a fundamental role. Its use in equilibrium statistical mechanics started in the late 1970s, see [118, 98, 157, 115, 116].

Infrared bound and long-range order in O(N) models. The infrared bound, Theorem 10.24, was first proved by Fröhlich, Simon and Spencer in [118]. In this paper, among other applications, they use this bound to establish existence of spontaneous magnetization in O(N) models on $\mathbb{Z}^d$, $d \geq 3$, at low temperature (Theorem 10.1 in this chapter).

Chessboard estimate and the anisotropic XY model. The chessboard estimate, in the form stated in Theorem 10.11, was first proved by Fröhlich and Lieb [117]. There were however earlier versions of it, see [134, Notes on Chapter 17]. The application to the anisotropic XY model, Theorem 10.2, was first established using other methods in [226] and [202]. The first proof relying on the chessboard estimate appeared in [117] and served as a basis for Section 10.2.

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook