9 Models with Continuous Symmetry

In Chapter 3, we have analyzed the phase transition occurring in the Ising model. We have seen, in particular, that the change of behavior observed (when $h = 0$ and $d \ge 2$) as the inverse temperature β crosses the critical value $\beta_c = \beta_c(d)$ was associated to the spontaneous breaking of a discrete symmetry: when $\beta < \beta_c$, there is a unique infinite-volume Gibbs measure, invariant under a global spin flip (that is, interchange of all + and − spins); on the contrary, when $\beta > \beta_c$, uniqueness fails, and we proved the existence of two distinct infinite-volume Gibbs measures $\mu^+_{\beta,0}$ and $\mu^-_{\beta,0}$, which are not invariant under a global spin flip (since $\langle\sigma_0\rangle^+_{\beta,0} > 0 > \langle\sigma_0\rangle^-_{\beta,0}$).

Our goal in the present chapter is to analyze the effect of the existence of a *continuous* symmetry (that is, corresponding to a Lie group) on phase transitions. We will see that, in one- and two-dimensional models, a global continuous symmetry is in general never spontaneously broken. In this sense, continuous symmetries are more robust.

9.1 O(N)-symmetric models

The systems we consider in this chapter are models for which the spins are N-dimensional unit vectors, living at the vertices of $\mathbb{Z}^d$.

Let us thus fix some $N \in \mathbb{N}$, and define the single-spin space

$$\Omega_0 \stackrel{\text{def}}{=} \left\{ v \in \mathbb{R}^N : \|v\|_2 = 1 \right\} \equiv \mathbb{S}^{N-1} .$$

Correspondingly, the set of configurations in a finite set $\Lambda \Subset \mathbb{Z}^d$ (resp. in $\mathbb{Z}^d$) is given by

$$\Omega_\Lambda \stackrel{\text{def}}{=} \Omega_0^\Lambda \qquad (\text{resp. } \Omega = \Omega_0^{\mathbb{Z}^d}) .$$

To each vertex $i \in \mathbb{Z}^d$, we associate the random variable $\mathbf{S}_i = (S_i^1, S_i^2, \dots, S_i^N)$ defined by

$$\mathbf{S}_i(\omega) \stackrel{\text{def}}{=} \omega_i ,$$

which we call, as usual, the **spin** at i. We assume that spins interact only with their nearest-neighbors and, most importantly, that *the interaction is invariant under* *simultaneous rotations of all the spins.* We can therefore assume that the interaction between two spins located at nearest-neighbor vertices i and j contribute an amount to the total energy which is a function of their scalar product $\mathbf{S}_i \cdot \mathbf{S}_j$.

Definition 9.1. *Let* $W : [-1,1] \to \mathbb{R}$. *The* ***Hamiltonian of an*** $\mathsf{O}(N)$***-symmetric model*** *in* $\Lambda \Subset \mathbb{Z}^d$ *is defined by*

$$\mathscr{H}_{\Lambda;\beta} \stackrel{\text{def}}{=} \beta \sum_{\{i,j\}\in\mathscr{E}^{\text{b}}_\Lambda} W(\mathbf{S}_i \cdot \mathbf{S}_j)\,. \tag{9.1}$$

A particularly important class of models is given by the $\mathsf{O}(N)$ **models**, for which $W(x) = -x$:

$$\mathscr{H}_{\Lambda;\beta} = -\beta \sum_{\{i,j\}\in\mathscr{E}^{\text{b}}_\Lambda} \mathbf{S}_i \cdot \mathbf{S}_j\,. \tag{9.2}$$

With this choice, different values of N then lead to different models, some of which have their own names. When $N = 1$, $\Omega_0 = \{\pm\mathbf{e}_1\}$ can be identified with $\{\pm 1\}$, so that the $O(1)$-model reduces to the Ising model. The case $N = 2$ corresponds to the ***XY* model**, and $N = 3$ corresponds to the **(classical) Heisenberg model**.

Given the Hamiltonian (9.1), we can define finite-volume Gibbs distributions and Gibbs measures in the usual way. We use the measurable structures on Ω_Λ and Ω, denoted respectively $\mathscr{F}_\Lambda$ and $\mathscr{F}$, introduced in Section 6.10. The reference measure for the spin at vertex i is the Lebesgue measure on $\mathbb{S}^{N-1}$, denoted simply $\mathrm{d}\omega_i$.

Given $\Lambda \Subset \mathbb{Z}^d$ and $\eta \in \Omega$, the Gibbs distribution of the $\mathsf{O}(N)$-symmetric models in Λ with boundary condition η is the probability measure $\mu^\eta_{\Lambda;\beta}$ on $(\Omega, \mathscr{F})$ defined by

$$\forall A \in \mathscr{F}, \quad \mu^\eta_{\Lambda;\beta}(A) \stackrel{\text{def}}{=} \int_{\Omega_\Lambda} \frac{e^{-\mathscr{H}_{\Lambda;\beta}(\omega_\Lambda\eta_{\Lambda^c})}}{\mathbf{Z}^\eta_{\Lambda;\beta}} 1_A(\omega_\Lambda\eta_{\Lambda^c}) \prod_{i\in\Lambda} \mathrm{d}\omega_i\,,$$

where the partition function is given by

$$\mathbf{Z}^\eta_{\Lambda;\beta} \stackrel{\text{def}}{=} \int_{\Omega_\Lambda} e^{-\mathscr{H}_{\Lambda;\beta}(\omega_\Lambda\eta_{\Lambda^c})} \prod_{i\in\Lambda} \mathrm{d}\omega_i\,.$$

As in Chapter 6, we can consider the specification associated to the kernels $(A,\eta) \mapsto \pi_\Lambda(A\,|\,\eta) \stackrel{\text{def}}{=} \mu^\eta_{\Lambda;\beta}(A)$, $\Lambda \Subset \mathbb{Z}^d$, and then denote by $\mathscr{G}(N)$ the set of associated infinite-volume Gibbs measures. (To lighten the notations, we do not indicate the dependence of $\mathscr{G}(N)$ on the choice of W and β.) Notice that Ω_0, and hence Ω, are compact, and so the results of Section 6.10.2 guarantee that the model has at least one infinite-volume Gibbs measure: $\mathscr{G}(N) \neq \varnothing$.

Even though our results below are stated in terms of infinite-volume Gibbs measures, the estimates in the proofs are actually valid for large finite systems. Therefore, readers not comfortable with the DLR formalism of Chapter 6 should be able to understand most of the content of this chapter.

9.1.1 Overview

Inspired by what was done for the Ising model, one of our goals in this chapter will be to determine whether suitable boundary conditions can lead to **orientational long-range order**, that is, whether spins align macroscopically along a preferred direction, giving rise to a non-zero spontaneous magnetization. For the sake of

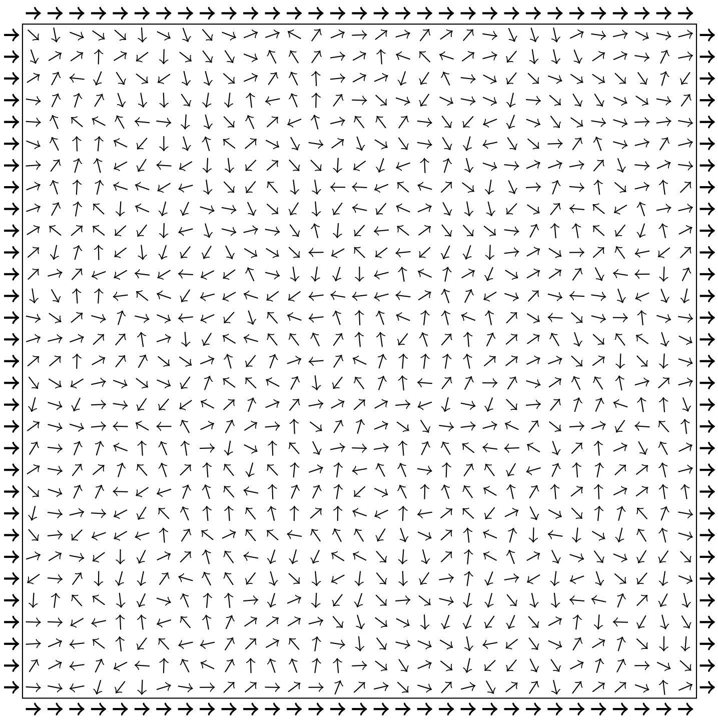


Figure 9.1: A configuration of the two-dimensional XY model with $\mathbf{e}_1$ boundary conditions, at high temperature: $\beta = 0.7$.

concreteness, one can think of those Gibbs measures obtained by fixing a boundary condition η and taking the thermodynamic limit:

$$\mu^{\eta}_{\mathrm{B}(n);\beta} \Rightarrow \mu\,.$$

- *Dimensions* 1 *and* 2*;* $N \geq 2$. We will see that, when $d = 1$ or $d = 2$, under any measure $\mu \in \mathscr{G}(N)$, the distribution $\mu(\mathbf{S}_i \in \cdot)$ of each individual spin $\mathbf{S}_i$ is *uniform* on $\mathbb{S}^{N-1}$; in particular,

  $$\langle \mathbf{S}_i \rangle_\mu = \mathbf{0}\,.$$

  Therefore, even in dimension 2 at very low temperature, orientational order *does not* occur in O(N)-symmetric models. This is due, as will be seen, to the existence of order-destroying excitations of arbitrarily low energy (Proposition 9.7 below). The above will actually be a consequence of a more general result, the celebrated *Mermin–Wagner Theorem* (Theorem 9.2).

- *Dimension* 2*;* $N = 2$. Even though there is no orientational long-range order when $d = 2$, we will see (but not prove) that there is **quasi-long-range order** at low temperatures: the 2-point correlation function decays only algebraically with the distance:

  $$\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_\mu \approx \| j - i \|_2^{-C/\beta}\,,$$

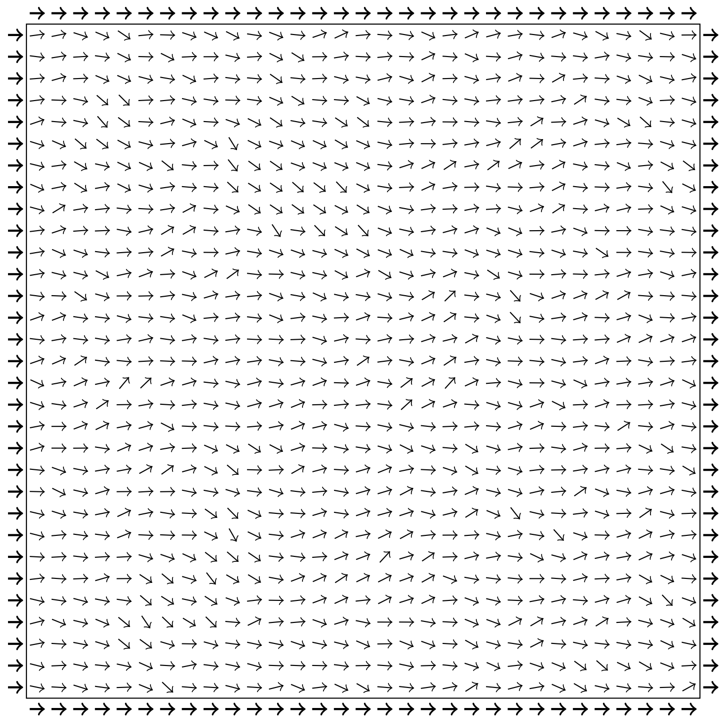


Figure 9.2: A configuration of the two-dimensional XY model with $\mathbf{e}_1$ boundary conditions, at low temperature: $\beta = 5$. In spite of the apparent ordering of the spins, we shall prove that, in large enough systems, there is no orientational long-range order at any temperature.

for some $C > 0$. This is in sharp contrast with $d = 1$ (for all $\beta \geq 0$), or with $d \geq 2$ at sufficiently high temperatures, where the 2-point correlation functions decay exponentially.

