

# **Reflection Positivity and Phase Transitions in Lattice Spin Models**

Lecture notes from

**the 5th Prague Summer School on  
Mathematical Statistical Mechanics**

September 11-22, 2006

Marek Biskup

University of California  
at Los Angeles

© 2006 by M. Biskup. Reproduction, by any means, of the entire article for non-commercial purposes is permitted without charge.



# Contents

<b>Preface</b>	<b>5</b>
<b>1 Lattice spin models</b>	<b>7</b>
1.1 Basic setup . . . . .	7
1.2 Examples . . . . .	8
1.3 Gibbs formalism . . . . .	11
1.4 Torus measures . . . . .	13
1.5 Some thermodynamics . . . . .	13
1.6 Literature remarks . . . . .	15
<b>2 IRB &amp; spin-wave condensation</b>	<b>17</b>
2.1 Random walk connections . . . . .	17
2.2 Infrared bound . . . . .	19
2.3 Spin-wave condensation in $O(n)$ -model . . . . .	21
2.4 Mermin-Wagner theorem . . . . .	23
2.5 Literature remarks . . . . .	24
<b>3 IRB &amp; Mean-field theory</b>	<b>27</b>
3.1 Mean-field theory . . . . .	27
3.2 Approximation theorem & applications . . . . .	30
3.3 Ideas from the proofs . . . . .	32
3.4 Literature remarks . . . . .	35
<b>4 Reflection positivity</b>	<b>37</b>
4.1 Reflection positive measures . . . . .	37
4.2 Gaussian domination . . . . .	40
4.3 Chessboard estimates . . . . .	42
4.4 Literature remarks . . . . .	45

<b>5 Applications of chessboard estimates</b>	<b>47</b>
5.1 Gaussian double-well model . . . . .	47
5.2 Proof of phase coexistence . . . . .	49
5.3 Gradient fields with non-convex potential . . . . .	52
5.4 Spin-waves vs infinite ground-state degeneracy . . . . .	55
5.5 Literature remarks . . . . .	59
<b>Three open problems</b>	<b>61</b>
<b>References</b>	<b>63</b>

# Preface

Phase transitions are one of the most fascinating, and also most perplexing, phenomena in equilibrium statistical mechanics. On the physics side, many approximate techniques and approaches are known but a complete mathematical understanding is available only in the simplest of all cases. This set of notes describes one successful approach to phase transitions in lattice spin models which is based on the technique of *reflection positivity*. This approach was developed in the late 1970s in the works of Dyson, Fröhlich, Israel, Lieb, Simon and Spencer who used it to prove phase transitions in a host of physically-interesting classical and quantum models; most notably, the classical Heisenberg ferromagnet and the quantum XY model and Heisenberg antiferromagnet.

The underlying structure of the arguments based on reflection positivity is very simple. Indeed, the latter is used to establish either the so called *infrared bound*, by which one infers phase transitions using a spin-condensation argument, or to produce the so called *chessboard estimate*, which allows one to implement a Peierls-type argument regardless of whether the model possesses an internal symmetry or not. Unfortunately, reflection positivity is a rather restrictive condition and, in a certain sense, applies only to a very small class of systems. Fortunately, the models to which it does apply constitute a large portion of what is interesting for physics—and to physicists. Thus, unless one is after exclusively universal statements—i.e., those robust under rather arbitrary perturbations—the route via reflection positivity is often fairly satisfactory.

The spectacular success of reflection positivity from the late 1970s was followed by many interesting developments. For instance, in various joint collaborations, Dobrushin, Kotecký and Shlosman showed how chessboard estimates can be used to prove a phase transition in a class of systems with naturally defined ordered and disordered components; most notably, the  $q$ -state Potts model for  $q \gg 1$ . Another unexpected application came in the papers of Aizenman from early 1980s in which he combined the infrared bound with his random-current representation to conclude mean-field critical behavior in the nearest-neighbor Ising ferromagnet above 4 dimensions. Yet another example is the proof, by Chayes, Kotecký and Shlosman, that the Fisher-renormalization scheme in annealed-diluted systems may be substituted by the emergence of an intermediate phase.

Recently, reflection positivity was used to prove asymptotic results which were part of a physics folklore but whose mathematical justification was missing. For instance, in papers by Crawford, Chayes and the present author, the infrared bound was shown to imply that, once a model undergoes a field or energy driven first-order transition in mean-field theory, a similar transition will occur in the lattice model provided the spatial dimension

is sufficiently high or the interaction is sufficiently spread-out (but still reflection positive). Another result—due to Chayes, Starr and the present author—asserts that if the classical limit of a quantum spin system admits a proof of phase transition by chessboard estimates, the corresponding conclusion extends also to the quantum system provided the magnitude of the quantum spin is sufficiently large.

There have also been cases where reflection positivity brought a definite end to a controversy that physics arguments were not able to resolve. One instance concerned certain non-linear vector and liquid-crystal models where there was a debate about whether a transition can occur already in 2 dimensions. This was settled in the work of van Enter and Shlosman. Another instance involved spin systems whose (infinite) set of ground states had a much larger set of symmetries than the Hamiltonian of the model; two competing physics reasonings argued for, and against, the survival of these states at low temperatures. Here, in papers of Chayes, Nussinov, Starr and the present author, spin-wave free energy calculations were combined with chessboard estimates to construct a rigorous proof of phase coexistence of only a *finite* number of low-temperature states.

The set of notes introduced herewith records nine hours of lectures on the subject of reflection positivity and phase transitions that were delivered at the 5th Prague School of Mathematical Statistical Mechanics in September 2006. The material of the notes is essentially identical to the lectures save for the last chapter which was not entirely covered for lack of time. The limited time-frame of the lectures, and the desire of the lecturer to explain all concepts in sufficient depth, caused these notes to be far from the initially-intended comprehensive review. Nevertheless, while only a modest fraction of this text is spent on discussing recent applications, it is hoped that the readership of these notes will get sufficiently prepared for getting through—and building further upon—the recent work without serious difficulties.

It is a pleasure to thank Roman Kotecký for organizing the summer school and for allowing me to speak on this subject. Next I would like to thank Lincoln Chayes with whom I have coauthored more than half-a-dozen papers on reflection positivity and who taught me many subtleties of this field. I am also much indebted to Aernout van Enter for our frequent email conversations on these subjects, and for many helpful suggestions on the first draft of this text. Finally, my presence at the school was made possible thanks to the support from the ESF-program “Phase Transitions and Fluctuation Phenomena for Random Dynamics in Spatially Extended Systems” and from the National Science Foundation under the grant DMS-0505356.

M.B.  
Praha & Los Angeles  
September 2006

# Chapter 1

## Lattice spin models: Crash course

This chapter prepares the ground for the rest of the course by introducing the main concepts from the theory of Gibbs measures for lattice spin models. The results introduced here are selected entirely for the purpose of this note; readers wishing a more comprehensive—and in-depth—treatment should consult classic textbooks on the subject.

### 1.1 Basic setup

Let us start discussing the setup of the models to which we will direct our attention throughout this course. The basic ingredients are as follows:

- *Lattice*: We will take the  $d$ -dimensional hypercubic lattice  $\mathbb{Z}^d$  as our underlying graph. This is a graph with vertices at all points in  $\mathbb{R}^d$  with integer coordinates and edges between any *nearest neighbor* pair of vertices; i.e., those at Euclidean distance one. We will use  $\langle x, y \rangle$  to denote an (unordered) nearest-neighbor pair.
- *Spins*: At each  $x \in \mathbb{Z}^d$  we will consider a spin  $S_x$ , by which we will mean a random variable taking values in a closed subset  $\Omega$  of  $\mathbb{R}^\nu$ , for some  $\nu \geq 1$ . We will use  $S_x \cdot S_y$  to denote a scalar product between  $S_x$  and  $S_y$  (Euclidean or otherwise).
- *Spin configurations*: For  $\Lambda \subset \mathbb{Z}^d$ , we will refer to  $S_\Lambda = (S_x)_{x \in \Lambda}$  as the spin configuration in  $\Lambda$ . We will be generically interested in describing the statistical properties of such spin configurations with respect to certain (canonical) measures.
- *Boundary conditions*: To describe the law of  $S_\Lambda$ , we will not be able to ignore that some spins are also outside  $\Lambda$ . We will refer to the configuration  $S_{\Lambda^c}$  of these spins as the boundary condition. The latter will usually be fixed and may often even be considered a parameter of the game. When both  $S_\Lambda$  and  $S_{\Lambda^c}$  are known, we will write

$$S = (S_\Lambda, S_{\Lambda^c}) \tag{1.1}$$

to denote their concatenation on all of  $\mathbb{Z}^d$ .

The above setting incorporates rather varied physical contexts. The spins may be thought of as describing magnetic moments of atoms in a crystal, displacement of atoms from their equilibrium positions or even orientation of grains in nearly-crystalline granular materials.

To define the dynamics of spin systems, we will need to specify the energetics. This is conveniently done by prescribing the *Hamiltonian* which is a function on the spin-configuration space  $\Omega^{\mathbb{Z}^d}$  that tells us how much energy each spin configuration has. Of course, to have all quantities well defined we need to fix a *finite* volume  $\Lambda \subset \mathbb{Z}^d$  and compute only the energy in  $\Lambda$ . The most general formula we will ever need is

$$H_\Lambda(S) = \sum_{\substack{A \subset \mathbb{Z}^d \text{ finite} \\ A \cap \Lambda \neq \emptyset}} \Phi_A(S) \quad (1.2)$$

where  $\Phi_A$  is a function that depends only on  $S_A$ . To make everything well defined, we require, e.g., that  $\Phi_A$  is translation invariant and that  $\sum_{A \ni 0} \|\Phi_A\|_\infty < \infty$ . (The infinity norm may be replaced by some other norm, should the need arise to talk about unbounded spins.) It is often more convenient to write the above as a formal sum

$$H(S) = \sum_A \Phi_A(S) \quad (1.3)$$

with the above specific understanding whenever a rigorous definition is desired.

The energy is not sufficient on its own to define the statistical mechanics of such spin systems; we also need to specify the *a priori measure* on the spins. This will be achieved by prescribing a Borel measure  $\mu_0$  on  $\Omega$  (which may or may not be finite). The spin configurations (in finite volume) will be “distributed” according to the product measure, e.g., the *a priori* law of  $S_\Lambda$  is  $\bigotimes_{x \in \Lambda} \mu_0(dS_x)$ . The full statistical-mechanical law is then given by a *Gibbs measure* which takes the form  $e^{-\beta H(S)} \prod_x \mu_0(dS_x)$ ; cf Sect. 1.3 for more details.

## 1.2 Examples

Here are a few examples of spin systems:

(1) *O(n)-model*: Here  $\Omega = \mathbb{S}^{n-1} = \{z \in \mathbb{R}^n : |z|_2 = 1\}$  with  $\mu_0 =$  surface measure. The Hamiltonian is

$$H(S) = -J \sum_{\langle x, y \rangle} S_x \cdot S_y \quad (1.4)$$

where the dot denotes the usual (Euclidean) dot-product in  $\mathbb{R}^n$  and  $J \geq 0$ . (The sign of  $J$  is reversed by reversing the spins on the odd sublattice of  $\mathbb{Z}^d$ .)

Note that if  $A \in O(n)$ —i.e.,  $A$  is an  $n$ -dimensional orthogonal matrix—then

$$AS_x \cdot AS_y = S_x \cdot S_y \quad (1.5)$$

and so  $H(AS) = H(S)$ . Since also  $\mu_0 \circ A^{-1} = \mu_0$ , the model possesses a *global rotation invariance* (with respect to simultaneous rotation of all spins).

Two instances of this model are known by other names:  $n = 2$  is the *rotor model* while  $n = 3$  is the (classical) *Heisenberg ferromagnet*.

(2) *Ising model*: Formally, this is the  $O(1)$ -model. Explicitly, the spin variables  $\sigma_x$  take values in  $\Omega = \{-1, +1\}$  with uniform *a priori* measure; the Hamiltonian is

$$H(\sigma) = -J \sum_{\langle x, y \rangle} \sigma_x \sigma_y \quad (1.6)$$

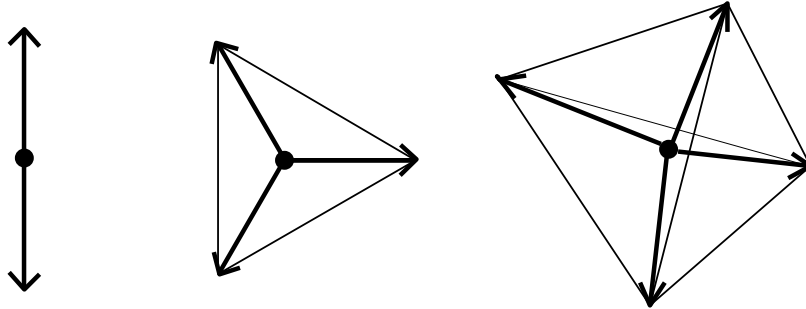
Note that the energy is smaller when the spins at nearest neighbors align and higher when they antialign. A similar statement holds, of course, for all  $O(n)$  models. This is due to the choice of the sign  $J \geq 0$  which makes these models *ferromagnets*.