- *Dimensions $d \geq 3$; $N \geq 2$.* It turns out that spontaneous breaking of the continuous symmetry does indeed occur at low enough temperatures in the $\mathsf{O}(N)$ models, in dimensions $d \geq 3$, as will be discussed in Remark 9.5, and proved later in Chapter 10 (Theorem 10.25).

Additional information, including some outstanding open problems, can be found in the complements.

9.2 Absence of continuous symmetry breaking

Symmetries in the study of Gibbs measures were described in Section 6.6. Let $R \in \mathsf{SO}(N)$ be any rotation on $\mathbb{S}^{N-1}$; R can of course be represented as an $N \times N$ orthogonal matrix of determinant 1. We can use R to define a global rotation r on a configuration $\omega \in \Omega$ by

$$(\mathsf{r}\omega)_i \stackrel{\text{def}}{=} R\omega_i, \quad \forall i \in \mathbb{Z}^d.$$

A global rotation can also be defined on events $A \in \mathscr{F}$, by letting $\mathsf{r}A \stackrel{\text{def}}{=} \{\mathsf{r}\omega : \omega \in A\}$, as well as on functions and probability measures:

$$\mathsf{r}f(\omega) \stackrel{\text{def}}{=} f(\mathsf{r}^{-1}\omega), \quad \mathsf{r}(\mu)(A) \stackrel{\text{def}}{=} \mu(\mathsf{r}^{-1}A).$$

We shall often write $\mathsf{r} \in \mathsf{SO}(N)$, meaning that r is a global rotation associated to some element of $\mathsf{SO}(N)$.

By construction, the Hamiltonian (9.1) is invariant under global rotations of the spins: for all $\mathsf{r} \in \mathsf{SO}(N)$,

$$\mathscr{H}_{\Lambda;\beta}(\mathsf{r}\omega) = \mathscr{H}_{\Lambda;\beta}(\omega) \quad \forall \omega \in \Omega. \tag{9.3}$$

Therefore, as a consequence of Theorem 6.45, $\mathscr{G}(N)$ is invariant under r: if $\mu \in \mathscr{G}(N)$, then $\mathsf{r}(\mu) \in \mathscr{G}(N)$. What Theorem 6.45 does not say is whether $\mathsf{r}(\mu)$ coincides with μ. A remarkable fact is that this is necessarily the case when $d = 1, 2$.

Theorem 9.2 (Mermin–Wagner Theorem)**.** *Assume that $N \geq 2$, and that W is twice continuously differentiable. Then, when $d = 1$ or 2, all infinite-volume Gibbs measures are invariant under the action of $\mathsf{SO}(N)$: for all $\mu \in \mathscr{G}(N)$,*

$$\mathsf{r}(\mu) = \mu, \qquad \forall \mathsf{r} \in \mathsf{SO}(N).$$

Of course, the claim is wrong when $N = 1$, since in this case the global spin flip symmetry can be broken at low temperature in $d = 2$. Let us make a few important comments.

Theorem 9.2 implies that, in an infinite system whose equilibrium properties are described by a Gibbs measure $\mu \in \mathscr{G}(N)$, the distribution of each individual spin $\mathbf{S}_i$ is uniform on $\mathbb{S}^{N-1}$. Namely, let $I \subset \mathbb{S}^{N-1}$; for any $\mathsf{r} \in \mathsf{SO}(N)$,

$$\mu(\mathbf{S}_i \in I) = \mathsf{r}(\mu)(\mathbf{S}_i \in I) = \mu(\mathbf{S}_i \in \mathsf{r}^{-1}(I)).$$

As a consequence, spontaneous magnetization (that is, some global orientation observed at the macroscopic level) cannot be observed in low-dimensional systems with continuous symmetries, even at very low temperature:

$$\langle \mathbf{S}_0 \rangle_\mu = \mathbf{0}, \qquad (d = 1, 2). \tag{9.4}$$

The above is in sharp contrast with the symmetry breaking observed in the two-dimensional Ising model at low temperature. There, when $h = 0$, the Hamiltonian was invariant under the discrete global spin flip, $\tau_{\text{g.s.f.}}$, but $\tau_{\text{g.s.f.}}(\mu^+_{\beta,0}) = \mu^-_{\beta,0} \neq \mu^+_{\beta,0}$ when $\beta > \beta_c(d)$.

Although this was not stated explicitly above, the $\mathsf{SO}(N)$-invariance of the infinite-volume Gibbs measures also implies absence of orientational long-range order. Namely, let $k \in \mathbb{Z}^d$ be fixed, far from the origin. Let n be large, but small enough to have $k \in \mathsf{B}(n)^c$ (for example: $n = \|k\|_\infty - 1$). If $\mu \in \mathscr{G}(N)$, then the DLR compatibility conditions $\mu = \mu\pi_\Lambda$, $\forall \Lambda \Subset \mathbb{Z}^d$, imply that

$$\langle \mathbf{S}_0 \cdot \mathbf{S}_k \rangle_\mu = \int \langle \mathbf{S}_0 \cdot \mathbf{S}_k \rangle^\eta_{\mathsf{B}(n);\beta} \, \mu(\mathrm{d}\eta) = \int \langle \mathbf{S}_0 \rangle^\eta_{\mathsf{B}(n);\beta} \cdot \mathbf{S}_k(\eta) \, \mu(\mathrm{d}\eta). \tag{9.5}$$

We shall actually obtain a quantitative version of (9.4) in Proposition 9.7, a consequence of which will be that $\lim_{n\to\infty} \langle \mathbf{S}_0 \rangle^\eta_{\mathsf{B}(n);\beta} = \mathbf{0}$, uniformly in the boundary condition η (see Exercise 9.5). Therefore, by dominated convergence,

$$\langle \mathbf{S}_0 \cdot \mathbf{S}_k \rangle_\mu \to 0 \quad \text{when } \|k\|_\infty \to \infty.$$

Exercise 9.1. *Prove that this also implies that*

$$\lim_{n\to\infty} \langle \| m_{\mathsf{B}(n)} \|_2^2 \rangle_\mu = 0\,,$$

where $m_{\mathsf{B}(n)} \stackrel{\text{def}}{=} |\mathsf{B}(n)|^{-1} \sum_{i\in\mathsf{B}(n)} \mathbf{S}_i$ *is the* ***magnetization density*** *in* $\mathsf{B}(n)$.

Remark 9.3. As explained above, Theorem 9.2 implies the absence of spontaneous magnetization and long-range order. Nevertheless, this theorem does not imply that there is a unique infinite-volume Gibbs measure [1]. ⋄

Remark 9.4. It is interesting to see what happens if one considers perturbations of the above models in which the continuous symmetry is explicitly broken. As an example, consider the **anisotropic** $\boldsymbol{XY}$ **model**, which has the same single-spin space as the XY model, but a more general Hamiltonian

$$\mathscr{H}_{\Lambda;\beta,\alpha} = -\beta \sum_{\{i,j\}\in\mathscr{E}_\Lambda^{\mathrm{b}}} \{S_i^1 S_j^1 + \alpha S_i^2 S_j^2\}$$

depending on an **anisotropy parameter** $\alpha \in [0,1]$. Observe that this Hamiltonian is $\mathsf{SO}(2)$-invariant only when $\alpha = 1$, in which case one recovers the usual XY model.

It turns out that there is always orientational long-range order at sufficiently low temperatures when $\alpha \in [0,1)$ and $d \geq 2$ (in $d = 1$, uniqueness always holds thanks to a suitable generalization of Theorem 6.40). Indeed, using reflection positivity, we will prove in Theorem 10.18 that, for any $\alpha \in [0,1)$ and all β sufficiently large, there exist at least two infinite-volume Gibbs measures μ^+ and μ^- such that

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

This shows that having continuous spins is not sufficient to prevent orientational long-range order in low dimensions: the presence of a continuous symmetry is essential. ⋄

Remark 9.5. Theorem 9.2 is restricted to dimensions 1 and 2. Let us briefly mention what happens in higher dimensions, restricting the discussion to the XY model: as soon as $d \geq 3$, for all β sufficiently large, there exist a number $m(\beta) > 0$ and a family of extremal infinite-volume Gibbs measures $(\mu_\beta^\psi)_{-\pi<\psi\leq\pi}$ such that

$$\langle \mathbf{S}_0 \rangle_\beta^\psi = m(\beta)\,(\cos\psi, \sin\psi)\,.$$

The proof of this claim (actually, for all values of N) will be given in Chapter 10.

Additional information on the role of the dimension, as well as on the corresponding results for $O(N)$ models with more general (not necessarily nearest-neighbor) interactions will be provided in Section 9.6.2. ⋄

Remark 9.6. Both the proof of Theorem 9.2 and the heuristic argument below rely in a seemingly crucial way on the smoothness of the interaction W. The reader might thus wonder whether the latter is a necessary condition. It turns out that Theorem 9.2 can be extended to all piecewise continuous interactions W; see Section 9.6.2. ⋄

9.2.1 Heuristic argument

Before turning to the proof of Theorem 9.2, let us emphasize a crucial difference between continuous and discrete spin systems. For this heuristic discussion, W

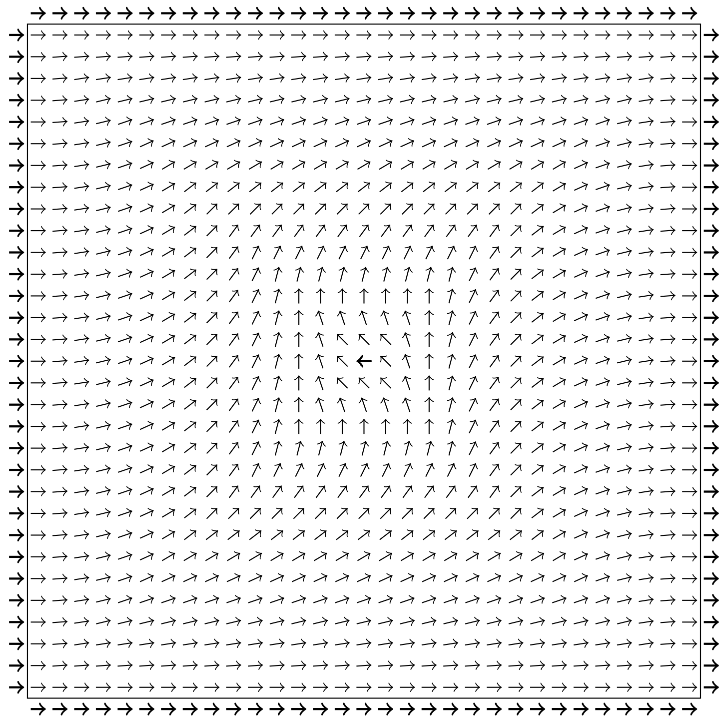


Figure 9.3: The spin wave ω^{SW} (see (9.6)), which flips the spin at the center of $\mathsf{B}(n)$, at a cost that can be made arbitrarily small by taking n large enough.

can be any twice continuously differentiable function, but we only consider the case $N = 2$ and, mostly, $d = 2$.

Let us therefore consider a two-dimensional $\mathsf{O}(2)$-symmetric model in the box $\mathsf{B}(n)$, with boundary condition $\eta_i = \mathbf{e}_1 = (1,0)$ for all $i \in \mathsf{B}(n)^{\mathrm{c}}$. If we assume that W is decreasing on $[-1,1]$, then the ground state (that is, the configuration with the lowest energy) is the one that agrees everywhere with the boundary condition. We denote it by $\omega^{\mathbf{e}_1}$: $\omega^{\mathbf{e}_1}_i = \mathbf{e}_1$ for all i. We would like to determine the energetic cost of flipping the spin in the middle of the box. More precisely: among all configurations ω that agree with the boundary condition outside $\mathsf{B}(n)$ but in which the spin at the origin is flipped, $\omega_0 = -\mathbf{e}_1$, which one minimizes the Hamiltonian, and what is the corresponding value of the energy?

Remember that for the two-dimensional Ising model ($\mathsf{O}(N)$ with $N = 1$), the energetic cost required to flip the spin at the center of the box, with + boundary condition, is at least 8β (since the shortest Peierls contour surrounding the origin has length 4), uniformly in the size of the box. Due to the presence of a continuous symmetry, the situation is radically different for the two-dimensional $\mathsf{O}(2)$-symmetric model: by slowly rotating the spins between the boundary and the center of the box, the spin at the origin can be flipped at an arbitrarily low cost (see Figure 9.3).

To understand this quantitatively, let us describe each configuration by the family $(\vartheta_i)_{i\in\mathbb{Z}^2}$, where $\vartheta_i \in (-\pi,\pi]$ is the angle such that $\mathbf{S}_i = (\cos\vartheta_i, \sin\vartheta_i)$. Let us also write $V(\theta) = W(\cos(\theta))$, so that

$$\mathscr{H}_{\mathsf{B}(n);\beta} = \beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}} W(\mathbf{S}_i \cdot \mathbf{S}_j) = \beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}} V(\vartheta_j - \vartheta_i)\,.$$

Let us consider the configuration $\omega_i^{\mathrm{SW}} = (\cos\theta_i^{\mathrm{SW}}, \sin\theta_i^{\mathrm{SW}})$, where

$$\theta_i^{\mathrm{SW}} \stackrel{\text{def}}{=} \Bigl(1 - \frac{\log(1+\|i\|_\infty)}{\log(1+n)}\Bigr)\pi, \quad i \in \mathsf{B}(n)\,, \tag{9.6}$$

and $\theta_i^{\mathrm{SW}} = 0$ for $i \not\in \mathsf{B}(n)$ (see Figure 9.3). Clearly, the only nonzero contributions to $\mathscr{H}_{\mathsf{B}(n);\beta}(\omega^{\mathrm{SW}})$ are those due to pairs of neighboring vertices i and j such that $\|i\|_\infty = \|j\|_\infty - 1$. For each such pair,

$$\theta_i^{\mathrm{SW}} - \theta_j^{\mathrm{SW}} = \pi \frac{\log(1+\frac{1}{\|j\|_\infty})}{\log(1+n)} \le \frac{\pi}{\log(1+n)} \frac{1}{\|j\|_\infty}\,.$$

Therefore, if n is large, each term $V(\theta_i^{\mathrm{SW}} - \theta_j^{\mathrm{SW}})$ can be estimated using a Taylor expansion of V at $\theta = 0$. Moreover, since V is twice continuously differentiable, there exists a constant C such that

$$\sup_{\theta\in(-\pi,\pi]} V''(\theta) \le C\,, \tag{9.7}$$

and we have, since $V'(0) = 0$,

$$V(\theta_i^{\mathrm{SW}} - \theta_j^{\mathrm{SW}}) \le V(0) + \tfrac{1}{2}C(\theta_i^{\mathrm{SW}} - \theta_j^{\mathrm{SW}})^2 \le V(0) + \frac{C\pi^2}{2\bigl(\log(1+n)\bigr)^2} \frac{1}{\|j\|_\infty^2}\,.$$

Summing over the contributing pairs of neighboring vertices i and j,

$$0 \le \mathscr{H}_{\mathsf{B}(n)}(\omega^{\mathrm{SW}}) - \mathscr{H}_{\mathsf{B}(n)}(\omega^{\mathbf{e}_1}) \le \frac{C\beta\pi^2}{2\bigl(\log(1+n)\bigr)^2} \sum_{r=1}^{n+1} 4(2r-1)\frac{1}{r^2} \le \frac{8C\beta\pi^2}{\log(1+n)}\,,$$

which indeed tends to 0 when $n \to \infty$.

It is the existence of configurations like ω^{SW}, representing collective excitations of arbitrarily low energy, called **spin waves**, which renders impossible the application of a naive Peierls-type argument. We shall see that spin waves are the key ingredient in the proof of the Mermin–Wagner Theorem, given in Section 9.2.2.

In the above argument, we only flipped the spin located at the center of the box. It is easy to check that similar spin waves can also be constructed if one wants to flip all the spins in an extended region.

Exercise 9.2. ($d = 2$) *Adapt the previous computation to show that the lowest energy required to flip all the spins in a smaller box* $\mathsf{B}(\ell) \subset \mathsf{B}(n)$ *goes to zero when* $n \to \infty$.

Let us now briefly discuss what happens in dimensions $d \neq 2$. First, in the next exercise, the reader is encouraged to check that one can also construct a spin wave as above in dimension 1 (actually, one can take a much simpler one in that case).

Exercise 9.3. *Construct a suitable spin wave for the one-dimensional* $\mathsf{O}(2)$-*symmetric model.*

In higher dimensions, $d \geq 3$, it is not possible anymore to repeat the argument given above. In the following exercise (see also Lemma 9.8), the reader is asked to show that the second-order term in the Taylor expansion remains bounded away from zero as $n \to \infty$.

Exercise 9.4. *Let $f : \mathbb{Z}_{\geq 0} \to \mathbb{R}$ be such that $f(r) = 1$ if $0 \leq r \leq \ell$ and $f(r) = 0$ if $r \geq n$. Show that*

$$\sum_{\{i,j\} \in \mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}} \big(f(\|i\|_\infty) - f(\|j\|_\infty) \big)^2 \geq \Big\{ \sum_{k \geq \ell} k^{-(d-1)} \Big\}^{-1}.$$

Conclude that, in contrast to the case $d = 2$ (Exercise 9.2), the minimal energy required to flip all spins in the box $\mathsf{B}(\ell)$ does not tend to zero when $n \to \infty$ when $d \geq 3$ (it is not even bounded in ℓ). Hint: *To derive the inequality, use the Cauchy–Schwarz inequality for functions $g : \{0, \dots, n\} \to \mathbb{R}$.*

9.2.2 Proof of the Mermin–Wagner Theorem for $N = 2$

We first give a proof of the result in the case $N = 2$, and then use it to address the general case in Section 9.2.3. We write $\mathbf{S}_i = (\cos\vartheta_i, \sin\vartheta_i)$ and set $V(\theta) \stackrel{\text{def}}{=} W(\cos(\theta))$, as in the heuristic argument above.

Let $\mu \in \mathscr{G}(2)$ and let $\mathsf{r}_\psi \in \mathsf{SO}(2)$ denote the rotation of angle $\psi \in (-\pi, \pi]$. To show that $\mathsf{r}_\psi(\mu) = \mu$, we shall show that $\langle f \rangle_\mu = \langle \mathsf{r}_\psi f \rangle_\mu$ for each local bounded measurable function f. But, by the DLR compatibility conditions, we can write, for any $\Lambda \Subset \mathbb{Z}^d$,

$$|\langle f \rangle_\mu - \langle \mathsf{r}_\psi f \rangle_\mu| = \Big| \int \big\{ \langle f \rangle^\eta_{\Lambda;\beta} - \langle \mathsf{r}_\psi f \rangle^\eta_{\Lambda;\beta} \big\} \mu(\mathrm{d}\eta) \Big| \leq \int \big| \langle f \rangle^\eta_{\Lambda;\beta} - \langle \mathsf{r}_\psi f \rangle^\eta_{\Lambda;\beta} \big| \mu(\mathrm{d}\eta). \quad (9.8)$$

We study the differences $|\langle f \rangle^\eta_{\Lambda;\beta} - \langle \mathsf{r}_\psi f \rangle^\eta_{\Lambda;\beta}|$ quantitatively in the following proposition. In view of (9.8), Theorem 9.2 is a direct consequence of the following proposition:

Proposition 9.7. *Assume that $d = 1$ or $d = 2$ and fix $N = 2$. Under the hypotheses of Theorem 9.2, there exist constants c_1, c_2 such that, for any boundary condition $\eta \in \Omega$, any inverse temperature $\beta < \infty$, any angle $\psi \in (-\pi, \pi]$ and any $\ell \in \mathbb{Z}_{\geq 0}$,*

$$\big| \langle f \rangle^\eta_{\mathsf{B}(n);\beta} - \langle \mathsf{r}_\psi f \rangle^\eta_{\mathsf{B}(n);\beta} \big| \leq \beta^{1/2} |\psi| \, \|f\|_\infty \times \begin{cases} \frac{c_1}{\sqrt{n-\ell}} & \text{if } d = 1, \\ \frac{c_2 \sqrt{\ell}}{\sqrt{\log(n-\ell)}} & \text{if } d = 2, \end{cases} \quad (9.9)$$

for all $n > \ell$ and all bounded functions f such that $\operatorname{supp}(f) \subset \mathsf{B}(\ell)$.

In the sequel, we will use the notation $T_n(1) \stackrel{\text{def}}{=} \sqrt{n}$, $T_n(2) \stackrel{\text{def}}{=} \sqrt{\log n}$ when we want to treat the cases $d = 1$ and $d = 2$ simultaneously.

Exercise 9.5. *Deduce from* (9.9) *that, under $\mu^\eta_{\mathsf{B}(n);\beta}$, the distribution of ϑ_0 converges to the uniform distribution on* $(-\pi, \pi]$. *In particular, for any η,*

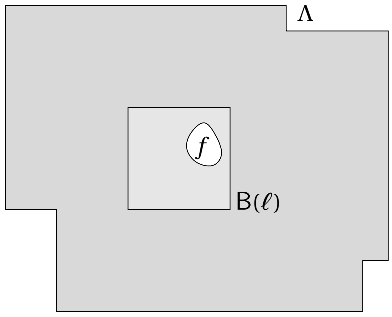
$$\lim_{n \to \infty} \big\| \langle \mathbf{S}_0 \rangle^\eta_{\mathsf{B}(n);\beta} \big\|_2 = 0.$$

Most of the proof of the bounds (9.9) does not depend on the shape of the system considered. So, let us first consider an arbitrary connected Λ, which will later be taken to be the box $\mathsf{B}(n)$. Our starting point is to express

$$\langle \mathsf{r}_\psi f\rangle^\eta_{\Lambda;\beta} = (\mathbf{Z}^\eta_{\Lambda;\beta})^{-1}\int f(\mathsf{r}_{-\psi}\omega)e^{-\mathscr{H}_{\Lambda;\beta}(\omega_\Lambda\eta_{\Lambda^c})}\prod_{i\in\Lambda}\mathrm{d}\omega_i$$

as the expectation of f under a modified distribution.

We let Λ and ℓ be large enough so that $\Lambda\supset\mathsf{B}(\ell)\supset\mathrm{supp}(f)$:



Let $\Psi:\mathbb{Z}^d\to(-\pi,\pi]$ satisfy $\Psi_i=\psi$ for all $i\in\mathsf{B}(\ell)$, and $\Psi_i=0$ for all $i\not\in\Lambda$. An explicit choice for Ψ will be made later. Let $\mathsf{t}_\Psi:\Omega\to\Omega$ denote the transformation under which

$$\vartheta_i(\mathsf{t}_\Psi\omega)=\vartheta_i(\omega)+\Psi_i\,,\quad\forall\omega\in\Omega\,.$$

That is, t_Ψ acts as the identity on spins located outside Λ and as the rotation r_ψ on spins located inside $\mathsf{B}(\ell)$. Observe that $\mathsf{t}_{-\Psi}=\mathsf{t}_\Psi^{-1}$. Now, since $\mathsf{t}_{-\Psi}\omega$ and $\mathsf{r}_{-\psi}\omega$ coincide on $\mathrm{supp}(f)\subset\mathsf{B}(\ell)$,

$$\begin{aligned}\int f(\mathsf{r}_{-\psi}\omega)e^{-\mathscr{H}_{\Lambda;\beta}(\omega_\Lambda\eta_{\Lambda^c})}\prod_{i\in\Lambda}\mathrm{d}\omega_i &= \int f(\mathsf{t}_{-\Psi}\omega)e^{-\mathscr{H}_{\Lambda;\beta}(\omega_\Lambda\eta_{\Lambda^c})}\prod_{i\in\Lambda}\mathrm{d}\omega_i\\ &=\int f(\omega)e^{-\mathscr{H}_{\Lambda;\beta}(\mathsf{t}_\Psi(\omega_\Lambda\eta_{\Lambda^c}))}\prod_{i\in\Lambda}\mathrm{d}\omega_i\,.\end{aligned}$$