(3) *Potts model*: This is a generalization of the Ising model beyond two spin states. Explicitly, we fix  $q \in \mathbb{N}$  and let  $\sigma_x$  take values in  $\{1, \dots, q\}$  (with uniform *a priori* measure). The Hamiltonian is

$$H(\sigma) = -J \sum_{\langle x, y \rangle} \delta_{\sigma_x, \sigma_y} \quad (1.7)$$

so the energy is  $-J$  when  $\sigma_x$  and  $\sigma_y$  “align” and zero otherwise. The  $q = 2$  case is the Ising model and  $q = 1$  may be related to bond percolation on  $\mathbb{Z}^d$  (via the so called *Fortuin-Kasteleyn representation* leading to a *random-cluster model*).

It turns out that the Hamiltonian (1.7) can be brought to the form (1.4). Indeed, let  $\Omega$  denote the set of  $q$  points uniformly spread on the unit sphere in  $\mathbb{R}^{q-1}$ ; we may think of these as the vertices of a  $q$ -simplex (or a regular  $q$ -hedron). The cases  $q = 2, 3, 4$  are depicted in this figure:



More explicitly, the elements of  $\Omega$  are vectors  $\hat{v}_\alpha$ ,  $\alpha = 1, \dots, q$ , such that

$$\hat{v}_\alpha \cdot \hat{v}_\beta = \begin{cases} 1, & \text{if } \alpha = \beta, \\ -\frac{1}{q-1}, & \text{otherwise.} \end{cases} \quad (1.8)$$

The existence of such vectors may be proved by induction on  $q$ . Clearly, if  $S_x$  corresponds to  $\sigma_x$  and  $S_y$  to  $\sigma_y$ , then

$$S_x \cdot S_y = \frac{q}{q-1} \delta_{\sigma_x, \sigma_y} - \frac{1}{q-1} \quad (1.9)$$

and so the Potts Hamiltonian is to within an additive constant of

$$H(S) = -\tilde{J} \sum_{\langle x, y \rangle} S_x \cdot S_y \quad (1.10)$$

with  $\tilde{J} = J \frac{q-1}{q}$ . This form will be far more useful for our purposes than (1.7).

(4) Liquid-crystal model: There are many models that describe certain granular materials known to many of us from digital displays: liquid crystals. A distinguished feature of such materials is the presence of *orientational long-range order* where a majority of the grains align with one another despite the fact that the system as a whole is rotationally invariant. One of the simplest models capturing this phenomenon is as follows: Consider spins  $S_x \in \mathbb{S}^{n-1}$  with uniform *a priori* measure. The Hamiltonian is

$$H(S) = -J \sum_{\langle x,y \rangle} (S_x \cdot S_y)^2 \quad (1.11)$$

The interaction features global rotation invariance and the square takes care of the fact that reflection of any spin does not change the energy (i.e., only the *orientation* rather than the *direction* of the spin matters).

As for the Potts model, the Hamiltonian can again be brought to the form reminiscent of the  $O(n)$ -model. Indeed, given a spin  $S \in \mathbb{S}^{n-1}$  with Cartesian components  $S^{(\alpha)}$ ,  $\alpha = 1, \dots, n$ , define an  $n \times n$  matrix  $Q$  by

$$Q_{\alpha\beta} = S^{(\alpha)} S^{(\beta)} - \frac{1}{n} \delta_{\alpha\beta} \quad (1.12)$$

(The subtraction of the identity is rather arbitrary at this point; the goal is to achieve zero trace and thus reduce the number of independent variables characterizing  $Q$  to  $n - 1$ ; i.e., exactly as many degrees of freedom as  $S$  has.) As is easy to check, if  $Q \leftrightarrow S$  and  $\tilde{Q} \leftrightarrow \tilde{S}$  are related via the above formula, then

$$\text{Tr}(Q\tilde{Q}) = (S \cdot \tilde{S})^2 - \frac{1}{n} \quad (1.13)$$

Since  $Q$  is symmetric, the trace evaluates to

$$\text{Tr}(Q\tilde{Q}) = \sum_{\alpha,\beta} Q_{\alpha\beta} \tilde{Q}_{\alpha\beta} \quad (1.14)$$

which is the canonical scalar product on  $n \times n$  matrices. In such language the Hamiltonian again takes the form known from the  $O(n)$  model.

At the point we pause to remark that all of the above Hamiltonians are of the following rather general form:

$$H(S) = +\frac{1}{2} \sum_{x,y} J_{x,y} |S_x - S_y|^2 \quad (1.15)$$

where  $(J_{x,y})$  is a collection of suitable *coupling constants*. This is possible because, in all cases, the norm of  $S_x$  is constant. The model thus obtained bears striking similarity to our last example:

(5) Gaussian free field: Let  $\Omega = \mathbb{R}$ ,  $\mu_0 =$  Lebesgue measure and let  $P(x, y)$  be the transition kernel of a random walk on  $\mathbb{Z}^d$ ; i.e.,  $P(x, y) = P(0, y - x)$ . In this case we will denote the

variables by  $\phi_x$ ; the Hamiltonian is

$$H(\phi) = \frac{1}{2} \sum_{x,y} P(x,y) (\phi_y - \phi_x)^2 \quad (1.16)$$

This can be rewritten as

$$H(\phi) = (\phi, (1 - P)\phi)_{L^2(\mathbb{Z}^d)} =: \mathcal{E}_{1-P}(\phi, \phi) \quad (1.17)$$

where experts on harmonic analysis of Markov chains will recognize  $\mathcal{E}_{1-P}(\phi, \phi)$  to be the *Dirichlet form* associated with the generator  $1 - P$  of the above random walk. In the Gibbs measure, the law of the  $\phi_x$ 's will be Gaussian with grad-squared interactions; hence the name of the model.

The sole difference between (1.15) and (1.16) is that, unlike the  $\phi_x$ 's, the spins  $S_x$  are generally confined to a subset of a Euclidean space and/or their *a priori* measure is not Lebesgue—which will ultimately mean their law is *not* Gaussian. One purpose of this course is to show how this similarity can nevertheless be exploited to provide information on the models (1.15).

### 1.3 Gibbs formalism

Now we are ready to describe the statistical-mechanical properties of the above models for which we resort to the formalism of Gibbs-Boltzmann distributions. First we define these in finite volume: Given a finite set  $\Lambda \subset \mathbb{Z}^d$  and a boundary condition  $S_{\Lambda^c}$  we define the *Gibbs measure* in  $\Lambda$  to be the measure on  $\Omega^\Lambda$  given by

$$\mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(dS_\Lambda) = \frac{e^{-\beta H_\Lambda(S)}}{Z_{\Lambda,\beta}(S_{\Lambda^c})} \prod_{x \in \Lambda} \mu_0(dS_x) \quad (1.18)$$

Here  $\beta \geq 0$  is the *inverse temperature*—in physics terms,  $\beta = \frac{1}{k_B T}$  where  $k_B$  is the Boltzmann constant and  $T$  is the temperature measured in Kelvins—and  $Z_{\Lambda,\beta}(S_{\Lambda^c})$  is the normalization constant called the *partition function*.

To extend this concept to infinite volume we have two options:

- (1) Consider all possible weak cluster points of the family  $\{\mu_{\Lambda,\beta}^{(S_{\Lambda^c})}\}$  as  $\Lambda \uparrow \mathbb{Z}^d$  (with the boundary condition possibly varying with  $\Lambda$ ) and all convex combinations thereof.
- (2) Identify a distinguishing property of Gibbs measures and use it to define infinite volume objects directly.

While approach (1) is ultimately very useful in practical problems, option (2) is more elegant at this level of generality. The requisite “distinguishing property” is as follows:

**Lemma 1.1 [DLR condition]** *Let  $\Lambda \subset \Delta \subset \mathbb{Z}^d$  be finite sets and let  $S_{\Delta^c} \in \Omega^{\Delta^c}$ . Then (for  $\mu_{\Delta,\beta}^{(S_{\Delta^c})}$ -a.e.  $S_{\Lambda^c}$ ),*

$$\mu_{\Delta,\beta}^{(S_{\Delta^c})}(\cdot | S_{\Lambda^c}) = \mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\cdot) \quad (1.19)$$

In simple terms, conditioning the Gibbs measure in  $\Delta$  on the configuration in  $\Delta \setminus \Lambda$ , we get the Gibbs measure in  $\Lambda$  with the corresponding boundary condition.

This leads to:

**Definition 1.2 [DLR Gibbs measures]** A probability measure on  $\Omega^{\mathbb{Z}^d}$  is called an infinite volume Gibbs measure for interaction  $H$  and inverse temperature  $\beta$  if for all finite  $\Lambda \subset \mathbb{Z}^d$  and  $\mu$ -a.e.  $S_{\Lambda^c}$ ,

$$\mu(\cdot | S_{\Lambda^c}) = \mu_{\Lambda, \beta}^{(S_{\Lambda^c})}(\cdot) \quad (1.20)$$

where  $\mu_{\Lambda, \beta}^{(S_{\Lambda^c})}$  is defined using the Hamiltonian  $H_\Lambda$ .

We will use  $\mathfrak{G}_\beta$  to denote the set of all infinite volume Gibbs measures at inverse temperature  $\beta$  (assuming the model is clear from the context).

Here are some straightforward, nonetheless important consequences of these definitions:

- (1) As a consequence of Lemma 1.1, any weak cluster point of  $(\mu_{\Lambda, \beta}^{(S_{\Lambda^c})})$  is in  $\mathfrak{G}_\beta$ .
- (2) By the Backward Martingale Convergence Theorem, if  $\Lambda_n \uparrow \mathbb{Z}^d$  and  $\mu \in \mathfrak{G}_\beta$ , then for  $\mu$ -a.e. spin configuration  $S$  the sequence  $\mu_{\Lambda_n, \beta}^{(S_{\Lambda_n^c})}$  has a weak limit.
- (3)  $\mathfrak{G}_\beta$  is a convex set. Moreover,  $\mu \in \mathfrak{G}_\beta$  is extremal in  $\mathfrak{G}_\beta$  iff  $\mu_{\Lambda_n, \beta}^{(S_{\Lambda_n^c})} \xrightarrow{w} \mu$  for  $\mu$ -almost every spin configuration  $S$ .

Similarly direct is the proof of the following ‘‘continuity’’ property:

- (4) Let  $H_n$  be a sequence of Hamiltonians converging—in the sup-norm on the potentials  $\Phi_A$ —to Hamiltonian  $H$ , and  $\beta_n$  is a sequence with  $\beta_n \rightarrow \beta < \infty$ . Let  $\mu_n$  be a sequence of corresponding Gibbs measures. Then every (weak) cluster point of  $(\mu_n)$  is an infinite-volume Gibbs measure for Hamiltonian  $H$  and inverse temperature  $\beta$ .

Now we give a meaning to the terms often used somewhat vaguely by physicists:

**Definition 1.3 [Phase coexistence]** We say that the model is at phase coexistence (or undergoes a 1st-order phase transition) whenever the parameters are such that  $|\mathfrak{G}_\beta| > 1$ .

The simplest example where this happens is the Ising model. Let  $\Lambda_L = \{1, \dots, L\}^d$  and consider the Ising model in  $\Lambda_L$  with all boundary spins set to  $+1$ . This is the so called *plus boundary condition*. As a consequence of stochastic domination—which we will not discuss here— $\mu_{\Lambda_L, \beta}^+$  tends weakly to a measure  $\mu^+$  as  $L \rightarrow \infty$ . Similarly,  $\mu_{\Lambda_L, \beta}^- \xrightarrow{w} \mu^-$ . It turns out that, in dimensions  $d \geq 2$  there exists  $\beta_c(d) \in (0, \infty)$  such that

$$\beta > \beta_c(d) \quad \Rightarrow \quad \mu^+ \neq \mu^- \quad (1.21)$$

while for  $\beta < \beta_c(d)$ , the set of all infinite volume Gibbs measures is a singleton. We will prove similar statements in all of the models introduced above.

## 1.4 Torus measures

In the above, we always put a boundary condition in the complement of the finite set  $\Lambda$ . However, it is sometimes convenient to consider other boundary conditions. One possibility is to ignore the existence of  $\Lambda^c$  altogether—this leads to the so called *free boundary condition*. Another possibility is to wrap  $\Lambda$  into a graph without boundary—typically a torus. This is the case of *periodic* or *torus boundary conditions*.

Consider the torus  $\mathbb{T}_L$  which we define as  $(\mathbb{Z}/L\mathbb{Z})^d$  endowed with the corresponding (periodicized) nearest-neighbor relation. For nearest-neighbor interactions, the corresponding Hamiltonian is defined easily, but some care is needed for interactions that can be of arbitrary range. If  $S \in \Omega^{\mathbb{T}_L}$  we define the *torus Hamiltonian*  $H_L(S)$  by

$$H_L(S) = H_{\Lambda_L}(\text{periodic extension of } S \text{ to } \mathbb{Z}^d) \quad (1.22)$$

where we recall  $\Lambda_L = \{1, \dots, L\}^d$ . For  $H(S) = \frac{1}{2} \sum_{x,y} J_{x,y} S_x \cdot S_y$  we thus get

$$H_L(S) = \frac{1}{2} \sum_{x,y} J_{x,y}^{(L)} S_x \cdot S_y \quad (1.23)$$

where  $J_{x,y}^{(L)}$  are the periodicized coupling constants

$$J_{x,y}^{(L)} = \sum_{z \in \mathbb{Z}^d} J_{x,y+Lz} \quad (1.24)$$

The Gibbs measure on  $\Omega^{\mathbb{T}_L}$  is then defined accordingly:

$$\mu_{L,\beta}(dS) = \frac{e^{-\beta H_L(S)}}{Z_{L,\beta}} \prod_{x \in \mathbb{T}_L} \mu_0(dS_x) \quad (1.25)$$

where  $Z_{L,\beta}$  is the torus partition function. The following holds:

**Lemma 1.4** *Every (weak) cluster point of  $(\mu_{L,\beta})_{L \geq 1}$  lies in  $\mathfrak{G}_\beta$ .*

There *is* something to prove here because, due to (1.24), the interaction depends on  $L$ .

## 1.5 Some thermodynamics

For historical, and also practical reasons, many accounts of statistical mechanics start with the notion of free energy. We will need this notion only tangentially—it suffices to think of the free energy as a cumulant generating function—in the proofs of phase coexistence. The relevant statement is as follows:

**Theorem 1.5** *For  $x \in \mathbb{Z}^d$  let  $\tau_x$  be the shift-by- $x$  which is defined by  $(\tau_x S)_y = S_{y-x}$ . Let  $g: \Omega^{\mathbb{Z}^d} \rightarrow \mathbb{R}$  be a bounded, local function—i.e., one that depends only on a finite number of spins—and let  $\mu_{L,\beta}$  be the torus Gibbs measures. Then:*

(1) *The limit*

$$f(h) = \lim_{L \rightarrow \infty} \frac{1}{L^d} \log E_{\mu_{L,\beta}} \left\{ \exp \left( h \sum_{x \in \mathbb{T}_L} g \circ \tau_x \right) \right\} \quad (1.26)$$

exists for all  $h \in \mathbb{R}$  and is convex in  $h$ .

(2) *If  $\mu \in \mathfrak{G}_\beta$  is translation invariant, then*

$$\frac{\partial f}{\partial h^-} \Big|_{h=0} \leq E_\mu(g) \leq \frac{\partial f}{\partial h^+} \Big|_{h=0} \quad (1.27)$$

(3) *There exist translation-invariant, ergodic measures  $\mu^\pm \in \mathfrak{G}_\beta$  such that*

$$E_{\mu^\pm}(g) = \frac{\partial f}{\partial h^\pm} \Big|_{h=0} \quad (1.28)$$

*Proof of (1).* For compact state-spaces and absolutely-summable interactions, the existence of the limit follows by standard subadditivity arguments. In fact, the measure  $\mu_{L,\beta}$  could be replaced by any sequence of Gibbs measures in  $\Lambda_L$  with (even variable) boundary conditions. The convexity of  $f$  follows by the Hölder inequality.  $\square$

*Proof of (2).* Let  $\mu \in \mathfrak{G}_\beta$  be translation invariant and abbreviate

$$Z_L(h) = E_\mu \left\{ \exp \left( h \sum_{x \in \Lambda_L} g \circ \tau_x \right) \right\} \quad (1.29)$$

Since  $\log Z_L$  is convex in  $h$ , we have for any  $h > 0$  that

$$\begin{aligned} \log Z_L(h) - \log Z_L(0) &\geq \frac{\partial}{\partial h} \log Z_L(h) \Big|_{h=0} h \\ &= h E_\mu \left( \sum_{x \in \Lambda_L} g \circ \tau_x \right) = h |\Lambda_L| E_\mu(g). \end{aligned} \quad (1.30)$$

Dividing by  $|\Lambda_L|$ , passing to  $L \rightarrow \infty$  and using that  $f$  is independent of the boundary condition, we get

$$f(h) - f(0) \geq h E_\mu(g) \quad (1.31)$$

Divide by  $h$  and let  $h \downarrow 0$  to get one half of (1.27). The other half is proved analogously.  $\square$

*Proof of (3).* Let  $\mathfrak{G}_{\beta,h}$  be the set of Gibbs measures for the Hamiltonian  $H - (h/\beta) \sum_x g \circ \tau_x$ . A variant of proof of (2) shows that if  $\mu_h \in \mathfrak{G}_{\beta,h}$  is translation-invariant, then

$$\frac{\partial f}{\partial h^-} \leq E_{\mu_h}(g) \leq \frac{\partial f}{\partial h^+} \quad (1.32)$$

In particular, if  $h > 0$  we have

$$E_{\mu_h}(g) \geq \frac{\partial f}{\partial h^-} \Big|_{h>0} \geq \frac{\partial f}{\partial h^+} \Big|_{h=0} \quad (1.33)$$

by the monotonicity of derivatives of convex functions. Taking  $h \downarrow 0$  and extracting a weak limit from  $\mu_h$ , we get a Gibbs measure  $\mu^+ \in \mathfrak{G}_\beta$  such that

$$E_{\mu^+}(g) \geq \left. \frac{\partial f}{\partial h^+} \right|_{h=0} \quad (1.34)$$

(The expectations converge because  $g$  is a local—and thus continuous, in the product topology—function.) Applying (2) we verify (1.28) for  $\mu^+$ .

The measure  $\mu^+$  is translation invariant and so it remains to show that  $\mu^+$  can actually be chosen ergodic. To that end let us first prove that

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} g \circ \tau_x \xrightarrow{L \rightarrow \infty} E_{\mu^+}(g), \quad \text{in } \mu^+\text{-probability} \quad (1.35)$$

The random variables on the left are bounded by the norm of  $g$  and have expectation  $E_{\mu^+}(g)$  so it suffices to prove that the limsup is no larger than the expectation. However, if that were not the case, we would have

$$\mu^+ \left( \sum_{x \in \Lambda_L} g \circ \tau_x > (E_{\mu^+}(g) + \epsilon) |\Lambda_L| \right) > \epsilon \quad (1.36)$$

for some  $\epsilon > 0$  and some sequence of  $L$ 's. But then for all  $h > 0$ ,

$$E_{\mu^+} \left\{ \exp \left( h \sum_{x \in \Lambda_L} g \circ \tau_x \right) \right\} \geq \epsilon e^{|\Lambda_L| h [E_{\mu^+}(g) + \epsilon]} \quad (1.37)$$

This implies

$$f(h) \geq h(E_{\mu^+}(g) + \epsilon) \quad (1.38)$$

which cannot hold for all  $h > 0$  should the right-derivative of  $f$  at  $h = 0$  be equal  $E_{\mu^+}(g)$ . Hence (1.35) holds.

By the Pointwise Ergodic Theorem, the convergence in (1.35) actually occurs—and, by (1.35), the limit equals  $E_{\mu^+}(g)$ —for  $\mu^+$ -almost every spin configuration. This implies that the same must be true for any measure in the decomposition of  $\mu^+$  into ergodic components. By classic theorems from Gibbs-measure theory, every measure in this decomposition is also in  $\mathfrak{G}_\beta$  and so we can choose  $\mu^+$  ergodic.  $\square$

The above theorem is very useful for the proofs of phase coexistence. Indeed, one can often prove some estimates that via (1.27) imply that  $f$  is not differentiable at  $h = 0$ . Then one applies (1.28) to infer the existence of two distinct, ergodic Gibbs measures saturating the bounds in (1.27). Examples of this approach will be discussed momentarily.

## 1.6 Literature remarks

This chapter contains only the absolute minimum we need for understanding the rest of the course. For a comprehensive treatment of Gibbs-measure theory, we refer to classic monographs by Israel [48], Simon [64] and Georgii [43]; further general background on statistical

mechanics of such systems can be found in Ruelle’s “blue” book [58]. The acronym DLR derives from the initials of Dobrushin and the team of Lanford & Ruelle who first introduced the idea of conditional definition of infinite volume Gibbs measures; cf e.g. [21].

The  $O(n)$  model goes back to Heisenberg (who introduced its quantum version); the Ising model was introduced by Lenz and given to Ising as a thesis problem and the Potts model was introduced by Domb and given to Potts as a thesis problem. An excellent reference for mathematical physics of liquid crystals is the monograph by de Gennes and Prost [42]. The tetrahedral representation of the Potts model can be found in Wu’s review article [70] on the Potts model; the matrix representation of the liquid-crystal model goes back to Angelescu and Zagrebnov [4]. Gradient fields—of which the GFF is the simplest example—have enjoyed considerable attention in recent years; cf the review articles by Funaki [38], Velenik [68] and Sheffield [60]. The GFF is sometimes called the *harmonic crystal*.

## Chapter 2

# Infrared bound & spin-wave condensation

The goal of this chapter is to explain the concept of infrared bound—postponing the proof until Chapter 4—and its use in the proof of symmetry breaking via the mechanism of “spin-wave condensation.” The presence, and absence, of symmetry breaking in  $O(n)$ -model with certain non-negative two-body interactions will be linked to recurrence vs transience of a naturally induced random walk.

### 2.1 Random walk connections

Consider the model with the Hamiltonian

$$H = -\frac{1}{2} \sum_{x,y} J_{xy} S_x \cdot S_y \quad (2.1)$$

where the spins  $S_x$  are *a priori* distributed according to a measure  $\mu_0$  which is supported in a compact set  $\Omega \subset \mathbb{R}^v$ . The interaction constants satisfy the following requirements:

- (I1)  $J_{xx} = 0$  and  $J_{x,y} = J_{0,y-x}$
- (I2)  $\sum_x |J_{0,x}| < \infty$  and  $\sum_x J_{0,x} = 1$

i.e., the coupling constants are translation invariant, absolutely summable and, for convenience, normalized. We will actually always consider the following specific examples:

- *n.n. interactions:*

$$J_{x,y} = \begin{cases} \frac{1}{2d}, & \text{if } |x - y| = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.2)$$

- *Yukawa potentials:*

$$J_{x,y} = C e^{-\mu|x-y|_1} \quad (2.3)$$

with  $\mu > 0$  and  $C > 0$ .

- *Power-law decaying potentials:*

$$J_{x,y} = \frac{C}{|x - y|^s} \quad (2.4)$$

with  $s > d$  and  $C > 0$ .

On top of these, we will also permit:

- Any convex combination of the above (with, of course, positive coefficients).

Note that we are using the  $\ell_1$ -distance (rather than the more natural  $\ell_2$ -distance). This is dictated by our methods of proof. Also note that the Yukawa potential is in the class of *Kac models* where the coupling constants take the form  $J_{x,y} = \gamma^d f(\gamma(x - y))$  for some rapidly decaying function  $f: \mathbb{R}^d \rightarrow [0, \infty)$  with unit  $L^1$ -norm

A unifying feature of all three examples is that  $J_{xy} \geq 0$  which allows us to interpret the coupling constants as the *transition probabilities* of a random walk on  $\mathbb{Z}^d$ . Explicitly, consider a Markov chain  $(X_n)$  on  $\mathbb{Z}^d$  with

$$P_z(X_{n+1} = y | X_n = x) = J_{xy} \quad (2.5)$$

where  $P_z$  is the law of the chain started at site  $z$ . Of particular interest will the question whether this random walk is recurrent or transient. Here is a criterion to this matter:

**Lemma 2.1** *Let  $\hat{J}(k) = \sum_x J_{0,x} e^{ik \cdot x}$ ,  $k \in [-\pi, \pi]^d$ . Then  $(X_n)$  is transient if and only if*

$$\int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} < \infty. \quad (2.6)$$

*Proof.* Recall that a random walk is transient iff the first return time to the origin,  $\tau_0 = \inf\{n > 0: X_n = 0\}$ , is infinite with positive probability,  $P_0(\tau_0 < \infty) < 1$ , which happens iff the number of visits back to the origin,  $N = \sum_{n \geq 0} \mathbf{1}_{\{X_n = 0\}}$ , is finite almost surely. By the formula  $E_0 N = [1 - P_0(\tau_0 < \infty)]^{-1}$ , this is also equivalent to  $E_0 N < \infty$ . To compute the expectation, we first note that

$$\mathbf{1}_{\{X_n = 0\}} = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} e^{ik \cdot X_n} \quad (2.7)$$

which via  $E_0 e^{ik \cdot X_n} = [E_0 e^{ik \cdot X_1}]^n = [\sum_x J_{0,x} e^{ik \cdot x}]^n = \hat{J}(k)^n$  implies

$$P_0(X_n = 0) = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \hat{J}(k)^n. \quad (2.8)$$

Summing over  $n \geq 0$  yields

$$E_0 N = \sum_{n \geq 0} \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \hat{J}(k)^n = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} \quad (2.9)$$

whereby the claim follows. (A careful proof of the latter identity requires justification of the exchange of the integral with the infinite sum; one has to represent the LHS as a power series, perform the sum and justify limits via appropriate convergence theorems.)  $\square$

As to the above examples, we have:

- *n.n. & Yukawa potentials:* As  $k \rightarrow 0$ ,

$$1 - \hat{J}(k) \sim C|k|^2 \quad (2.10)$$

and so  $(X_n)$  is transient iff  $d \geq 3$ .