In the second equality, we used the fact that the mapping $\omega_\Lambda\mapsto(\mathsf{t}_{-\Psi}\omega)_\Lambda$ has a Jacobian equal to 1. Let $\langle\cdot\rangle^{\eta,\Psi}_{\Lambda;\beta}$ denote the expectation under the probability measure

$$\mu^{\eta;\Psi}_{\Lambda;\beta}(A)\stackrel{\text{def}}{=}(\mathbf{Z}^{\eta;\Psi}_{\Lambda;\beta})^{-1}\int_{\Omega_\Lambda}e^{-\mathscr{H}_{\Lambda;\beta}(\mathsf{t}_\Psi(\omega_\Lambda\eta_{\Lambda^c}))}\,1_A(\omega_\Lambda\eta_{\Lambda^c})\prod_{i\in\Lambda}\mathrm{d}\omega_i\,,\quad A\in\mathscr{F}\,.$$

Observe that, for the same reasons as above (the Jacobian being equal to 1 and the boundary condition being preserved by t_Ψ), the partition function is actually left unchanged:

$$\mathbf{Z}^{\eta;\Psi}_{\Lambda;\beta}=\mathbf{Z}^\eta_{\Lambda;\beta}\,. \tag{9.10}$$

We can then write $\langle\mathsf{r}_\psi f\rangle^\eta_{\Lambda;\beta}=\langle f\rangle^{\eta,\Psi}_{\Lambda;\beta}$, and therefore

$$\big|\langle f\rangle^\eta_{\Lambda;\beta}-\langle\mathsf{r}_\psi f\rangle^\eta_{\Lambda;\beta}\big|=\big|\langle f\rangle^\eta_{\Lambda;\beta}-\langle f\rangle^{\eta;\Psi}_{\Lambda;\beta}\big|,$$

which reduces the problem to comparing the expectation of f under the measures $\mu^\eta_{\Lambda;\beta}$ and $\mu^{\eta;\Psi}_{\Lambda;\beta}$.

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

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

The measures $\mu^{\eta}_{\Lambda;\beta}$ and $\mu^{\eta;\Psi}_{\Lambda;\beta}$ only differ by the "addition" of the spin wave Ψ. However, we saw in the Section 9.2.1 that the latter can be chosen such that its energetic cost is arbitrarily small. One would thus expect such excitations to proliferate in the system, and thus the two Gibbs distributions to be "very close" to each other. ⋄

One convenient way of measuring the "closeness" of two measures μ, ν is the **relative entropy**

$$\mathsf{h}(\mu\,|\,\nu) \stackrel{\text{def}}{=} \begin{cases} \langle \frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu} \rangle_\nu, & \text{if } \mu \ll \nu, \\ \infty & \text{otherwise,} \end{cases}$$

where $\frac{d\mu}{d\nu}$ is the Radon–Nikodym derivative of μ with respect to ν. The relevant properties of the relative entropy can be found in Appendix B.12. Particularly well-suited to our needs, **Pinsker's inequality**, see Lemma B.67, states that, for any measurable function f with $\|f\|_\infty \le 1$,

$$|\langle f \rangle_\mu - \langle f \rangle_\nu| \le \sqrt{2\mathsf{h}(\mu\,|\,\nu)}\,.$$

In our case, thanks to (9.10),

$$\frac{d\mu^{\eta}_{\Lambda;\beta}}{d\mu^{\eta,\Psi}_{\Lambda;\beta}}(\omega) = e^{\mathscr{H}_{\Lambda;\beta}(\mathsf{t}_\Psi\omega) - \mathscr{H}_{\Lambda;\beta}(\omega)}\,.$$

Using Pinsker's inequality,

$$\begin{aligned} |\langle f \rangle^{\eta}_{\Lambda;\beta} - \langle f \rangle^{\eta,\Psi}_{\Lambda;\beta}| &\le \|f\|_\infty \sqrt{2\mathsf{h}(\mu^{\eta}_{\Lambda;\beta}\,|\,\mu^{\eta;\Psi}_{\Lambda;\beta})} \\ &= \|f\|_\infty \sqrt{2\big\langle \mathscr{H}_{\Lambda;\beta}\circ \mathsf{t}_\Psi - \mathscr{H}_{\Lambda;\beta}\big\rangle^{\eta}_{\Lambda;\beta}}\,. \end{aligned} \tag{9.11}$$

A second-order Taylor expansion yields, using again (9.7),

$$\begin{aligned} \big\langle \mathscr{H}_{\Lambda;\beta}\circ \mathsf{t}_\Psi - \mathscr{H}_{\Lambda;\beta}\big\rangle^{\eta}_{\Lambda;\beta} &= \beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_\Lambda} \Big\langle V\big(\vartheta_j - \vartheta_i + \Psi_j - \Psi_i\big) - V\big(\vartheta_j - \vartheta_i\big)\Big\rangle^{\eta}_{\Lambda;\beta} \\ &\le \beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_\Lambda} \Big\{ \big\langle V'\big(\vartheta_j - \vartheta_i\big)\big\rangle^{\eta}_{\Lambda;\beta}(\Psi_j - \Psi_i) + \tfrac{C}{2}(\Psi_j - \Psi_i)^2 \Big\}\,. \end{aligned}$$

Note the parallel with the heuristic discussion of Section 9.2.1. There, however, the first-order terms trivially vanished. We need an alternative way to see that the same occurs here, since the contribution of these terms would be too large to prove our claim. In order to get rid of them, we use the following trick: since the relative entropy is always nonnegative (Lemma B.65), we can write

$$\mathsf{h}(\mu^{\eta}_{\Lambda;\beta}\,|\,\mu^{\eta;\Psi}_{\Lambda;\beta}) \le \mathsf{h}(\mu^{\eta}_{\Lambda;\beta}\,|\,\mu^{\eta;\Psi}_{\Lambda;\beta}) + \mathsf{h}(\mu^{\eta}_{\Lambda;\beta}\,|\,\mu^{\eta;-\Psi}_{\Lambda;\beta})\,. \tag{9.12}$$

The second term in the right-hand side of the latter expression can be treated as above, and gives rise to the same first-order terms *but with the opposite sign*. These thus cancel, and we are left with (remember the notation $(\nabla\Psi)_{ij} \stackrel{\text{def}}{=} \Psi_j - \Psi_i$)

$$\mathsf{h}(\mu^{\eta}_{\Lambda;\beta}\,|\,\mu^{\eta;\Psi}_{\Lambda;\beta}) \le C\beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_\Lambda} (\nabla\Psi)^2_{ij}\,. \tag{9.13}$$

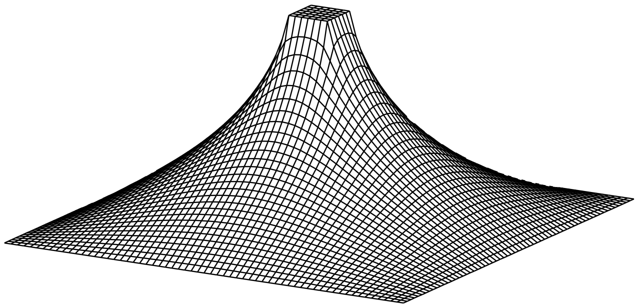


Figure 9.4: The minimizer of the Dirichlet energy given by (9.14). In this picture, $\Lambda = \mathsf{B}(30)$ and $\ell = 3$.

We will now choose the values Ψ_i, $i \in \Lambda \setminus \mathsf{B}(\ell)$. One possible choice is to take $\Lambda = \mathsf{B}(n)$, $n > \ell$, and to take for Ψ the spin wave introduced in Exercise 9.2. Nevertheless, since it is instructive, we shall provide a more detailed study of the problem of minimizing the above sum under constraints. This will also shed some light on the role played by the dimension d.

Define the **Dirichlet energy** (in $\Lambda \setminus \mathsf{B}(\ell)$) of a function $\Psi : \mathbb{Z}^d \to \mathbb{R}$ by

$$\mathscr{E}(\Psi) \stackrel{\text{def}}{=} \frac{1}{2} \sum_{\{i,j\} \in \mathscr{E}^{\mathrm{b}}_{\Lambda \setminus \mathsf{B}(\ell)}} (\nabla \Psi)^2_{ij} .$$

We will determine the minimizer of $\mathscr{E}(\Psi)$ among all functions Ψ such that $\Psi \equiv \psi$ on $\mathsf{B}(\ell)$ and $\Psi \equiv 0$ on Λ^{c}. As we shall see, in dimensions 1 and 2, the minimum value of that minimizer tends to zero when $\Lambda \uparrow \mathbb{Z}^d$.

Lemma 9.8. *The Dirichlet energy possesses a unique minimizer among all functions $u : \mathbb{Z}^d \to \mathbb{R}$ satisfying $u_i = 0$ for all $i \notin \Lambda$, and $u_i = 1$ for all $i \in \mathsf{B}(\ell)$. This minimizer is given by (see Figure 9.4)*

$$u^*_i \stackrel{\text{def}}{=} \mathbb{P}_i\big(X \textit{ enters } \mathsf{B}(\ell) \textit{ before exiting } \Lambda\big), \tag{9.14}$$

where $X = (X_k)_{k \geq 0}$ is the symmetric simple random walk on $\mathbb{Z}^d$ and $\mathbb{P}_i(X_0 = i) = 1$. Moreover,

$$\mathscr{E}(u^*) = d \sum_{j \in \partial^{\mathrm{int}} \mathsf{B}(\ell)} \mathbb{P}_j\big(X \textit{ exits } \Lambda \textit{ before returning to } \mathsf{B}(\ell)\big). \tag{9.15}$$

Proof of Lemma 9.8: Let us first characterize the critical points of $\mathscr{E}$. Namely, assume u is a critical point of $\mathscr{E}$, satisfying the constraints. Then we must have

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathscr{E}(u + s\delta)\Big|_{s=0} = 0, \tag{9.16}$$

for all perturbations $\delta : \mathbb{Z}^d \to \mathbb{R}$ such that $\delta_i = 0$ for all $i \notin \Lambda \setminus \mathsf{B}(\ell)$. However, a simple computation yields

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathscr{E}(u + s\delta)\Big|_{s=0} = \sum_{\{i,j\} \in \mathscr{E}^{\mathrm{b}}_{\Lambda \setminus \mathsf{B}(\ell)}} (\nabla u)_{ij} (\nabla \delta)_{ij} ,$$

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

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

and, by the Discrete Green Identity (8.14),

$$\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\Lambda\setminus \mathsf{B}(\ell)}} (\nabla u)_{ij}(\nabla\delta)_{ij} = -\sum_{i\in\Lambda\setminus\mathsf{B}(\ell)} \delta_i(\Delta u)_i + \sum_{\substack{i\in\Lambda\setminus\mathsf{B}(\ell)\\ j\notin\Lambda\setminus\mathsf{B}(\ell),j\sim i}} \delta_j(\nabla u)_{ij}. \tag{9.17}$$

The second sum in the right-hand side vanishes since $\delta_j = 0$ outside $\Lambda\setminus\mathsf{B}(\ell)$. In order for the first sum to be equal to zero for all δ, Δu must vanish everywhere on $\Lambda\setminus\mathsf{B}(\ell)$. This shows that the minimizer we are after is harmonic on $\Lambda\setminus\mathsf{B}(\ell)$. We know from Lemma 8.15 that the solution to the Dirichlet problem on $\Lambda\setminus\mathsf{B}(\ell)$ with boundary condition η is unique and given by

$$u^*_i \stackrel{\text{def}}{=} \mathbb{E}_i[\eta_{X_{\tau_{(\Lambda\setminus\mathsf{B}(\ell))^{\mathrm{c}}}}}], \tag{9.18}$$

where $\mathbb{P}_i$ is the law of the simple random walk on $\mathbb{Z}^d$ with initial condition $X_0 = i$. Using our boundary condition ($\eta_i = 1$ on $\mathsf{B}(\ell)$, $\eta_i = 0$ on Λ^{c}), we easily write u^*_i as in (9.14).