- *Power-law potentials:* Here as  $k \rightarrow 0$ ,

$$1 - \hat{J}(k) \sim C \begin{cases} |k|^{s-d}, & \text{if } s < d + 2, \\ |k|^2 \log \frac{1}{|k|}, & \text{if } s = d + 2, \\ |k|^2, & \text{if } s > d + 2. \end{cases} \quad (2.11)$$

Hence  $(X_n)$  is transient iff  $d \geq 3$  OR  $s < \min\{d + 2, 2d\}$ .

(Note that the walk with  $s < d + 2$  has a stable-law tail with parameter  $\alpha = s - d$ .) A convex combination (with non-zero coefficients) of the three coupling constants will lead to a transient walk provided at least one of the interactions involved therein is transient.

## 2.2 Infrared bound

The principal claim of this chapter is that the finiteness of the integral in (2.6) is sufficient for the existence of a symmetry-breaking phase transition in many spin systems of the kind (2.1). The reason is the connection of the above random walk to the Gaussian free field (1.16) (GFF) with  $P(x, y) = J_{xy}$ . Indeed, consider the field in a square box  $\Lambda$  with, say, zero boundary condition. It turns out that

$$\text{Cov}_\Lambda(\phi_x, \phi_y) = \sum_{n \geq 0} P_x(X_n = y, \tau_{\Lambda^c} = y) =: G_\Lambda(x, y) \quad (2.12)$$

where  $\tau_{\Lambda^c}$  is the first exit time of the walk from  $\Lambda$  and  $G_\Lambda$  denotes the so called *Green's function*. In particular, we have

$$\text{Var}_\Lambda(\phi_0) = G_\Lambda(0, 0) \quad (2.13)$$

which, as we will see, tends to the integral (2.6) as  $\Lambda \uparrow \mathbb{Z}^d$ . Since  $E_\Lambda(\phi_0) = 0$  due to our choice of the boundary condition, we conclude

$$(\phi_0)_{\Lambda \subset \mathbb{Z}^d} \text{ is tight iff } (X_n) \text{ is transient} \quad (2.14)$$

Physicists actually prefer to think of this in terms of symmetry breaking: Formally, the Hamiltonian of the GFF is invariant under the transformation  $\phi_x \rightarrow \phi_x + c$ , i.e., the model possesses a global spin-translation symmetry. The symmetry group is not compact and so, to define the model even in finite volume, the symmetry needs to be broken by boundary conditions. The existence of a limit law for  $\phi_0$  means the breaking survives the thermodynamic limit (while non-existence means that the invariance is restored in this limit).

Our goal is to show that qualitatively the same conclusions hold also for the  $O(n)$ -spin system. Explicitly, we will prove:

**Theorem 2.2** *Let  $(J_{xy})$  be one of the 3 interactions above. Then:*

$$\begin{array}{l} \text{Global rotation symmetry} \\ \text{of } O(n)\text{-model is broken} \\ \text{at low temperatures} \end{array} \iff \begin{array}{l} \text{Random walk driven} \\ \text{by } (J_{xy}) \text{ is transient} \end{array}$$

We begin with the proof of the implication  $\Leftarrow$ . The principal tool will be our next theorem which, for technical reasons, is formulated for torus boundary conditions:

**Theorem 2.3 [Infrared bound]** *Let  $L$  be an even integer and consider the model (2.1) on torus  $\mathbb{T}_L$  with Gibbs measure  $\mu_{L,\beta}$ . Suppose  $(J_{xy})$  is one of the three interactions above and let*

$$c_{L,\beta}(x) = E_{\mu_{L,\beta}}(S_0 \cdot S_x). \quad (2.15)$$

Define  $\hat{c}_{L,\beta}(k) = \sum_{x \in \mathbb{T}_L} c_{L,\beta}(x) e^{ik \cdot x}$ . Then

$$\hat{c}_{L,\beta}(k) \leq \frac{\nu}{2\beta} \frac{1}{1 - \hat{J}(k)}, \quad k \in \mathbb{T}_L^* \setminus \{0\} \quad (2.16)$$

where  $\nu$  is the dimension of the spin vectors and  $\mathbb{T}_L^* = \{\frac{2\pi}{L}(n_1, \dots, n_d) : 0 \leq n_i < L\}$  is the reciprocal torus.

The proof will require developing the technique of reflection positivity and is therefore postponed to Chapter 4.

Note that  $c_{L,\beta}(x)$  is the spin-spin correlation function which, in light of translation invariance of  $\mu_{L,\beta}$  is a function of only the displacement of the two spins. The result has the following equivalent formulation: For all  $(v_x) \in \mathbb{C}^{\mathbb{T}_L}$  with  $\sum_x v_x = 0$ ,

$$\sum_{x,y \in \mathbb{T}_L} v_x \bar{v}_y E_{\mu_{L,\beta}}(S_0 \cdot S_x) \leq \frac{\nu}{2\beta} \sum_{x,y \in \mathbb{T}_L} v_x \bar{v}_y G_L(x,y) \quad (2.17)$$

where

$$G_L(x,y) = \frac{1}{L^d} \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \frac{e^{ik \cdot (x-y)}}{1 - \hat{J}(k)} \quad (2.18)$$

Observe that the latter is the covariance matrix of the GFF on  $\mathbb{T}_L$ , projected on configurations with total integral zero (i.e., on the orthogonal complement of constant functions). This is a meaningful object because while the  $\phi_x$  are not really well defined—due to the absence of the boundary—the differences  $\phi_y - \phi_x$  are.

A short formulation of the infrared bound is thus:

*The correlation of the spins in models (2.1) with one of the three interactions above is dominated—as a matrix on the orthogonal complement of constant functions in  $L^2(\mathbb{T}_L)$ —by the covariance of the GFF.*

This fact is often referred to as *Gaussian domination*.

### 2.3 Spin-wave condensation in $O(n)$ -model

Now we will continue in our original line of thought. The above theorem implies:

**Corollary 2.4 [Spin-wave condensation]** *Suppose  $|S_x| = 1$  for all  $x$ . Then*

$$E_{\mu_L, \beta} \left( \left| \frac{1}{L^d} \sum_{x \in \mathbb{T}_L} S_x \right|^2 \right) \geq 1 - \frac{\nu}{2\beta} G_L(0, 0). \quad (2.19)$$

*Proof.* Let  $\hat{S}_k = \sum_{x \in \mathbb{T}_L} S_x e^{ik \cdot x}$  be the Fourier coefficient of the decomposition of  $(S_x)$  into the so called *spin waves*. The IRB yields

$$E_{\mu_L, \beta} |\hat{S}_k|^2 \leq \frac{\nu}{2\beta} \frac{L^d}{1 - \hat{J}(k)}, \quad k \in \mathbb{T}_L^* \setminus \{0\}. \quad (2.20)$$

On the other hand, Parseval's identity along with  $|S_x| = 1$  implies

$$\sum_{k \in \mathbb{T}_L^*} |\hat{S}_k|^2 = L^d \sum_{x \in \mathbb{T}_L} |S_x|^2 = L^{2d}. \quad (2.21)$$

The IRB makes no statement about  $\hat{S}_0$  so we split it from the rest of the sum:

$$\frac{1}{L^{2d}} |\hat{S}_0|^2 = 1 - \frac{1}{L^{2d}} \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} |\hat{S}_k|^2 \quad (2.22)$$

Now take expectation and apply (2.20):

$$E_{\mu_L, \beta} \left( \frac{1}{L^{2d}} |\hat{S}_0|^2 \right) \geq 1 - \frac{\nu}{2\beta} \frac{1}{L^d} \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \frac{1}{1 - \hat{J}(k)}. \quad (2.23)$$

In light of (2.18), this is (2.19).  $\square$

With (2.19) in the hand we can apply the same reasoning as for the GFF: In the transient cases,  $G_L(0, 0)$  converges to the integral (2.6) and so the right-hand side has a finite limit. By taking  $\beta$  sufficiently large, the limit is actually strictly positive uniformly in  $L$ . This in turn implies that the zero mode of the spin-wave decomposition is *macroscopically* populated—very much like the ground state of the free Bose gas at Bose-Einstein condensation. Here is how we pull the corresponding conclusions from  $\mathbb{T}_L$  onto  $\mathbb{Z}^d$ :

**Theorem 2.5 [Phase coexistence in  $O(n)$ -model]** *Consider  $O(n)$ -model with  $n \geq 1$  and one of the three interactions above. Let*

$$\beta_0 = \frac{n}{2} \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} \quad (2.24)$$

*Then for any  $\beta > \beta_0$  and any  $\theta \in \mathbb{S}^{n-1}$  there exists  $\mu_\theta \in \mathfrak{G}_\beta$  which is translation invariant and ergodic such that*

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} S_x \xrightarrow[L \rightarrow \infty]{} m_\star \theta, \quad \mu_\theta\text{-a.s.}, \quad (2.25)$$

*for some  $m_\star = m_\star(\beta) > 0$ .*

Note that (2.25) implies that the measures  $\mu_\theta$  are mutually singular with respect to one another.

*Proof.* Suppose, without loss of generality, that we are in the transient case, i.e.,  $\beta_0 < \infty$ . The idea of the proof is quite simple: We use (2.19) to show that the free energy is not differentiable in an appropriate external field when this field is set to zero. Then we apply Theorem 1.5 to conclude the existence of the required distinct, ergodic Gibbs measures.

Fix  $\theta \in \mathbb{S}^{n-1}$  and let

$$f(h) = \lim_{L \rightarrow \infty} \frac{1}{L^d} \log E_{\mu_{L,\beta}}[e^{h\theta \cdot \hat{S}_0}]. \quad (2.26)$$

(The limit exists by Theorem 1.5.) We want to show that  $\frac{\partial}{\partial h^+} f(0) > 0$  (and thus, by symmetry,  $\frac{\partial}{\partial h^-} f(0) < 0$ ). Corollary 2.4 yields

$$E_{\mu_{L,\beta}}(L^{-2d} |\hat{S}_0|^2) \geq \frac{\beta - \beta_0}{\beta} + o(1). \quad (2.27)$$

Since  $|\hat{S}_0| \leq L^d$ , for any  $0 < \epsilon < 1$  we have

$$E_{\mu_{L,\beta}}(L^{-2d} |\hat{S}_0|^2) \leq \epsilon + \mu_{L,\beta}(|\hat{S}_0| \geq \epsilon L^d) \quad (2.28)$$

and so

$$\mu_{L,\beta}(|\hat{S}_0| \geq \frac{1}{2} \frac{\beta - \beta_0}{\beta} L^d) \geq \frac{1}{2} \frac{\beta - \beta_0}{\beta} + o(1). \quad (2.29)$$

By the  $O(n)$  symmetry of  $\mu_{L,\beta}$ , the law of  $\hat{S}_0/L^d$  is rotationally invariant with non-degenerate “radius” distribution. Specializing to  $n = 2$  to get the explicit constants, we have

$$\mu_{L,\beta}(\theta \cdot \hat{S}_0 \geq \frac{1}{4} \frac{\beta - \beta_0}{\beta} L^d) \geq \frac{1}{6} \frac{\beta - \beta_0}{\beta} + o(1) \quad (2.30)$$

But this means that the exponent in the definition of  $f$  is at least  $\frac{1}{4} \frac{\beta - \beta_0}{\beta} L^d$  with uniformly positive probability and so

$$\frac{\partial f}{\partial h^+} \Big|_{h=0} \geq \frac{\beta - \beta_0}{4\beta} \geq - \frac{\partial f}{\partial h^-} \Big|_{h=0} \quad (2.31)$$

Applying Theorem 1.5, for  $\beta > \beta_0$  and any  $\theta \in \mathcal{S}^1$  there exists a translation invariant, ergodic Gibbs state  $\mu_\theta \in \mathfrak{G}_\beta$  such that

$$E_{\mu_\theta}(\theta \cdot S_x) = \frac{\partial f}{\partial h^+} \Big|_{h=0} > 0. \quad (2.32)$$

Next we need to show that the states  $\mu_\theta$  are actually distinct. The Ergodic Theorem implies

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} S_x \xrightarrow{L \rightarrow \infty} m_* \tilde{\theta}, \quad \mu_\theta\text{-a.s.} \quad (2.33)$$

where  $\tilde{\theta} \in \mathbb{S}^{n-1}$  and where  $m_*$  is a positive number such that, in light of (2.32),

$$m_* \theta \cdot \tilde{\theta} = \frac{\partial f}{\partial h^+} \Big|_{h=0}. \quad (2.34)$$

The distinctness of  $\mu_\theta$  will follow once we prove (2.25), i.e.,  $\theta = \tilde{\theta}$ . (This is intuitively obvious because by the way  $\mu_\theta$  is constructed, it should be tilted in the direction of  $\theta$ .)