We still have to check that u^* is actually a minimizer of the Dirichlet energy. But this follows from (9.17), since, for all δ as above, the latter implies that $\mathscr{E}(u^*+\delta) = \mathscr{E}(u^*)+\mathscr{E}(\delta) \geq \mathscr{E}(u^*)$.

Finally, using now (9.17) with $u = \delta = u^*$,

$$\begin{aligned}
\mathscr{E}(u^*) &= \tfrac{1}{2}\sum_{i\in\Lambda\setminus\mathsf{B}(\ell)}\ \sum_{j\in\mathsf{B}(\ell),j\sim i}(u^*_j - u^*_i)\\
&= \tfrac{1}{2}\sum_{j\in\mathsf{B}(\ell)}\ \sum_{\substack{i\in\partial^{\mathrm{ext}}\mathsf{B}(\ell)\\ i\sim j}} \mathbb{P}_i\big(X \text{ exits } \Lambda \text{ before hitting } \mathsf{B}(\ell)\big)\\
&= \tfrac{1}{2}\sum_{j\in\mathsf{B}(\ell)}\sum_{i\sim j}\mathbb{P}_i\big(X \text{ exits } \Lambda \text{ before hitting } \mathsf{B}(\ell)\big)\\
&= d\sum_{j\in\partial^{\mathrm{int}}\mathsf{B}(\ell)} \mathbb{P}_j\big(X \text{ exits } \Lambda \text{ before returning to } \mathsf{B}(\ell)\big),
\end{aligned}$$

where we used the Markov property for the fourth equality. □

We can now complete the proof of the Mermin–Wagner Theorem for $N = 2$:

Proof of Proposition 9.7: Take $\Lambda = \mathsf{B}(n)$. Let u^* be the minimizer (9.18) and set $\Psi = \psi u^*$. Observe that this choice of Ψ has all the required properties and that $\mathscr{E}(\Psi) = \psi^2\mathscr{E}(u^*)$. Using (9.11) and (9.13), we thus have

$$\big|\langle f\rangle^{\eta}_{\mathsf{B}(n);\beta} - \langle \mathsf{r}_\psi f\rangle^{\eta}_{\mathsf{B}(n);\beta}\big| \leq \|f\|_\infty\sqrt{4C\beta\psi^2\mathscr{E}(u^*)}.$$

Since

$$\begin{aligned}
&\mathbb{P}_j\big(X \text{ exits } \mathsf{B}(n) \text{ before returning to } \mathsf{B}(\ell)\big)\\
&\qquad\leq \mathbb{P}_j\big(X \text{ exits } \mathsf{B}(n-\ell)+j \text{ before returning to } j\big)\\
&\qquad= \mathbb{P}_0\big(X \text{ exits } \mathsf{B}(n-\ell) \text{ before returning to } 0\big),
\end{aligned}$$

we finally get

$$\begin{aligned}
\big|\langle f\rangle^{\eta}_{\mathsf{B}(n);\beta} - \langle \mathsf{r}_\psi f\rangle^{\eta}_{\mathsf{B}(n);\beta}\big| &\leq \|f\|_\infty\sqrt{4Cd\beta\psi^2\,|\partial^{\mathrm{int}}\mathsf{B}(\ell)|}\\
&\qquad\times\mathbb{P}_0\big(X \text{ exits } \mathsf{B}(n-\ell) \text{ before returning to } 0\big)^{1/2}.
\end{aligned}$$

In dimensions $d = 1$ and $d = 2$, recurrence of the symmetric simple random walk implies that the latter probability goes to zero as $n\to\infty$. The rate at which this occurs is given in Theorem B.74. □

9.2.3 Proof of the Mermin–Wagner theorem for $N \geq 3$

To prove Theorem 9.2 when $N \geq 3$, we essentially reduce the problem to the case $N = 2$. The main observation is that, given an arbitrary rotation $R \in \mathsf{SO}(N)$, there exists an orthonormal basis, an integer $n \leq N/2$ and n numbers $\psi_i \in (-\pi,\pi]$, such that R can be represented as a block diagonal matrix of the following form [2]:

$$\begin{pmatrix} M(\psi_1) & & & & \\ & M(\psi_2) & & & \\ & & \ddots & & \\ & & & M(\psi_n) & \\ & & & & \boldsymbol{I}_{N-2n} \end{pmatrix},$$

where $\boldsymbol{I}_{N-2n}$ is the identity matrix of dimension $N-2n$, and the matrix $M(\psi)$ is given by

$$M(\psi) = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}.$$

In particular, R is the composition of n two-dimensional rotations. Therefore, it suffices to prove that any infinite-volume Gibbs measure μ is invariant under such a rotation. This can be achieved almost exactly as was done in the case $N = 2$, as we briefly explain now.

In view of the above, we can assume without loss of generality that R has the following block diagonal matrix representation

$$\begin{pmatrix} M(\psi) & 0 \\ 0 & \boldsymbol{I}_{N-2} \end{pmatrix},$$

for some $-\pi < \psi \leq \pi$. Let r be the global rotation associated to R. Since r only affects non-trivially the first two components S_i^1 and S_i^2 of the spins $\mathbf{S}_i$, we introduce the random variables r_i and ϑ_i, $i \in \mathbb{Z}^d$, such that

$$S_i^1 = r_i \cos\vartheta_i, \qquad S_i^2 = r_i \sin\vartheta_i .$$

(Notice that $r_i > 0$ almost surely, so that ϑ_i is almost surely well defined.)

As in the case $N = 2$, we consider an application $\Psi : \mathbb{Z}^d \to (-\pi,\pi]$ such that $\Psi_i = \psi$ for all $i \in \mathsf{B}(\ell)$, and $\Psi_i = 0$ for all $i \notin \mathsf{B}(n)$, and let $\mathsf{t}_\Psi : \Omega \to \Omega$ be the transformation such that $\vartheta_i(\mathsf{t}_\Psi\omega) = \vartheta_i(\omega) + \Psi_i$ for all configurations $\omega \in \Omega$.

From this point on, the proof is identical to the one given in Section 9.2.2. The only thing to check is that the relative entropy estimate still works in the same way. But $W(\mathbf{S}_i \cdot \mathbf{S}_j)$ is actually a function of $\vartheta_j - \vartheta_i$, r_i, r_j, and the components S_i^l, S_j^l with $l \geq 3$. Since all these quantities except the first one remain constant under the action of t_Ψ, and since the first one becomes $\vartheta_j - \vartheta_i + \Psi_j - \Psi_i$, the conclusion follows exactly as before. □

9.3 Digression on gradient models

Before turning to the study of correlations in $\mathsf{O}(N)$-symmetric models, we take advantage of the technique developed in the previous section, to take a new look at the gradient models of Chapter 8.

We proved in Theorem 8.19 that the massless GFF possesses no infinite-volume Gibbs measures in dimensions 1 and 2, a consequence of the divergence of the variance of the field in the thermodynamic limit. Our aim here is to explain how this divergence can actually be seen as resulting from the presence of a continuous symmetry at the level of the Hamiltonian. As a by-product, this will allow us to extend the proof of non-existence of infinite-volume Gibbs measures in dimensions 1 and 2 to a rather large class of models. So, only for this section, we switch to the models and notations of Chapter 8. In particular, spins take their values in $\mathbb{R}$, and Ω now represents $\mathbb{R}^{\mathbb{Z}^d}$.

Remember that gradient models have Hamiltonians of the form

$$\mathscr{H}_\Lambda = \sum_{\{i,j\}\in\mathscr{E}_\Lambda^{\mathrm{b}}} V(\varphi_j - \varphi_i),$$

where the inverse temperature has been included in $V : \mathbb{R} \to \mathbb{R}_{\geq 0}$. One must assume that V increases fast enough at infinity to make the finite-volume Gibbs measure well defined (that is, to make the partition function finite). The massless GFF corresponds to taking $V(x) = \frac{1}{4d}x^2$.

Let $t \in \mathbb{R}$, and consider the transformation $\mathsf{v}_t : \Omega \to \Omega$ defined by

$$(\mathsf{v}_t\omega)_i \overset{\text{def}}{=} \omega_i - t.$$

Since the interaction, by definition, depends only on the gradients $\omega_j - \omega_i$,

$$\mathscr{H}_\Lambda(\mathsf{v}_t\omega) = \mathscr{H}_\Lambda(\omega), \quad \forall t \in \mathbb{R}.$$

Of course, the setting here differs from the one we studied earlier in this chapter, in particular because the transformation group now is non-compact (it is actually isomorphic to $(\mathbb{R},+)$*). Let us assume for a moment that an analogue of the Mermin–Wagner theorem still applies in the present setting. Suppose also that* μ *is an infinite-volume Gibbs measure. We would then conclude that the distribution of* φ_0 *under* μ *should be uniform over* $\mathbb{R}$*, but then it would not be a probability distribution. This contradiction would show that such an infinite-volume Gibbs measure* μ *cannot exist!* ⋄

We will now show how the above can be turned into a rigorous argument. In fact, we will obtain (rather good) lower bounds on fluctuations for finite-volume Gibbs distributions:

Theorem 9.9. *($d = 1,2$) Consider the gradient model introduced above, with* $V : \mathbb{R} \to \mathbb{R}_{\geq 0}$ *even, twice differentiable, satisfying* $V(0) = 0$ *and* $\sup_{x\in\mathbb{R}} V''(x) < C < \infty$*. Then there exists a constant* $c > 0$ *such that, for any boundary condition* η*, the following holds: for all* $K > 0$*, when* n *is large enough,*

$$\mu^\eta_{\mathsf{B}(n)}\big(|\varphi_0| > K\big) \geq \begin{cases} \frac{1}{c}\exp\{-cCK^2/n\} & \text{if } d = 1, \\ \frac{1}{c}\exp\{-cCK^2/\log n\} & \text{if } d = 2. \end{cases}$$

Exercise 9.6. *Using Theorem 9.9, show that there exist no Gibbs measures for such gradient models in* $d = 1,2$. Hint: *Argue as in the proof of Theorem 8.19.*

Proof of Theorem 9.9: We assume that $\mu^\eta_{\mathsf{B}(n)}(\varphi_0 > 0) \geq \frac{1}{2}$ (if this fails, then consider the boundary condition $-\eta$). Let $T_n(d)$ be defined as in Exercise 9.5. To study

$\mu^\eta_{\mathsf{B}(n)}(|\varphi_0|>K)$, we change $\mu^\eta_{\mathsf{B}(n)}$ into a new measure under which the event is likely to occur. Namely, let $\Psi:\mathbb{Z}^d\to\mathbb{R}$ be such that $\Psi_0=K$, and $\Psi_i=0$ when $i\not\in\mathsf{B}(n)$. Let then $\mathsf{v}_\Psi:\Omega\to\Omega$ be defined by $(\mathsf{v}_\Psi\omega)_i\stackrel{\text{def}}{=}\omega_i-\Psi_i$. Let us consider the following deformed probability measure:

$$\mu^{\eta;\Psi}_{\mathsf{B}(n)}(A)\stackrel{\text{def}}{=}\mu^\eta_{\mathsf{B}(n)}(\mathsf{v}_\Psi A),\qquad \forall A\in\mathscr{F}.$$

Under this new measure, the event we are considering has probability

$$\mu^{\eta;\Psi}_{\mathsf{B}(n)}\big(|\varphi_0|>K\big)\geq\mu^{\eta;\Psi}_{\mathsf{B}(n)}\big(\varphi_0>K\big)=\mu^\eta_{\mathsf{B}(n)}\big(\varphi_0>0\big)\geq\tfrac{1}{2}. \tag{9.19}$$

Then, the probability of the same event under the original measure can be estimated by using the relative entropy inequality of Lemma B.68. The latter allows to compare the probability of an event A under two different (non-singular) probability measures μ,ν:

$$\mu(A)\geq\nu(A)\exp\Big(-\frac{\mathsf{h}(\nu\,|\,\mu)+e^{-1}}{\nu(A)}\Big).$$

Together with (9.19), this gives in our case:

$$\mu^\eta_{\mathsf{B}(n)}\big(|\varphi_0|>K\big)\geq\tfrac{1}{2}\exp\big\{-2\big(\mathsf{h}(\mu^{\eta;\Psi}_{\mathsf{B}(n)}\,|\,\mu^\eta_{\mathsf{B}(n)})+e^{-1}\big)\big\}.$$