Suppose  $\tilde{\theta} \neq \theta$ . Find a rotation  $A \in O(n)$  such that  $A\tilde{\theta} = \theta$ . Let  $\tilde{\mu} = \mu_\theta \circ A^{-1}$ —since both the Hamiltonian and the *a priori* measure are  $O(n)$ -invariant,  $\tilde{\mu} \in \mathfrak{G}_\beta$ . Since (2.33) implies  $E_{\mu_\theta} = m_\star \tilde{\theta}$ , we have

$$E_{\tilde{\mu}}(\theta \cdot S_x) = E_{\mu_\theta}(\theta \cdot AS_x) = m_\star |\theta|^2 \underset{\tilde{\theta} \neq \theta}{>} m_\star \theta \cdot \tilde{\theta} = \left. \frac{\partial f}{\partial h^+} \right|_{h=0}. \quad (2.35)$$

However,  $\tilde{\mu}$  is a Gibbs measure and so this contradicts the general bounds in Theorem 1.5. Hence, we must have  $\theta = \tilde{\theta}$  after all.  $\square$

## 2.4 Mermin-Wagner theorem

Our last item of business in this chapter is the complementary part of Theorem 2.2 on the *absence* of symmetry breaking in the recurrent cases. This argument predates the existence part by 20 years and bears the name of its discoverers:

**Theorem 2.6 [Mermin-Wagner theorem]** *Consider the  $O(n)$ -model,  $n \geq 1$ , with non-negative interactions constants  $(J_{x,y})$  satisfying the conditions (II,I2) from Sect. 2.1. Suppose the corresponding random walk is recurrent. Then every  $\mu \in \mathfrak{G}_\beta$  is invariant under simultaneous rotation of all spins.*

*Proof.* We will show that spins can be arbitrarily rotated at arbitrary small cost of the total energy. We will have to work with inhomogeneous rotations to achieve this, so let  $\varphi_x$  be a collection of numbers with  $\{x : \varphi_x \neq 0\}$  finite and let  $iR$  be a unit element of the Lie algebra  $\mathfrak{o}(n)$ , i.e.,  $e^{iR\alpha}$  is a rigid rotation of the unit sphere by angle  $\alpha$  about a particular axis. Let  $\omega_\varphi$  be the map on configuration space acting on individual spins via

$$\omega_\varphi(S_x) = e^{i\varphi_x R} S_x. \quad (2.36)$$

To investigate the effect of such transformation on the Hamiltonian, note that

$$\omega_\varphi(S_x) \cdot \omega_\varphi(S_y) = S_x \cdot e^{i(\varphi_y - \varphi_x)R} S_y = S_x \cdot S_y - S_x \cdot [1 - e^{i(\varphi_y - \varphi_x)R}] S_y. \quad (2.37)$$

Hence the energy of a configuration in any block  $\Lambda \supset \{x : \varphi_x \neq 0\}$  transforms as

$$H_\Lambda(\omega_\varphi(S)) = H_\Lambda(S) - \Delta H \quad (2.38)$$

where

$$\Delta H = \frac{1}{2} \sum_{x,y} J_{xy} S_x \cdot [1 - e^{i(\varphi_y - \varphi_x)R}] S_y. \quad (2.39)$$

Using that  $\Delta H$  depends only on the portion of the spin configuration in  $\Lambda$ , a simple application of the DLR condition shows that, for any local function  $f$ ,

$$E_\mu(f \circ \omega_\varphi) = E_\mu(f e^{\beta \Delta H}). \quad (2.40)$$

Suppose now that  $\varphi_x \rightarrow \alpha$ . Then the LHS tends to expectation of  $f$  with respect to the *uniformly* rotated measure,  $\mu \circ \omega_\alpha^{-1}$ . The theorem will thus follow if we can find a sequence of  $(\varphi_x)$  such that  $\varphi_x \rightarrow \alpha$  for all  $x$  but  $\Delta H \rightarrow 0$ .

To express the  $\varphi$ -dependence of  $\Delta H$ , we expand the exponential:

$$\begin{aligned} \Delta H &= -\frac{i}{2} \sum_{x,y} J_{xy} (S_x \cdot RS_y) (\varphi_y - \varphi_x) \\ &\quad + \frac{1}{4} \sum_{x,y} J_{xy} (RS_x \cdot RS_y) (\varphi_y - \varphi_x)^2 + \dots \end{aligned} \quad (2.41)$$

In the first term we note that the self-adjointness of  $R$ —valid by the choice of  $iR$  as an element of the Lie algebra—implies that  $J_{xy} (S_x \cdot RS_y)$  is symmetric under the exchange of  $x$  and  $y$ . Since  $(\varphi_y - \varphi_x)$  is antisymmetric and finitely supported, the sum is zero. Estimating the remainder by the quadratic term, we thus get

$$|\Delta H| \leq C \sum_{x,y} J_{xy} (\varphi_y - \varphi_x)^2 = 2C \mathcal{E}_{1-J}(\varphi, \varphi) \quad (2.42)$$

where we used that  $(RS_x \cdot RS_y)$  is bounded and recalled the definition of the Dirichlet form of the random walk driven by the  $(J_{x,y})$ 's.

Our task now boils down to minimizing the Dirichlet form subject to the condition that  $\varphi$  tends to one in every finite set. To that end we fix  $0 < R < \infty$  and let

$$\varphi(x) = \alpha \mathbb{P}_x(\tau_0 < \tau_{\Lambda_R^c}). \quad (2.43)$$

This function equals  $\alpha$  at  $x = 0$ , zero on  $\Lambda_R^c$  and is harmonic (wrt the generator of the random walk) in  $\Lambda_R \setminus \{0\}$ . A calculation shows

$$\mathcal{E}_{1-J}(\varphi, \varphi) = \sum_x \varphi(x) \sum_y J_{x,y} (\varphi_x - \varphi_y) \underset{\substack{\text{harmonic or} \\ \text{zero in } \{0\}^c}}{=} \alpha \sum_y J_{0,y} (\alpha - \varphi_y) \quad (2.44)$$

But recurrence implies that  $\varphi_y \rightarrow \alpha$  as  $R \rightarrow \infty$  for every  $y$  and since the  $J_{xy}$ 's are summable, the right-hand side tends to zero by the Dominated Convergence Theorem.  $\square$

## 2.5 Literature remarks

The content of the entire chapter is very classical. The Infrared Bound was discovered in the seminal work by Fröhlich, Simon and Spencer [37] from 1976 and used to prove a phase transition in the  $O(n)$ -model. The technique was immediately extended to other classical, and in some more limited cases, quantum models in papers of Fröhlich, Israel, Lieb and Simon [33, 34] and Dyson, Lieb and Simon [26].

The Mermin-Wagner theorem goes back to 1966 [54]. Various mathematical treatments followed [23, 36, 57]; the approach presented here is inspired by the exposition in Simon's book [64]. The probabilistic approach to this result, discovered by Dobrushin and Shlosman [23], has the advantage that no regularity conditions need to be posed on the spin-spin

interaction provided it takes the form  $V(S_x - S_y)$ ; cf the recent paper by Ioffe, Shlosman and Velenik [47]. Finally, we remark that a beautiful and more in-depth exposition of this material—including quantum systems—was presented at the Prague School in 1996 by Bálint Tóth; his handwritten lecture notes should be available online [67].

The connection with random walk is, of course, made possible by our choice to work with non-negative couplings. However, most of the quantitative conclusions of this chapter hold without reference to random walks.



## Chapter 3

# Infrared bound & mean-field theory

In the preceding lecture we recalled some classic results based on the infrared bound (IRB). The purpose of this lecture is to discuss how the IRB can be used to estimate how close is the so called mean-field approximation to the actual lattice model. A distinctive feature of this application is that while the control of spin-wave condensation is based primarily on the infrared—i.e., small- $k$  or large spatial scale—content of the IRB, here will make the predominant use of the finite- $k$ —i.e., short range—part of the IRB. (Notwithstanding, the finiteness of the integral (2.6) will still be required.)

### 3.1 Mean-field theory

Mean-field theory is a versatile approximation technique frequently used by physicists to analyze realistic physical models. We begin by a simple derivation that underscores the strengths, and the shortcomings, of this approach.

Consider the model with the usual Hamiltonian (2.1). Pick a translation invariant Gibbs measure  $\mu \in \mathfrak{G}_\beta$  and consider the expectation of the spin at the origin. The conditional definition of Gibbs measures (the DLR condition) allows us to compute this expectation by first conditioning on the spins outside the origin. The one-spin Gibbs measure is determined by the (one-spin) Hamiltonian

$$H_{\{0\}}(S) = - \sum_x J_{0,x} S_0 \cdot S_x = S_0 \cdot \sum_x J_{0,x} S_x \quad (3.1)$$

Denoting  $M_0 = \sum_x J_{0,x} S_x$ , we get

$$E_\mu(S_0) = E_\mu \left( \frac{E_{\mu_0}(S_0 e^{\beta S_0 \cdot M_0})}{E_{\mu_0}(e^{\beta S_0 \cdot M_0})} \right) \quad (3.2)$$

Here, abusing the notation slightly, the “inner” expectations are over  $S_0$ —while  $M_0$  acts as a constant here—and the outer expectation is over the spins in  $\mathbb{Z}^d \setminus \{0\}$ , and thus over  $M_0$ .

So far the derivation has been completely rigorous but now comes an *ad hoc* step: We suppose that the random variable  $M_0$  is strongly concentrated about its average so that we

can replace it by this average. Denoting

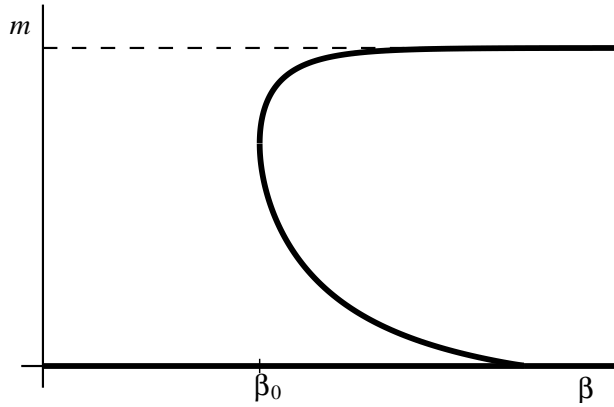
$$m = E_{\mu}(S_0) \quad (3.3)$$

we thus get that  $m$  should be an approximate solution to

$$m = \frac{E_{\mu_0}(S e^{\beta S \cdot m})}{E_{\mu_0}(e^{\beta S \cdot m})} \quad (3.4)$$

This is the so called *mean-field equation for the magnetization*.

Apart from the unjustified approximation, the problem with this equation is that it often has multiple solutions. For instance, for the set of points  $(\beta, m)$  that obey this equation, one typically gets a picture like this:



It is clear that as  $\beta$  varies, the physical solution must undergo some sort of jump, but it is not possible to tell where this jump occurs on the basis of equation (3.4) alone. For that one has to go beyond the heuristic derivation presented above. Here guidance will be provided by the fact that there is actually one system where the above derivation can be made rigorous: the same model on the complete graph. (In this case there are  $N$  vertices and  $J_{xy} = 1/N$  for all  $x$  and  $y$ .) The second key idea is to invoke the language of large-deviation theory.