To conclude, we must now choose Ψ so as to bound $\mathsf{h}(\mu^{\eta;\Psi}_{\mathsf{B}(n)}\,|\,\mu^\eta_{\mathsf{B}(n)})$ uniformly in n. Proceeding as in the proof of Theorem 9.7, we obtain

$$\mathsf{h}(\mu^{\eta;\Psi}_{\mathsf{B}(n)}\,|\,\mu^\eta_{\mathsf{B}(n)})\leq C\sum_{\{i,j\}\in\mathscr{E}^{\mathsf{b}}_{\mathsf{B}(n)}}(\nabla\Psi)^2_{ij}=2C\mathscr{E}(\Psi).$$

Lemma 9.8 thus implies that the choice of Ψ that minimizes $\mathscr{E}$ is

$$\Psi_i=K\,\mathbb{P}_i(X\text{ hits 0 before exiting }\mathsf{B}(n)).$$

Moreover, for this choice of Ψ, it follows from (9.15) that

$$\begin{aligned}\mathscr{E}(\Psi)&=\tfrac{1}{2}K^2\sum_{i\sim 0}\mathbb{P}_i(\text{X exits }\mathsf{B}(n)\text{ before hitting 0})\\&=dK^2\mathbb{P}_0(\text{X exits }\mathsf{B}(n)\text{ before returning to 0}).\end{aligned}$$

Since the latter probability is of order $T_n(d)^{-2}$ (Theorem B.74), this concludes the proof. □

9.4 Decay of correlations

We have already seen the following consequence of Theorem 9.2: for any $\mu\in\mathscr{G}(2)$, there is no orientational long-range order in dimensions 1 and 2:

$$\langle\mathbf{S}_i\cdot\mathbf{S}_j\rangle_\mu\to 0,\quad \|j-i\|_2\to\infty.$$

The estimates in the proof of Proposition 9.7 can be used to provide some information on the speed at which these correlations decay to zero. Namely, using (9.5) and Exercise 9.5 with $n=\|j-i\|_\infty-1$, one obtains the upper bound

$$\big|\langle\mathbf{S}_i\cdot\mathbf{S}_j\rangle_\mu\big|\leq\begin{cases}\frac{C}{\sqrt{\|j-i\|_\infty}} & \text{in } d=1,\\ \frac{C}{\sqrt{\log\|j-i\|_\infty}} & \text{in } d=2.\end{cases}$$

Unfortunately, these bounds are far from being optimal. In the next sections, we discuss various improvements.

9.4.1 One-dimensional models

For one-dimensional models, it can actually be proved that the 2-point function decays exponentially fast in $\|j-i\|_\infty$ for all $\beta<\infty$. In this section, we will prove this result for $\mathsf{O}(N)$ models.

There are several ways of obtaining this result; we will proceed by comparison with the Ising model, for which this issue has already been considered in Chapter 3. The main result we will use is the following simple inequality between 2-point functions in the $\mathsf{O}(N)$ and the Ising models.

Theorem 9.10. *For any $d\ge 1$, any $N\ge 1$, any $\beta\ge 0$, any Gibbs measure μ of the $\mathsf{O}(N)$ model at inverse temperature β on $\mathbb{Z}^d$,*

$$|\langle \mathbf{S}_0\cdot\mathbf{S}_i\rangle_\mu| \le N\,\langle\sigma_0\sigma_i\rangle^{+,\mathrm{Ising}}_{\beta,0},$$

where the expectation in the right-hand side is with respect to the Gibbs measure $\mu^+_{\beta,0}$ of the Ising model on $\mathbb{Z}^d$ at inverse temperature β and $h=0$.

Proof. Let n be such that $\{0,i\}\subset \mathsf{B}(n)$. By the DLR compatibility conditions,

$$\langle \mathbf{S}_0\cdot\mathbf{S}_i\rangle_\mu = \big\langle \langle \mathbf{S}_0\cdot\mathbf{S}_i\rangle^{\cdot}_{\mathsf{B}(n)}\big\rangle_\mu = \sum_{\ell=1}^N \big\langle \langle S_0^\ell S_i^\ell\rangle^{\cdot}_{\mathsf{B}(n)}\big\rangle_\mu .$$

It is thus sufficient to prove that

$$|\langle S_0^1 S_i^1\rangle^\eta_{\mathsf{B}(n)}| \le \langle\sigma_0\sigma_i\rangle^{+,\mathrm{Ising}}_{\beta,0},$$

for any boundary condition η.

Let $\sigma_j \stackrel{\text{def}}{=} S_j^1/|S_j^1| \in\{\pm1\}$ (of course, $S_j^1\neq 0$, for all $j\in\mathsf{B}(n)$, almost surely). Since

$$S_j^1 = |S_j^1|\sigma_j = \Big\{1-\sum_{\ell=2}^N (S_j^\ell)^2\Big\}^{1/2}\sigma_j ,$$

conditionally on the values of $S_j^2,\dots,S_j^N$, all the randomness in S_j^1 is contained in the sign σ_j. Introducing the σ-algebra $\mathscr{F}^{\neq1}_{\mathsf{B}(n)} \stackrel{\text{def}}{=} \sigma\{S_j^\ell : j\in\mathsf{B}(n), \ell\neq 1\}$, we can thus write

$$\langle S_0^1 S_i^1\rangle^\eta_{\mathsf{B}(n)} = \big\langle \langle S_0^1 S_i^1 \,|\, \mathscr{F}^{\neq1}_{\mathsf{B}(n)}\rangle^\eta_{\mathsf{B}(n)}\big\rangle^\eta_{\mathsf{B}(n)} = \big\langle |S_0^1||S_i^1| \langle \sigma_0\sigma_i \,|\, \mathscr{F}^{\neq1}_{\mathsf{B}(n)}\rangle^\eta_{\mathsf{B}(n)}\big\rangle^\eta_{\mathsf{B}(n)} .$$

Observe now that the joint distribution of the random variables $(\sigma_j)_{j\in\mathsf{B}(n)}$ is given by an inhomogeneous Ising model in $\mathsf{B}(n)$, with Hamiltonian

$$\mathscr{H}_{\mathsf{B}(n);\mathbf{J}} \stackrel{\text{def}}{=} -\sum_{\{u,v\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}} J_{uv}\sigma_u\sigma_v ,$$

where the coupling constants are given by $J_{uv}\stackrel{\text{def}}{=}\beta|S_u^1||S_v^1|$ and the boundary condition by $\hat\eta = (\eta_j^1/|\eta_j^1|)_{j\in\mathbb{Z}^d}$. Since $0\le J_{uv}\le\beta$, it follows from Exercise 3.31 that, almost surely,

$$\big\langle \sigma_0\sigma_i \,\big|\, \mathscr{F}^{\neq1}_{\mathsf{B}(n)}\big\rangle^\eta_{\mathsf{B}(n)} = \langle\sigma_0\sigma_i\rangle^{\hat\eta,\mathrm{Ising}}_{\mathsf{B}(n);\mathbf{J}} \le \langle\sigma_0\sigma_i\rangle^{+,\mathrm{Ising}}_{\mathsf{B}(n);\beta,0} .$$

For the lower bound, set $\tilde J_{0j} = -J_{0j}$ for all $j\sim 0$ and $\tilde J_{uv}=J_{uv}$ for all other pairs and use $\langle\sigma_0\sigma_i\rangle^{\hat\eta,\mathrm{Ising}}_{\mathsf{B}(n);\mathbf{J}} = -\langle\sigma_0\sigma_i\rangle^{\hat\eta,\mathrm{Ising}}_{\mathsf{B}(n);\tilde{\mathbf{J}}},\ge -\langle\sigma_0\sigma_i\rangle^{+,\mathrm{Ising}}_{\mathsf{B}(n);\beta,0}$ as before. □

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

Applying this lemma in dimension 1, we immediately deduce the desired estimate from Exercise 3.25.

Corollary 9.11. *Let μ be the unique Gibbs measure of the $\mathsf{O}(N)$ model on $\mathbb{Z}$. Then, for any $0 \le \beta < \infty$,*

$$|\langle \mathbf{S}_0 \cdot \mathbf{S}_i \rangle_\mu| \le N (\tanh\beta)^{|i|}.$$

Alternatively, one can compute explicitly the 2-point function, by integrating one spin at a time.

Exercise 9.7. *Consider the one-dimensional XY model at inverse temperature β. Let μ be its unique Gibbs measure. Compute the pressure and the correlation function $\langle \mathbf{S}_0 \cdot \mathbf{S}_i \rangle_\mu$ in terms of the* ***modified Bessel functions of the first kind****:*

$$I_n(x) \stackrel{\text{def}}{=} \frac{1}{\pi}\int_0^\pi e^{x\cos t}\cos(nt)\mathrm{d}t.$$

Hint: *Use free boundary conditions.*

9.4.2 Two-dimensional models

We investigate now whether it is also possible to improve the estimate in dimension 2. To keep the matter as simple as possible, we only consider the XY model, although similar arguments apply for a much larger class of two-dimensional models, as described in Section 9.6.2.

Heuristic argument

Let us start with some heuristic considerations, which lead to a conjecture on the rate at which $\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_\mu$ should decrease to 0 at low temperature.

As before, we write the spin at i as $\mathbf{S}_i = (\cos\vartheta_i, \sin\vartheta_i)$. We are interested in the asymptotic behavior of

$$\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle^{\mathbf{e}_1}_{\Lambda;\beta} = \langle \cos(\vartheta_j - \vartheta_i) \rangle^{\mathbf{e}_1}_{\Lambda;\beta} = \langle e^{\mathrm{i}(\vartheta_j - \vartheta_i)} \rangle^{\mathbf{e}_1}_{\Lambda;\beta},$$

where the expectation is taken with boundary condition $\eta \equiv \mathbf{e}_1$; the last identity relies on the symmetry, which makes the imaginary part vanish.

At very low temperatures, most neighboring spins are typically nearly aligned, $|\vartheta_i - \vartheta_j| \ll 1$. In this regime, it makes sense to approximate the interaction term in the Hamiltonian using a Taylor expansion to second order:

$$-\beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_\Lambda} \mathbf{S}_i \cdot \mathbf{S}_j = -\beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_\Lambda} \cos(\vartheta_j - \vartheta_i) \cong -\beta|\mathscr{E}^{\mathrm{b}}_\Lambda| + \tfrac{1}{2}\beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_\Lambda} (\vartheta_j - \vartheta_i)^2.$$

We may also assume that, when β is very large, the behavior of the field is not much affected by replacing the angles ϑ_i, which take their values in $(-\pi,\pi]$, by variables φ_i taking values in $\mathbb{R}$, especially since we are interested in the expectation value of the 2π-periodic function $e^{\mathrm{i}(\varphi_j - \varphi_i)}$. This discussion leads us to conclude that the very-low temperature properties of the XY model should be closely approximated by those of the GFF at inverse temperature 4β.