Consider the cumulant generating function of the measure  $\mu_0$ ,

$$G(h) = \log E_{\mu_0}[e^{h \cdot S}], \quad h \in \mathbb{R}^v. \quad (3.5)$$

The Legendre transform,

$$\mathcal{S}(m) = \inf_{h \in \mathbb{R}^v} [G(h) - h \cdot m] \quad (3.6)$$

defines the *entropy* which, according to Cramér's theorem, is the rate of large-deviation decay in

$$\mu_0\left(\sum_{x=1}^N S_x \approx mN\right) = e^{-N\mathcal{S}(m)+o(N)}. \quad (3.7)$$

Next we inject the energy into the mix and look at the Gibbs measure. To describe what configurations dominate the partition function, and thus the Gibbs measure, we identify the decay rate of the probability

$$\mu_0\left(e^{\frac{\beta}{2N} \sum_{x,y=1}^N S_x \cdot S_y} \mathbf{1}_{\{\sum_x S_x \approx mN\}}\right) = e^{-N\Phi_{\beta}(m)+o(N)} \quad (3.8)$$

Here the rate function,

$$\Phi_\beta(m) = -\frac{\beta}{2}|m|^2 - \mathcal{S}(m), \quad (3.9)$$

is the so called *mean-field free-energy function*. The physical solutions are clearly obtained as the *absolute* minima of  $m \mapsto \Phi_\beta(m)$ . This is actually completely consistent with (3.4):

**Lemma 3.1** *We have*

$$\nabla \Phi_\beta(m) = 0 \quad \Leftrightarrow \quad m = \nabla G(\beta m) \quad (3.10)$$

*Explicitly, the solutions to (3.4) are in bijection with the extreme points of  $m \mapsto \Phi_\beta(m)$ .*

*Proof.* This is a simple exercise on the Legendre transform. First we note that  $\nabla \Phi_\beta(m) = 0$  is equivalent to  $\beta m = -\nabla \mathcal{S}(m)$ . The convexity of  $G$  implies that there is a unique  $h_m$  such that  $\mathcal{S}(m) = G(h_m) - m \cdot h_m$ . Furthermore,  $h_m$  depends smoothly on  $m$  and we have  $\nabla G(h_m) = m$ . It is easy to check that then  $\nabla \mathcal{S}(m) = -h_m$ . Putting this together with our previous observations, we get that  $\nabla \Phi_\beta(m) = 0 \Leftrightarrow \beta m = h_m \Leftrightarrow m = \nabla G(\beta m)$ . To get the second claim, note that  $m = \nabla G(\beta m)$  is a concise way to write (3.4).  $\square$

In short, the appearance of multiple solutions to (3.4) coincides with the emergence of secondary local maxima/minima. Let us check this on an explicit example:

*Mean field theory of the Potts model:* Recall the definition of the Potts model and our representation on the spin space  $\Omega = \{\hat{v}_1, \dots, \hat{v}_q\}$ . The mean-field free energy function is best expressed by means of the *mole fractions*,  $x_1, \dots, x_q$ , which on the complete graph represent the fraction of all vertices with spins pointing in the directions  $\hat{v}_1, \dots, \hat{v}_q$ , respectively. Clearly,  $\sum_{i=1}^q x_i = 1$  and the corresponding magnetization vector is

$$m = x_1 \hat{v}_1 + \dots + x_q \hat{v}_q. \quad (3.11)$$

In this notation we have

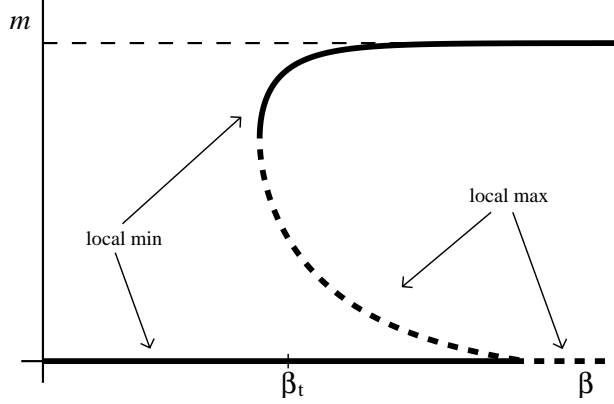
$$\Phi_\beta(m) = \sum_{k=1}^q \left( -\frac{\beta}{2} x_k^2 + x_k \log x_k \right). \quad (3.12)$$

As it turns out, all interesting behavior of  $\Phi_\beta$  occurs “on-axes” that is, in the directions of the spin states. The following picture shows the qualitative look of the function  $m \mapsto \Phi_\beta(m \hat{v}_1)$  at four increasing values of  $\beta$ :



Here the function first starts convex and, as  $\beta$  increases, develops a secondary local minimum (plus an inevitable local maximum). For  $\beta$  even larger, the secondary minimum

becomes degenerate with the one at  $m = 0$  and eventually takes over the role of the global minimum. With these new distinctions, the plot of solutions to the mean-field equation for the magnetization becomes:



Note that the local maximum eventually merges with the local minimum at zero—at which point zero becomes a local maximum. The jump in the position of the global minimum occurs at some  $\beta_t$ , which is strictly larger than the  $\beta_0$ , i.e., the point where the secondary minima/maxima first appear.

## 3.2 Approximation theorem & applications

The goal of this section is to show that, with the help of the IRB, the conclusions of mean-field theory can be given a quantitative form. Throughout we restrict ourselves to interactions of the form (2.1) and the coupling constants being one of the 3 types above.

**Definition 3.2** We say that a measure  $\mu \in \mathfrak{G}_\beta$  is a torus state if it is either a (weak) cluster point of measures  $\mu_{L,\beta}$  or can be obtained from such cluster points by perturbing either  $\beta$  or  $\mu_0$  or the inner product between spins.

The reason for the second half of this definition is that the “operations” thus specified preserve the validity of the IRB. For such states we prove:

**Theorem 3.3** Suppose  $|S_x| \leq 1$ . Let  $\mu \in \mathfrak{G}_\beta$  be a translation-invariant, ergodic, torus state and define

$$m_\star = E_\mu(S_0). \quad (3.13)$$

Let  $\Phi_\beta$  be the mean-field free energy function corresponding to this model. Then

$$\Phi_\beta(m_\star) \leq \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + \frac{v\beta}{2} \mathcal{I}_d \quad (3.14)$$

where

$$\mathcal{I}_d = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{\hat{J}(k)^2}{1 - \hat{J}(k)} \quad (3.15)$$

Note that the integral is finite iff the random walk corresponding to  $(J_{xy})$  is transient. However, unlike for Green's function,  $\mathcal{I}_d$  represents the expected number of returns back to the origin after the walk has left the origin. In strongly transient situation, one should expect that  $\mathcal{I}_d$  is fairly small. And, indeed, we have the following asymptotics:

- *n.n. interactions:*

$$\mathcal{I}_d \sim \frac{1}{2d}, \quad d \rightarrow \infty. \tag{3.16}$$

- *Yukawa potentials:* If  $d \geq 3$ ,

$$\mathcal{I}_d \leq C\mu^d. \tag{3.17}$$

- *Power-law potentials:* If  $d \geq 3$  OR  $s < \min\{d + 2, 2d\}$ ,

$$\mathcal{I}_d \leq C(s - d). \tag{3.18}$$

Of course, one is able to make the integral small for interactions with power law tails even when  $s$  is not too close to  $d$ : Just take a mixture of Yukawa and power-law with positive coefficients and let  $\mu$  be sufficiently small. Within the class of above models, we can rephrase Theorem 3.3 as:

Physical magnetizations nearly minimize the mean-field free energy function

This is justified because, as it turns out, all relevant magnetizations can be achieved in ergodic torus states. Let us again demonstrate the conclusion on the example of the  $q$ -state Potts model:

**Theorem 3.4** *Let  $q \geq 3$  and suppose that  $\mathcal{I}_d \ll 1/q$ . Then there exists  $\beta_t \in (0, \infty)$  and translation-invariant, ergodic measures  $\nu_0, \nu_1, \dots, \nu_q \in \mathfrak{G}_{\beta_t}$  such that*

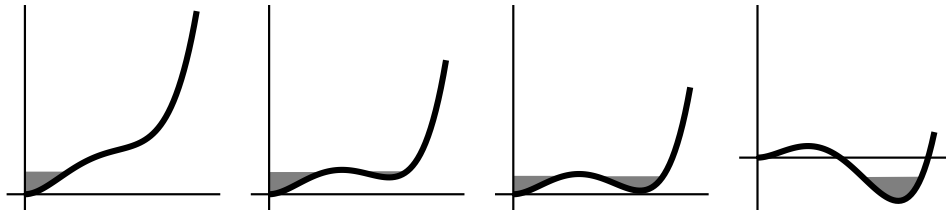
$$|E_{\nu_0}(S_x)| \ll 1 \tag{3.19}$$

and

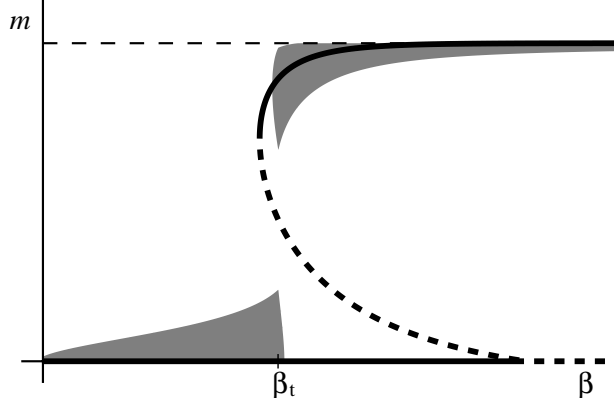
$$E_{\nu_j}(S_x) = m_* \hat{\nu}_j, \quad j = 1, \dots, q, \tag{3.20}$$

where  $m_* \geq 1/2$ . In particular, the 3-state Potts model undergoes a first-order phase transition provided the spatial dimension is sufficiently large.

This result is pretty much the consequence of the pictures in Sect. 3.1. Indeed, including the error bound (3.14), the physical magnetization is confined to the shaded regions:



Thus, once the error is smaller than the “hump” separating the two local minima, there is no way that the physical magnetization can change continuously as the temperature varies. This is seen even more dramatically if we depict the set of allowed values of the magnetization directly into the mean-field magnetization plot:



(To make the effect visible, the plots are done for the  $q = 10$  state Potts model rather than the most interesting case of  $q = 3$ .) Notice also that the transition is bound to occur rather sharply and very near the mean-field value of  $\beta_t$ ; we actually have explicit error bounds but there is no need to state them here.

### 3.3 Ideas from the proofs

The fundamental technical ingredient of the proof is again provided by the IRB. However, we will need the following enhanced version:

**Lemma 3.5 [IRB enhanced]** *Suppose the random walk driven by the  $(J_{xy})$  is transient and let  $G(x, y)$  denote the corresponding Green’s function on  $\mathbb{Z}^d$ . Let  $\mu \in \mathfrak{G}_\beta$  be a translation-invariant, ergodic, torus state and let  $m_\star = E_\mu(S_0)$ . Then for all  $(v_x)_{x \in \mathbb{Z}^d} \in \mathbb{C}^{\mathbb{Z}^d}$  with finite support,*

$$\sum_{x,y} v_x \bar{v}_y E_\mu((S_x - m_\star) \cdot (S_y - m_\star)) \leq \frac{\nu}{2\beta} \sum_{x,y} v_x \bar{v}_y G(x, y). \quad (3.21)$$

*Proof.* The IRB on torus survives weak limits and so we know that, for every  $(w_x)$  with finite support and  $\sum_x w_x = 0$ ,

$$\sum_{x,y} w_x \bar{w}_y E_\mu(S_x \cdot S_y) \leq \frac{\nu}{2\beta} \sum_{x,y} w_x \bar{w}_y G(x, y) \quad (3.22)$$

where

$$G(x, y) = \lim_{L \rightarrow \infty} G_L(x, y) = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{e^{ik \cdot (x-y)}}{1 - \hat{J}(k)} \quad (3.23)$$

What separates (3.22) from (3.21) are the  $m_\star$  terms in the expectation on the left and the absence of the restriction on the sum of  $v_x$ . The former is remedied easily; indeed, the restriction  $\sum_x w_x = 0$  allows us to put the  $m_\star$  terms at no additional cost.

To address the latter issue, suppose  $\Lambda_L$  contains the support of  $(v_x)$  and let

$$a_L = \frac{1}{|\Lambda_L|} \sum_x v_x \quad (3.24)$$

Define  $w_x = v_x - a_L 1_{\Lambda_L}(x)$ . Then

$$\begin{aligned} \sum_{x,y} w_x \bar{w}_y E_\mu((S_x - m_\star) \cdot (S_y - m_\star)) &= \sum_{x,y} v_x \bar{v}_y E_\mu((S_x - m_\star) \cdot (S_y - m_\star)) \\ &\quad - 2E_\mu\left(\left[ a_L \sum_{x \in \Lambda_L} (S_x - m_\star) \right] \cdot \left[ \sum_y v_y (S_y - m_\star) \right]\right) \\ &\quad + E_\mu\left(\left| a_L \sum_{x \in \Lambda_L} (S_x - m_\star) \right|^2\right) \end{aligned} \quad (3.25)$$

But ergodicity of  $\mu$  implies that

$$E_\mu\left(\left| \frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} (S_x - m_\star) \right|^2\right) \xrightarrow{L \rightarrow \infty} 0 \quad (3.26)$$

and so, by Cauchy-Schwarz, the last two terms in (3.25) tend to zero as  $L \rightarrow \infty$ . Similarly, a direct calculation (and the Riemann-Lebesgue lemma) shows that

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} G(x, y) \xrightarrow{L \rightarrow \infty} 0 \quad (3.27)$$

and so the corresponding terms on the right-hand side of (3.21) suffer a similar fate. This means that the left-hand sides of (3.21–3.22) tend to each other, and same for the right-hand sides. The desired bound (3.21) is thus a limiting version of (3.22).  $\square$

Clearly, the restriction to finitely-supported  $(v_x)$  is not necessary; instead, one can consider many reasonable completions as well. The above formulation has an immediate, but rather fundamental, consequence:

**Corollary 3.6 [Key estimate]** *Let  $\mu \in \mathfrak{G}_\beta$  be an ergodic torus state and let  $m_\star = E_\mu(S_x)$ . Then we have*

$$E_\mu\left(\left|\sum_x J_{0,x} S_x - m_\star\right|^2\right) \leq \frac{\nu}{2\beta} \mathcal{I}_d. \quad (3.28)$$

*Proof.* Choose  $v_x = J_{0,x}$  and note that with this choice the left-hand side of (3.21) becomes the left-hand side of (3.28). As to the right-hand side of (3.21), we get

$$\frac{\nu}{2\beta} \sum_{x,y} \int_{[-\pi,\pi]^d} \frac{dk}{(2\pi)^d} \frac{e^{ik \cdot (x-y)}}{1 - \hat{J}(k)} J_{0,x} J_{0,y} \quad (3.29)$$

Recalling the definition of  $\hat{J}(k)$ , this yields the desired error term.  $\square$

This corollary provides the justification of the *ad hoc* step in the derivation of mean-field theory: Indeed, once  $\mathcal{I}_d$  is small, the variance of  $M_0$  is small and so  $M_0$  is with high probability close to its average.