In particular, if we temporarily denote the expectations of the XY model with boundary condition $\eta_i \equiv \mathbf{e}_1$ by $\langle\cdot\rangle^{XY}_{\Lambda;\beta}$, and the expectation of the corresponding GFF with boundary condition $\eta_i \equiv 0$ by $\langle\cdot\rangle^{GFF}_{\Lambda,4\beta}$, we conclude that

$$\langle \mathbf{S}_i\cdot\mathbf{S}_j\rangle^{XY}_{\Lambda;\beta} = \langle e^{\mathrm{i}(\vartheta_j-\vartheta_i)}\rangle^{XY}_{\Lambda;\beta} \cong \langle e^{\mathrm{i}(\varphi_j-\varphi_i)}\rangle^{GFF}_{\Lambda;4\beta}.$$

Now, since $(\varphi_i)_{i\in\Lambda}$ is Gaussian, (8.8) gives

$$\langle e^{\mathrm{i}(\varphi_j-\varphi_i)}\rangle^{GFF}_{\Lambda;4\beta} = e^{-\frac{1}{8\beta}(G_\Lambda(i,i)+G_\Lambda(j,j)-2G_\Lambda(i,j))}, \tag{9.20}$$

where $G_\Lambda(i,j)$ is the Green function of the simple random walk in Λ (see Section 8.4.1). We will see at the end of the section that, as $\|j-i\|_2\to\infty$,

$$\tfrac{1}{2}\lim_{\Lambda\uparrow\mathbb{Z}^2}(G_\Lambda(i,i)+G_\Lambda(j,j)-2G_\Lambda(i,j)) \simeq \frac{2}{\pi}\log\|j-i\|_2\,,$$

which leads to the following conjectural behavior for correlations at low temperatures:

$$\langle \mathbf{S}_i\cdot\mathbf{S}_j\rangle^{XY}_{\Lambda;\beta} \cong e^{-\frac{1}{2\pi\beta}\log\|j-i\|_2} = \|j-i\|_2^{-1/(2\pi\beta)}. \tag{9.21}$$

Algebraic decay at low temperature

The following theorem provides, for large β, an essentially optimal upper bound of the type (9.21). The lower bound will be discussed (but not proved) in Section 9.6.1.

Theorem 9.12. *Let μ be an infinite-volume Gibbs measure associated to the two-dimensional XY model at inverse temperature β. For all $\epsilon>0$, there exists $\beta_0(\epsilon)<\infty$ such that, for all $\beta>\beta_0(\epsilon)$ and all $i\neq j\in\mathbb{Z}^2$,*

$$|\langle \mathbf{S}_i\cdot\mathbf{S}_j\rangle_\mu| \le \|j-i\|_2^{-(1-\epsilon)/(2\pi\beta)}.$$

Before turning to the proof, let us try to motivate the approach that will be used, which might otherwise seem rather uncanny. To do this, let us return to (9.20). To actually compute the expectation $\langle e^{\mathrm{i}(\varphi_j-\varphi_i)}\rangle^{GFF}_{\Lambda;4\beta}$, one should remember that $\varphi_j-\varphi_i$ has a normal distribution $\mathscr{N}(0,\sigma^2)$, with (see (8.6))

$$\sigma^2 = \frac{1}{4\beta}(G_\Lambda(i,i)+G_\Lambda(j,j)-2G_\Lambda(i,j)).$$

Its characteristic function can be computed by first completing the square:

$$\langle e^{\mathrm{i}(\varphi_j-\varphi_i)}\rangle^{GFF}_{\Lambda;4\beta} = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{\mathrm{i}x-\frac{1}{2\sigma^2}x^2}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{\sigma^2}{2}}\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(x-\mathrm{i}\sigma^2)^2}\,\mathrm{d}x.$$

Once the leading term $e^{-\frac{\sigma^2}{2}}$ is extracted, the remaining integral can be computed by translating the path of integration from $\mathbb{R}$ to $\mathbb{R}+\mathrm{i}\sigma^2$, an operation vindicated, through Cauchy's integral theorem, by the analyticity and rapid decay at infinity of the integrand:

$$\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(x-\mathrm{i}\sigma^2)^2}\,\mathrm{d}x = \int_{\mathbb{R}+\mathrm{i}\sigma^2} e^{-\frac{1}{2\sigma^2}(z-\mathrm{i}\sigma^2)^2}\,\mathrm{d}z = \int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}x^2}\,\mathrm{d}x = \sqrt{2\pi\sigma^2}.$$

The proof below follows a similar scheme, but applied directly to the XY spins instead of the GFF approximation.

Proof of Theorem 9.12: Without loss of generality, we consider $i = 0$, $j = k$. Similarly to what was done in (9.5), we first rely on the DLR property: for all n such that $\mathsf{B}(n) \ni k$,

$$\big\langle \mathbf{S}_0 \cdot \mathbf{S}_k \big\rangle_\mu = \int \big\langle \mathbf{S}_0 \cdot \mathbf{S}_k \big\rangle^\eta_{\mathsf{B}(n);\beta}\, \mu(\mathrm{d}\eta)\,. \tag{9.22}$$

We will estimate the expectation in the right-hand side, uniformly in the boundary condition η. Observe first that

$$\big|\langle \mathbf{S}_0 \cdot \mathbf{S}_k\rangle^\eta_{\mathsf{B}(n);\beta}\big| = \big|\langle \cos(\vartheta_k - \vartheta_0)\rangle^\eta_{\mathsf{B}(n);\beta}\big| \le \big|\langle e^{\mathrm{i}(\vartheta_k-\vartheta_0)}\rangle^\eta_{\mathsf{B}(n);\beta}\big|,$$

since $|\mathfrak{Re}\, z| \le |z|$ for all $z \in \mathbb{C}$. We will write the expectation $\langle\cdot\rangle^\eta_{\mathsf{B}(n);\beta}$ using explicit integrals over the angle variables $\vartheta_i \in (-\pi,\pi]$, $i \in \mathsf{B}(n)$. As a shorthand, we use the notation

$$\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \prod_{i\in\mathsf{B}(n)} \mathrm{d}\theta_i \equiv \int \mathrm{d}\theta_{\mathsf{B}(n)}\,.$$

Therefore,

$$\big\langle e^{\mathrm{i}(\vartheta_k-\vartheta_0)}\big\rangle^\eta_{\mathsf{B}(n);\beta} = \frac{1}{\mathbf{Z}^\eta_{\mathsf{B}(n);\beta}} \int \mathrm{d}\theta_{\mathsf{B}(n)} \exp\Big\{\mathrm{i}(\theta_k - \theta_0) + \beta \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}} \cos(\theta_i - \theta_j)\Big\},$$

where we have set $\theta_i = \vartheta_i(\eta)$ for each $i \not\in \mathsf{B}(n)$. Following the approach sketched before the proof, we add an imaginary part to the variables θ_j, $j \in \mathsf{B}(n)$. Since the integrand is clearly analytic, we can easily deform the integration path associated to the variable θ_j away from the real axis: we shift the integration interval from $[-\pi,\pi]$ to $[-\pi,\pi] + \mathrm{i}r_j$, where r_j will be chosen later (also as a function of β and n):

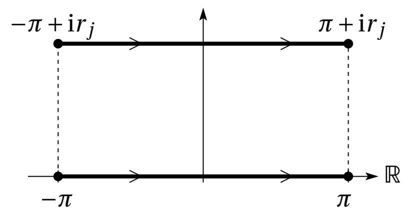


Figure 9.5: Shifting the integration path of θ_j. The shift depends on the vertex j, on n and on β.

Notice that the periodicity of the integrand guarantees that the contributions coming from the two segments connecting these two intervals cancel each other. We extend the r_is to a function $r : \mathbb{Z}^2 \to \mathbb{R}$, with $r_i = 0$ for all $i \not\in \mathsf{B}(n)$. Observe now that

$$|e^{\mathrm{i}(\theta_k + \mathrm{i}r_k - \theta_0 - \mathrm{i}r_0)}| = e^{-(r_k - r_0)},$$
$$|e^{\cos(\theta_i + \mathrm{i}r_i - \theta_j - \mathrm{i}r_j)}| = e^{\cosh(r_i - r_j)\cos(\theta_i - \theta_j)}.$$

We thus have, letting ω_θ denote the spin configuration associated to the angles θ_i,

$$\begin{aligned}
\big|\langle \mathbf{S}_0\cdot\mathbf{S}_k\rangle^{\eta}_{\mathsf{B}(n);\beta}\big| &\le \frac{e^{-(r_k-r_0)}}{\mathbf{Z}^{\eta}_{\mathsf{B}(n);\beta}}\int \mathrm{d}\theta_{\mathsf{B}(n)}\exp\Big\{\beta\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}}\cosh(r_i-r_j)\cos(\theta_i-\theta_j)\Big\}\\
&= e^{-(r_k-r_0)}\int \mathrm{d}\theta_{\mathsf{B}(n)}\exp\Big\{\beta\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}}\big(\cosh(r_i-r_j)-1\big)\cos(\theta_i-\theta_j)\Big\}\frac{e^{-\mathscr{H}_{\mathsf{B}(n);\beta}(\omega_\theta)}}{\mathbf{Z}^{\eta}_{\mathsf{B}(n);\beta}}\\
&= e^{-(r_k-r_0)}\Big\langle \exp\Big\{\beta\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}}\big(\cosh(r_i-r_j)-1\big)\cos(\vartheta_i-\vartheta_j)\Big\}\Big\rangle^{\eta}_{\mathsf{B}(n);\beta}\\
&\le e^{-(r_k-r_0)}\exp\Big\{\beta\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}}\big(\cosh(r_i-r_j)-1\big)\Big\}. \qquad (9.23)
\end{aligned}$$

In the last inequality, we used the fact that $\cosh(r_i-r_j)\ge 1$ and $\cos(\vartheta_i-\vartheta_j)\le 1$. Assume that r can be chosen in such a way that

$$|r_i-r_j|\le C/\beta,\quad \forall\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}, \qquad (9.24)$$

for some constant C. This allows us to replace the cosh term by a simpler quadratic term: given $\epsilon>0$, we can assume that β_0 is large enough to ensure that $\beta\ge\beta_0$ implies $\cosh(r_i-r_j)-1\le\frac12(1+\epsilon)(r_i-r_j)^2$ for all $\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}$. In particular, we can write

$$\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}}\big(\cosh(r_i-r_j)-1\big)\le\tfrac12(1+\epsilon)\sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}}(r_i-r_j)^2=(1+\epsilon)\mathscr{E}(r), \qquad (9.25)$$

where $\mathscr{E}(\cdot)$ is the Dirichlet energy functional defined on maps $r:\mathbb{Z}^2\to\mathbb{R}$ that vanish outside $\mathsf{B}(n)$. We thus have

$$\big|\langle \mathbf{S}_0\cdot\mathbf{S}_k\rangle^{\eta}_{\mathsf{B}(n);\beta}\big|\le\exp\{-\mathscr{D}(r)\}, \qquad (9.26)$$

with $\mathscr{D}(\cdot)$ the functional defined by

$$\mathscr{D}(r)\stackrel{\mathrm{def}}{=} r_k-r_0-\beta'\mathscr{E}(r),$$

where we have set $\beta'\stackrel{\mathrm{def}}{=}(1+\epsilon)\beta$. We now search for a maximizer of $\mathscr{D}$.

Lemma 9.13. *For a fixed $0\ne k\in\mathsf{B}(n)$, the functional $\mathscr{D}$ possesses a unique maximizer r^* among the functions r that satisfy $r_i=0$ for all $i\notin\mathsf{B}(n)$. That maximizer is the unique such function that satisfies*

$$(\Delta r)_i=(1_{\{i=0\}}-1_{\{i=k\}})/\beta',\quad i\in\mathsf{B}(n). \qquad (9.27)$$

It can be expressed explicitly as (see Figure 9.6)

$$r^*_i=\big(G_{\mathsf{B}(n)}(i,k)-G_{\mathsf{B}(n)}(i,0)\big)/(4\beta'),\quad i\in\mathsf{B}(n), \qquad (9.28)$$

where $G_{\mathsf{B}(n)}(\cdot,\cdot)$ is the Green function of the symmetric simple random walk in $\mathsf{B}(n)$.

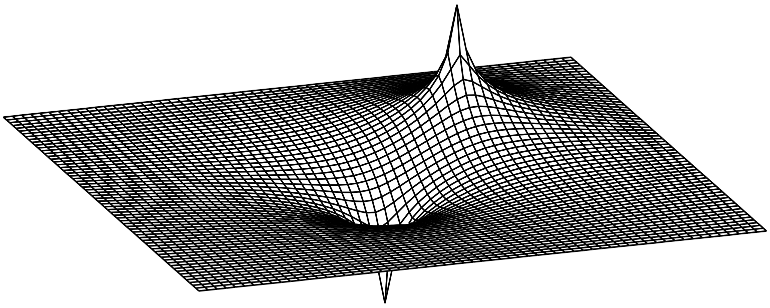


Figure 9.6: The maximizer (9.28) of the functional $\mathscr{D}$. In this picture, $\Lambda = \mathsf{B}(30)$ and $k = (12, 12)$.