The rest of the proof is based on inequalities linking the mean-field free energy with the actual magnetization of the system; this part of the proof works for general non-negative coupling constants satisfying conditions (I1-I2) from Sect. 2.1. The relevant observations are as follows:

**Proposition 3.7** *Let  $\mu \in \mathfrak{G}_\beta$  be translation invariant and let  $m_\star = E_\mu(S_x)$ .*

(1) *We have*

$$\Phi_\beta(m_\star) \leq \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + \frac{\beta}{2} \sum_{x \in \mathbb{Z}^d} J_{0,x} [E_\mu(S_0 \cdot S_x) - |m_\star|^2] \quad (3.30)$$

(2) *Suppose also  $J_{0,x} \geq 0$  and  $|S_x| \leq 1$ . Then*

$$\sum_{x \in \mathbb{Z}^d} J_{0,x} [E_\mu(S_0 \cdot S_x) - |m_\star|^2] \leq \beta E_\mu \left( \left| \sum_x J_{0,x} S_x - m_\star \right|^2 \right) \quad (3.31)$$

*Proof of (1).* The proof is based on convexity inequalities linking the mean-field free energy and the characteristics of the actual system. Fix  $\Lambda \subset \mathbb{Z}^d$  and let  $Z_\Lambda$  be the partition function in  $\Lambda$ . A standard example of such convexity inequality is

$$Z_\Lambda \geq \exp \left\{ -|\Lambda| \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + O(\partial\Lambda) \right\} \quad (3.32)$$

whose proof is standard (and uses the same arguments as are to follow) and so we omit it.

Let  $M_\Lambda = \sum_{x \in \Lambda} S_x$  and fix  $h \in \mathbb{R}^v$ . First we note that the DLR condition implies

$$E_\mu(e^{+\beta H_\Lambda + h \cdot M_\Lambda} Z_\Lambda) = E_{\otimes \mu_0}[e^{h \cdot M_\Lambda}] = e^{|\Lambda|G(h)}. \quad (3.33)$$

The  $Z_\Lambda$  term can be bounded away via (3.32); Jensen's inequality then gives

$$\beta E_\mu(H_\Lambda) + |\Lambda| h \cdot m_\star - |\Lambda| \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + O(\partial\Lambda) \leq |\Lambda|G(h) \quad (3.34)$$

Next, translation invariance of  $\mu$  yields  $E_\mu(H_\Lambda) = -|\Lambda| \frac{1}{2} \sum_x J_{0,x} E_\mu(S_0 \cdot S_x) + O(\partial\Lambda)$  and so dividing by  $\Lambda$  and taking  $\Lambda \uparrow \mathbb{Z}^d$  along cubes gets us

$$-\frac{\beta}{2} \sum_x J_{0,x} E_\mu(S_0 \cdot S_x) - \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) \leq G(h) - h \cdot m_\star \quad (3.35)$$

Optimizing over  $h$  turns the right-hand side into  $\mathcal{S}(m_\star)$ . Adding  $\frac{1}{2}|m_\star|^2$  on both sides and invoking (3.9) now proves the claim.  $\square$

*Proof of (2).* The left-hand side can be written as  $E_\mu(S_0 \cdot M_0) - |m_\star|^2$ . Since  $J_{0,0} = 0$ , an application of the DLR condition yields

$$E_\mu(M_0 \cdot S_0) = E_\mu(M_0 \cdot \nabla G(\beta M_0)) \quad (3.36)$$

As  $E_\mu(M_0) = E_\mu[\nabla G(\beta M_0)] = m_\star$ , we have

$$E_\mu(S_0 \cdot M_0) - |m_\star|^2 = E_\mu \left( (M_0 - m_\star) \cdot (\nabla G(\beta M_0) - \nabla G(\beta m_\star)) \right). \quad (3.37)$$

But  $|S_x| \leq 1$  implies  $\nabla \nabla G(m) \leq \text{id}$  at any  $m \in \text{Conv}(\Omega)$ —assuming  $J_{x,y} \geq 0$ —and so

$$(M_0 - m_*) \cdot (\nabla G(\beta M_0) - \nabla G(\beta m_*)) \leq \beta |M_0 - m_*|^2 \quad (3.38)$$

by the Mean-Value Theorem. Taking expectations proves (3.31).  $\square$

Theorem 3.3 now follows by combining Proposition 3.7 with Corollary 3.6.

Interestingly, (3.30) gives

$$\sum_{x \in \mathbb{Z}^d} J_{0,x} E_\mu(S_0 \cdot S_x) \geq |m_*|^2 \quad (3.39)$$

i.e., the actual energy density always exceeds the mean-field energy density.

### 3.4 Literature remarks

Mean field theory dates back to Curie [20] and Weiss [69]. One of the early connections to the models on the complete graph appears in Ellis' textbook on large-deviation theory [28]. Most of this section is based on the papers of Biskup and Chayes [6] and Biskup, Chayes and Crawford [7]. The Key Estimate has been known for some while in specific cases; e.g., for the Ising model in the paper by Bricmont, Kesten, Lebowitz and Schonmann [16] and for the  $q$ -state Potts model in the paper by Kesten and Schonmann [49]. Both these works deal with the limit of the magnetization as  $d \rightarrow \infty$ ; notwithstanding, no conclusions were extracted for the presence of first-order phase transitions in finite-dimensional systems.

The first-order phase transition in the  $q$ -state Potts model has first been proved by Kotecký and Shlosman [49] but the technique works only for extremely large  $q$ . The case of small  $q$  has been open. The upshot of the present technique is that it replaces  $q$  by  $d$  or interaction range in its role of a “large parameter.” The price to pay is the lack of explicit control over symmetry: We expect that the measure  $\nu_0$  in Theorem 3.4 is actually “disordered” and  $E_{\nu_0}(S_x) = 0$ . This would follow if we knew that the magnetization in the Potts model can be discontinuous only at the percolation threshold—for the Ising model this was recently proved by Bodineau [15]—but this is so far known only in  $d = 2$  (or for  $q$  very large).

It is natural to ask whether one can say anything about the continuum- $q$  extension of the Potts model, the random cluster model; see [45]. Unfortunately, here it has been shown that the main condition for proving the IRB, reflection positivity, holds for random-cluster measure if and only if  $q$  is integer [5].

As already mentioned, the special cases of Yukawa interaction belongs to the class of Kac models. The 3-state Kac-Potts model has recently been studied by Gobron and Merola [44] who proved the existence of the first-order phase transition mentioned above using robust contour techniques. An advantage of contour technology over the RP is the explicit control of many quantities, including the number of coexisting phases. A drawback is the restriction to smeared-out interaction.

Another model for which the methods of this chapter yield novel results is the liquid-crystal model discussed in Sect. 1.2. Here Angelescu and Zagrebnoy [4] proved that symmetry breaking (for the order parameter  $\max_\alpha E_\mu[S_x^{(\alpha)}]^2 - 1/n$ ) occurs at low temperatures by exhibiting spin-wave condensation; cf Chapter 2. In [6] it has been shown that, for  $n \geq 3$ , the

order parameter undergoes a discontinuous transition at intermediate temperatures (cf also the papers of van Enter and Shlosman [29, 30] for proofs of such transitions in highly “non-linear” cases). Further “mean-field driven” first order phase transitions have been proved for the cubic model [6] and the Blume-Capel model [7].

The IRB has been connected to mean-field theory before; namely, in the work of Aizenman [1] (cf also Fröhlich [32] and Sokal [65]) in the context of lattice field theories and Aizenman and Fernández in the context of Ising systems in either high spatial dimensions [2] or for spread-out interactions [3]. A representative result from these papers is that the critical behavior in the Ising model is of mean-field nature above 4 dimensions. The IRB enters as a tool to derive a one-way bound on the critical exponents. Unfortunately, the full conclusions are restricted to interactions that are reflection positive; a non-trivial generalization was obtained recently by Sakai [59] who proved the IRB—and the corresponding mean-field conclusions—directly via a version of the lace expansion.

## Chapter 4

# Reflection positivity

In the last two chapters we have made extensive use of the infrared bound. Now is the time to prove it. This will require introducing the technique of reflection positivity which, somewhat undesirably, links long-range correlation properties of the spin models under consideration to the explicit structure of the underlying graph. Apart from the infrared bound, reflection positivity yields also the so called chessboard estimates which we will use extensively in Chapter 5.

### 4.1 Reflection positive measures

We begin by introducing the basic setup for the definition of reflection positivity: Consider the torus  $\mathbb{T}_L$  of side  $L$  with  $L$  even. The torus has a natural reflection symmetry along planes orthogonal to one of the lattice directions. (For that purpose we may think of  $\mathbb{T}_L$  as embedded into a continuum torus.) The corresponding “plane of reflection”  $P$  has two components, one at the “front” of the torus and the other at the “back.” The plane either passes through the sites of  $\mathbb{T}_L$  or bisects bonds; we speak of reflections *through sites* or *through bonds*, respectively. The plane splits the torus into two halves,  $\mathbb{T}_L^+$  and  $\mathbb{T}_L^-$ , which are disjoint for reflections through bonds and obey  $\mathbb{T}_L^+ \cap \mathbb{T}_L^- = P$  for reflections through sites.

Let  $\mathfrak{A}^\pm$  denote the set of all functions  $f : \Omega^{\mathbb{T}_L} \rightarrow \mathbb{R}$  that depend only on the spins in  $\mathbb{T}_L^\pm$ . Let  $\vartheta$  denote the reflection operator,  $\vartheta : \mathfrak{A}^\pm \rightarrow \mathfrak{A}^\mp$ , which acts on spins via  $\vartheta(S_x) = S_{\vartheta(x)}$ . Clearly,  $\vartheta$  is a morphism of algebra  $\mathfrak{A}^+$  onto  $\mathfrak{A}^-$  and  $\vartheta^2 = \text{id}$ .

**Definition 4.1 [Reflection positivity]** *A measure  $\mu$  on  $\Omega^{\mathbb{T}_L}$  is reflection positive (RP) with respect to  $\vartheta$  if*

(1) For all  $f, g \in \mathfrak{A}^+$ ,

$$E_\mu(f \vartheta g) = E_\mu(g \vartheta f) \quad (4.1)$$

(2) For all  $f \in \mathfrak{A}^+$ ,

$$E_\mu(f \vartheta f) \geq 0. \quad (4.2)$$

Note that the above implies that  $f, g \mapsto E_\mu(f \vartheta g)$  is a positive-semidefinite bilinear form. The first condition is usually automatic—it requires only  $\vartheta$ -invariance of  $\mu$ —so it is

the second condition that makes this concept non-trivial (hence also the name). As it turns out, one has always at least one RP measure:

**Lemma 4.2** *The product measure,  $\mu = \otimes \mu_0$ , is RP with respect to all reflections.*

*Proof.* First consider reflections through bonds. Let  $f, g \in \mathfrak{A}^+$ . Since  $\mathbb{T}_L^+ \cap \mathbb{T}_L^- = \emptyset$ , the random variables  $f$  and  $\vartheta g$  are independent under  $\mu$ . Hence,

$$E_\mu(f \vartheta g) = E_\mu(f)E_\mu(\vartheta g) = E_\mu(f)E_\mu(g) \quad (4.3)$$

whereby both conditions in Definition 4.1 follow.