Proof. As in the proof of Lemma 9.8, we start by observing that a critical point r of $\mathscr{D}$ must be such that $\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{D}(r+s\delta)\big|_{s=0} = 0$ for all perturbations $\delta : \mathbb{Z}^d \to \mathbb{R}$ which vanish outside $\mathsf{B}(n)$. But, as a straightforward computation shows,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathscr{D}(r+s\delta)\Big|_{s=0} &= \delta_k - \delta_0 + \beta' \sum_{i\in\mathsf{B}(n)} \delta_i (\Delta r)_i \\ &= \delta_k\big(1+\beta'(\Delta r)_k\big) - \delta_0\big(1-\beta'(\Delta r)_0\big) + \beta' \sum_{i\in\mathsf{B}(n)\setminus\{0,k\}} \delta_i (\Delta r)_i \,.\end{aligned}$$

Since this sum of three terms must vanish for all δ, we see that r must satisfy (9.27). Since $r_i = 0$ outside $\mathsf{B}(n)$, we have $(\Delta r)_i = (\Delta_{\mathsf{B}(n)} r)_i$ (remember Remark 8.9), so that (9.27) can be written

$$(\Delta_{\mathsf{B}(n)} r)_i = (1_{\{i=0\}} - 1_{\{i=k\}})/\beta' \,.$$

In Lemma 8.13, we saw that $G_{\mathsf{B}(n)}$ is precisely the inverse of $-\frac{1}{4}\Delta_{\mathsf{B}(n)}$. Therefore, multiplying by $G_{\mathsf{B}(n)}(j,i)$ on both sides of the previous display and summing over i gives (9.28). To prove that r^* actually maximizes $\mathscr{D}(\cdot)$, let δ be such that $\delta_i = 0$ outside $\mathsf{B}(n)$. Proceeding as we have already done several times before,

$$\begin{aligned}\mathscr{D}(r^*+\delta) &= \mathscr{D}(r^*) - \beta'\mathscr{E}(\delta) + \delta_k - \delta_0 - \beta' \sum_{\{i,j\}\in\mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}} (\nabla\delta)_{ij}(\nabla r^*)_{ij} \\ &= \mathscr{D}(r^*) - \beta'\mathscr{E}(\delta) + \delta_k - \delta_0 + \beta' \sum_{i\in\mathsf{B}(n)} \delta_i \underbrace{(\Delta r^*)_i}_{\text{use (9.27)}} \\ &= \mathscr{D}(r^*) - \beta'\mathscr{E}(\delta)\,.\end{aligned}$$

Since $\mathscr{E}(\delta) \geq 0$, we conclude that $\mathscr{D}(r^*+\delta) \leq \mathscr{D}(r^*)$. □

It follows from Theorem B.77 that there exists a constant C such that $|G_{\mathsf{B}(n)}(i,v) - G_{\mathsf{B}(n)}(j,v)| \leq 2C$, uniformly in n, in $v \in \mathsf{B}(n)$ and in $\{i,j\} \in \mathscr{E}^{\mathrm{b}}_{\mathsf{B}(n)}$. In particular, $|r^*_i - r^*_j| \leq C/\beta$, meaning that (9.24) is satisfied. We now use (9.26) with r^*. First, one easily verifies that

$$\mathscr{D}(r^*) = \tfrac{1}{2}(r^*_k - r^*_0) = \tfrac{1}{8\beta'}\big\{(G_{\mathsf{B}(n)}(k,k) - G_{\mathsf{B}(n)}(k,0)) + (G_{\mathsf{B}(n)}(0,0) - G_{\mathsf{B}(n)}(0,k))\big\}\,.$$

We then let $n \to \infty$; using Exercise B.24,

$$\lim_{n\to\infty}(G_{\mathsf{B}(n)}(k,k) - G_{\mathsf{B}(n)}(k,0)) = \lim_{n\to\infty}(G_{\mathsf{B}(n)}(0,0) - G_{\mathsf{B}(n)}(0,k)) = a(k)\,, \qquad (9.29)$$

where $a(k)$ is called the **potential kernel** of the symmetric simple random walk on $\mathbb{Z}^2$, defined by

$$a(k) \stackrel{\text{def}}{=} \sum_{m\geq 0} \{\mathbb{P}_0(X_m = 0) - \mathbb{P}_k(X_m = 0)\}.$$

Therefore,

$$|\langle \mathbf{S}_0 \cdot \mathbf{S}_k \rangle_\mu| \leq e^{-a(k)/4\beta'}.$$

The conclusion now follows since, by Theorem B.76,

$$a(k) = \frac{2}{\pi} \log \|k\|_2 + O(1) \quad \text{as } \|k\|_2 \to \infty. \qquad \square$$

9.5 Bibliographical references

The problems treated in this chapter find their origin in a celebrated work by Mermin and Wagner [242, 241]. The latter triggered a long series of subsequent investigations, leading to stronger claims under weaker assumptions: [83, 237, 317, 179, 273, 193, 32, 243, 240, 252, 169, 126] to mention just a few.

Mermin–Wagner theorem. The proof of Theorem 9.2 follows the approach of Pfister [273], with some improvements from [169]. Being really classical material, alternative presentations of this material can be found in many books, such as [312, 282, 134, 308].

Effective interface models in $d = 1$ and 2. That the same type of arguments can be used to prove the absence of any Gibbs state in models with unbounded spins, as in our Theorem 9.9, was first realized by Dobrushin and Shlosman [84] and by Fröhlich and Pfister [119], although they did not derive quantitative lower bounds. Generalizations of Theorem 9.9 to more general potentials can be found in [169, 246]. Let us also mention that results of this type can also be derived by a very different method. Namely, relying on an inequality derived by Brascamp and Lieb in [42], it is possible to compare, under suitable assumptions, the variance of an effective interface models with that of the GFF; this alternative approach is described in [41].

Comparison with the Ising model. Our proof of Theorem 9.10 is original, as far as we know. However, a similar claim can be found in [250], with a proof based on correlation inequalities. Arguments similar to those used in the proof of Theorem 9.10 have already been used, for example, in [69].

Algebraic decay of correlation in two dimensions. The proof of Theorem 9.12 is originally due to McBryan and Spencer [237]; the argument presented in the chapter is directly based on their work. Again, alternative presentations can be found in many places, such as [140, 308].

9.6 Complements and further reading

9.6.1 The Berezinskiĭ–Kosterlitz–Thouless phase transition

Theorem 9.12 provides an algebraically decaying upper bound on the 2-point function of the two-dimensional XY model at low temperatures, which improves substantially on the bound that can be extracted from Proposition 9.7. Nevertheless, one might wonder whether this bound could be further improved, an issue which we briefly discuss now.

Consider again the two-dimensional XY model. Theorem 9.2 shows that all Gibbs measures are $\mathsf{SO}(2)$-invariant, but, as already mentioned, it does not imply uniqueness [1]. It is however possible to prove, using suitable correlation inequalities, that absence of spontaneous magnetization for the XY model entails the existence of a unique *translation-invariant* infinite-volume Gibbs measure [47]. Consequently, the Mermin–Wagner theorem implies that there is a unique translation-invariant infinite-volume Gibbs measure, at all $\beta \ge 0$, for the XY model on $\mathbb{Z}^2$. Moreover, this Gibbs measure is extremal. It is in fact expected that uniqueness holds for this model, but this has not yet been proved.

The following remarkable result proves that a phase transition of a more subtle kind nevertheless occurs in this model.

Theorem 9.14. *Consider the unique translation-invariant Gibbs measure of the two-dimensional XY model. There exist* $0 < \beta_1 < \beta_2 < \infty$ *such that*

- *for all* $\beta < \beta_1$*, there exist* $C(\beta)$ *and* $m(\beta) > 0$ *such that*

$$|\langle \mathbf{S}_0 \cdot \mathbf{S}_k \rangle_\beta| \le C(\beta)\, \exp(-m(\beta)\|k\|_2),$$

  for all $k \in \mathbb{Z}^2$*;*

- *for all* $\beta > \beta_2$*, there exist* $c(\beta) > 0$ *and* $D > 0$ *such that*

$$\langle \mathbf{S}_0 \cdot \mathbf{S}_k \rangle_\beta \ge c(\beta)\, \|k\|_2^{-D/\beta},$$

  for all $k \in \mathbb{Z}^2$.

Proof. The first claim follows immediately from Theorem 9.10 and Exercise 3.24. The proof of the second part, which is due to Fröhlich and Spencer [122], is however quite involved and goes beyond the scope of this book. □

Note that, combined with the upper bound of Theorem 9.12, this shows that the 2-point function of the two-dimensional XY model really decays algebraically at low temperature.

It is expected that there is a sharp transition between the two regimes (exponential vs. algebraic decay) described in Theorem 9.14, at a value β_{BKT} of the inverse temperature. This so-called **Berezinskiĭ–Kosterlitz–Thouless phase transition**, named after the physicists who studied this problem in the early 1970s [19, 195], exhibits several remarkable properties, among which the fact that the pressure remains infinitely differentiable (but not analytic) at the transition. One should point out, however, that the analytic properties at this phase transition are not universal, and other $\mathsf{O}(2)$-symmetric models display very different behavior, such as a first-order phase transition [344].

To conclude this discussion, let us mention an outstanding open problem in this area. As explained in Section 9.6.2, the proof of Theorem 9.12 can be adapted to obtain similar upper bounds for a general class of $\mathsf{O}(N)$-symmetric models, and in particular for all $\mathsf{O}(N)$ models with $N > 2$. However, it is conjectured that this upper bound is very poor when $N > 2$. Namely, it is expected that the 2-point function then *decays exponentially at all temperatures*. Interestingly, it is the fact that

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

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

SO(2) is abelian while the groups SO(N), $N \ge 3$, are not, which is deemed to be responsible for the difference of behavior [307].

9.6.2 Generalizations.

For pedagogical reasons, we have restricted our discussion to the simplest setup. The results presented here can however be extended in various directions. We briefly describe one possible such framework and provide some relevant references.

We assume that the spins $\mathbf{S}_i$ take values in some topological space $\mathscr{S}$, on which a compact, connected Lie group G acts continuously (we simply denote the action of $\mathrm{g} \in \mathsf{G}$ on $x \in \mathscr{S}$ by $\mathrm{g}x$). We replace the Hamiltonian in (9.1) by

$$\sum_{\substack{\{i,j\}\subset\mathbb{Z}^d:\\ \{i,j\}\cap\Lambda\neq\varnothing}} J_{j-i}\tilde{W}(S_i,S_j)\,,$$

where $(J_i)_{i\in\mathbb{Z}^d}$ is a collection of real numbers such that $\sum_{i\in\mathbb{Z}^d}|J_i| = 1$ and $\tilde{W}:\mathscr{S}\times\mathscr{S}\to\mathbb{R}$ is continuous and G-invariant, in the sense that $\tilde{W}(\mathrm{g}x,\mathrm{g}y) = \tilde{W}(x,y)$ for all $x,y\in\mathscr{S}$ and all $\mathrm{g}\in\mathsf{G}$.

Theorem 9.2 (and the more quantitative Proposition 9.7) can then be extended to this more general setup, under the assumption that the random walk on $\mathbb{Z}^d$, which jumps from i to j with probability $|J_{j-i}|$, is recurrent. This result was proved by Ioffe, Shlosman and Velenik [169], building on earlier works by Dobrushin and Shlosman [83] and Pfister [273]. We emphasize that the recurrence assumption cannot be improved in general, as there are examples of models for which spontaneous symmetry breaking at low temperatures occurs as soon as the corresponding random walk is transient [201, 32], see also [134, Theorem 20.15].

Using a similar approach and building on the earlier works of McBryan and Spencer [237] and of Messager, Miracle-Solé and Ruiz [243], Theorem 9.12 has been extended by Gagnebin and Velenik [126] to O(N)-symmetric models with a Hamiltonian as above, provided that $|J_i| \le J\|i\|_1^{-\alpha}$ for some $J<\infty$ and $\alpha > 4$.

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

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