For reflections through sites, we note that  $f$  and  $\vartheta g$  are independent conditional on  $S_P$ . Invoking the reflection symmetry of  $\mu(\cdot|S_P)$ , we get

$$E_\mu(f \vartheta g|S_P) = E_\mu(f|S_P)E_\mu(\vartheta g|S_P) = E_\mu(f|S_P)E_\mu(g|S_P). \quad (4.4)$$

Again the conditions of RP follow by inspection.  $\square$

A fundamental consequence of reflection positivity is the Cauchy-Schwarz inequality

$$[E_\mu(f \vartheta g)]^2 \leq E_\mu(f \vartheta f)E_\mu(g \vartheta g) \quad (4.5)$$

Here is an enhanced, but extremely useful, version of this inequality:

**Lemma 4.3** *Let  $\mu$  be RP with respect to  $\vartheta$  and let  $A, B, C_\alpha, D_\alpha \in \mathfrak{A}^+$ . Then*

$$[E_\mu(e^{A+\vartheta B+\sum_\alpha C_\alpha \vartheta D_\alpha})]^2 \leq [E_\mu(e^{A+\vartheta A+\sum_\alpha C_\alpha \vartheta C_\alpha})][E_\mu(e^{B+\vartheta B+\sum_\alpha D_\alpha \vartheta D_\alpha})] \quad (4.6)$$

*Proof.* Clearly, in the absence of the  $C_\alpha \vartheta D_\alpha$  terms, this simply reduces to (4.5). To include these terms we use expansion into Taylor series:

$$E_\mu(e^{A+\vartheta B+\sum_\alpha C_\alpha \vartheta D_\alpha}) = \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_\mu((e^A C_{\alpha_1} \dots C_{\alpha_n}) \vartheta (e^B D_{\alpha_1} \dots D_{\alpha_n})). \quad (4.7)$$

Now we apply (4.5) to the expectation on the right-hand side and then one more time to the resulting sum:

$$\begin{aligned} & E_\mu(e^{A+\vartheta B+\sum_\alpha C_\alpha \vartheta D_\alpha}) \\ & \leq \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} \left[ E_\mu((e^A C_{\alpha_1} \dots C_{\alpha_n}) \vartheta (e^A C_{\alpha_1} \dots C_{\alpha_n}))^{1/2} \right. \\ & \quad \left. \times E_\mu((e^B D_{\alpha_1} \dots D_{\alpha_n}) \vartheta (e^B D_{\alpha_1} \dots D_{\alpha_n}))^{1/2} \right] \\ & \leq \left( \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_\mu((e^A C_{\alpha_1} \dots C_{\alpha_n}) \vartheta (e^A C_{\alpha_1} \dots C_{\alpha_n})) \right)^{1/2} \\ & \quad \times \left( \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_\mu((e^B D_{\alpha_1} \dots D_{\alpha_n}) \vartheta (e^B D_{\alpha_1} \dots D_{\alpha_n})) \right)^{1/2} \end{aligned} \quad (4.8)$$

Resummation via (4.7) now yields the desired expression.  $\square$

Finally, let us also check that all 3 interactions that we focused our attention on in previous lectures are of the form in Lemma 4.3:

**Lemma 4.4** *For any plane  $P$ , the n.n. (ferromagnet) interaction, Yukawa potentials and the power-law decaying potentials, the torus Hamiltonian can be written as*

$$-H_L = A + \vartheta A + \sum_{\alpha} C_{\alpha} \vartheta C_{\alpha} \quad (4.9)$$

for some  $A, C_{\alpha} \in \mathfrak{A}^+$ .

*Proof.* We focus on reflections through bonds. Given  $P$ , the terms in the Hamiltonian can naturally be decomposed into three groups: those between the sites in  $\mathbb{T}_L^+$ , those between the sites in  $\mathbb{T}_L^-$  and those involving both halves of the torus:

$$-H_L = \underbrace{\frac{1}{2} \sum_{x,y \in \mathbb{T}_L^+} J_{xy}^{(L)} S_x \cdot S_y}_A + \underbrace{\frac{1}{2} \sum_{x,y \in \mathbb{T}_L^-} J_{xy}^{(L)} S_x \cdot S_y}_{\vartheta A} + \sum_{i=1}^d \underbrace{\sum_{\substack{x \in \mathbb{T}_L^+ \\ y \in \mathbb{T}_L^-}} J_{xy}^{(L)} S_x^{(i)} S_y^{(i)}}_{R_i} \quad (4.10)$$

where we used the reflection symmetry of the  $J_{xy}^{(L)}$  to absorb the  $1/2$  into the sum at the cost of confining  $x$  to  $\mathbb{T}_L^+$  and  $y$  to  $\mathbb{T}_L^-$ . The first two terms identify  $A$  and  $\vartheta A$ ; it remains to show that the  $R_i$ -term can be written as  $\sum_{\alpha} C_{\alpha} \vartheta C_{\alpha}$ . We proceed on case-by-case basis:

*n.n. interactions:* Here

$$R_i = \frac{1}{2d} \sum_{\substack{\langle x,y \rangle \\ x \in \mathbb{T}_L^+ \\ y \in \mathbb{T}_L^-}} S_x^{(i)} S_y^{(i)} \quad (4.11)$$

which is of the desired form since  $S_y = \vartheta(S_x)$ .

*Yukawa potentials:* We will only prove this in  $d = 1$ ; the higher dimensions are harder but similar. Note that if  $P$  passes through the origin and  $x \in \mathbb{T}_L^+$  and  $y \in \mathbb{T}_L^-$ ,

$$J_{xy}^{(L)} = C \sum_{n \geq 0} e^{-\mu(|x|+|y|+nL)} \quad (4.12)$$

Hence,

$$R_i = C \sum_{n \geq 0} e^{-\mu n L} \left( \sum_{x \in \mathbb{T}_L^+} e^{-\mu|x|} S_x^{(i)} \right) \left( \sum_{y \in \mathbb{T}_L^-} e^{-\mu|y|} S_y^{(i)} \right) \quad (4.13)$$

which is of the desired form.

*Power-law potentials:* Here we note

$$\frac{1}{|x-y|_1^s} = \int_0^{\infty} d\mu \mu^{s-1} e^{-\mu|x-y|_1} \quad (4.14)$$

which, after a moment's thought, reduces the problem to the Yukawa case.  $\square$

As a side remark we note that Lemma 4.4 gives us a sufficient condition for a Gibbs measure to be reflection positive:

**Lemma 4.5** *Fix a plane of reflection  $P$  and let  $\vartheta$  be the corresponding reflection operator. Fix  $\beta \geq 0$  and suppose that the torus Hamiltonian takes the form (4.9) with  $A, C_\alpha \in \mathfrak{A}^+$ . Then the torus Gibbs measure,  $\mu_{L,\beta}$ , is RP with respect to  $\vartheta$ .*

*Proof.* The proof is a simple modification of the argument in Lemma 4.5: Fix  $f, g \in \mathfrak{A}^+$ . Expansion of the exponential term in  $\sum_\alpha C_\alpha \vartheta C_\alpha$  yields

$$\begin{aligned} E_{\mu_{L,\beta}} &= \frac{1}{Z_L} E_{\otimes \mu_0} (f(\vartheta g) e^{\beta(A+\vartheta A+\sum_\alpha C_\alpha \vartheta C_\alpha)}) \\ &= \frac{1}{Z_L} \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_{\otimes \mu_0} ((f e^{\beta A} C_{\alpha_1} \cdots C_{\alpha_n}) \vartheta (g e^{\beta A} C_{\alpha_1} \cdots C_{\alpha_n})). \end{aligned} \quad (4.15)$$

The conditions of RP for  $\mu_{L,\beta}$  are now direct consequences of the fact that the product measure,  $\otimes \mu_0$ , is itself RP (cf Lemma 4.3).  $\square$

The previous proof actually yields the following fact: *if a torus measure  $\mu$  is RP, and a torus Hamiltonian  $H_L$  takes the form (4.9), then also the measure  $e^{-\beta H_L} d\mu$  is RP.* This seems like a useful tool for constructing RP measures; unfortunately, we do not know any natural measures other than product measures for which RP can be shown directly—that is, without invoking the proof of Lemma 4.3.

## 4.2 Gaussian domination

Now we are in a position to start proving the infrared bound. First we introduce its integral version known under the name Gaussian domination:

**Theorem 4.6 [Gaussian domination]** *Let  $(J_{xy})$  be one of the three interactions above. Fix  $\beta \geq 0$  and for  $h = (h_x)_{x \in \mathbb{T}_L} \in (\mathbb{R}^V)^{\mathbb{T}_L}$  define*

$$Z_L(h) = E_{\otimes \mu_0} \left( \exp \left\{ -\beta \sum_{x,y \in \mathbb{T}_L} J_{xy}^{(L)} |S_x - S_y + h_x - h_y|^2 \right\} \right). \quad (4.16)$$

Then

$$Z_L(h) \leq Z_L(0). \quad (4.17)$$

*Proof.* Let  $H_L$  denote the sum in the exponent. It is easy to check that  $H_L$  is of the form

$$-H_L = A + \vartheta B + \sum_\alpha C_\alpha \vartheta D_\alpha. \quad (4.18)$$

Indeed, for  $h \equiv 0$  this is simply Lemma 4.4 as the diagonal terms can always be absorbed into the *a priori* measure. To get  $h \neq 0$  we replace  $S_x$  by  $S_x + h_x$  at each  $x$ . This changes the meaning of the original terms  $A$  and  $C_\alpha$ —and makes them different on the two halves of the torus—but preserves the overall structure of the expression.

A fundamental ingredient is provided by Lemma 4.3 which yields

$$Z_L(h)^2 \leq Z_L(h_+) Z_L(h_-) \quad (4.19)$$

where  $h_+ = h$  on  $\mathbb{T}_L^+$  and  $h_+ = \vartheta h$  on  $\mathbb{T}_L^-$ , and similarly for  $h_-$ . Now let us show how this yields (4.17): Noting that  $Z_L(h) \rightarrow 0$  whenever any component of  $h$  tends to  $\pm\infty$ , the maximum of  $Z_L(h)$  is achieved at some finite  $h$ . Let  $h^*$  be a maximizer for which

$$N(h) = \#\{ \langle x, y \rangle : h_x \neq h_y \} \quad (4.20)$$

is minimal among all maximizers. We claim that  $N(h^*) = 0$ . Indeed, if  $N(h^*) > 0$  then there exists a plane of reflection  $P$  through bonds such that  $P$  intersects at least one bond  $\langle x, y \rangle$  with  $h_x^* \neq h_y^*$ . Observe that then

$$\min\{N(h_+^*), N(h_-^*)\} < N(h^*) \quad (4.21)$$

Suppose that  $N(h_+^*) < N(h^*)$ . Then the fact that  $h^*$  was a maximizer implies

$$Z_L(h^*)^2 \leq Z_L(h_+^*)Z_L(h_-^*) \leq Z_L(h_+^*)Z_L(h^*) \quad (4.22)$$

which means

$$Z_L(h^*) \leq Z_L(h_+^*), \quad (4.23)$$

i.e.,  $h_+^*$  is also a maximizer. But that contradicts the choice of  $h^*$  by which  $N(h^*)$  was already minimal possible. It follows that  $N(h^*) = 0$ , i.e.,  $h^*$  is a constant. Since  $Z(h+c) = Z(h)$  for any constant  $c$ , (4.17) follows.  $\square$

Now we can finally pay an old debt and prove the infrared bound:

*Proof of Theorem 2.3.* To keep the proof succinct, we will write  $\langle \eta, \zeta \rangle = \sum_{x \in \mathbb{T}_L} \eta_x \zeta_x$  for the natural inner product on  $L^2(\mathbb{T}_L)$ . Note that for any  $(\eta_x) \in (\mathbb{R}^\nu)^{\mathbb{T}_L}$ ,

$$\sum_{x, y \in \mathbb{T}_L} J_{xy}^{(L)} |\eta_x - \eta_y|^2 = \langle \eta, G_L^{-1} \eta \rangle \quad (4.24)$$

where  $G_L$  is as in (2.18). (Indeed, in Fourier components,  $\hat{G}_L^{-1}(k) = 1 - \hat{J}(k)$ .) As is easy to check,

$$\begin{aligned} Z_L(h) &= E_{\otimes \mu_0} \left( e^{-\beta \langle S+h, G_L^{-1}(S+h) \rangle} \right) \\ &= Z_L(0) E_{\mu_{L,\beta}} \left( e^{-2\beta \langle h, G_L^{-1} S \rangle - \beta \langle h, G_L^{-1} h \rangle} \right), \end{aligned} \quad (4.25)$$

where  $\mu_{L,\beta}$  is the torus Gibbs measure. The statement of Gaussian domination (4.17) is thus equivalent to

$$E_{\mu_{L,\beta}} \left( e^{-2\beta \langle h, G_L^{-1} S \rangle} \right) \leq e^{\beta \langle h, G_L^{-1} h \rangle} \quad (4.26)$$

We will now use invertibility of  $G_L$  to replace  $G_L^{-1} h$  by  $h$ . This yields

$$E_{\mu_{L,\beta}} \left( e^{-2\beta \langle h, S \rangle} \right) \leq e^{\beta \langle h, G_L h \rangle} \quad \text{whenever} \quad \sum_{x \in \mathbb{T}_L} h_x = 0, \quad (4.27)$$

where the latter condition comes from the fact that  $G_L^{-1}$  annihilates constant functions. Next we expand both sides to quadratic order in  $h$ :

$$1 - 2\beta E_{\mu_{L,\beta}} (\langle h, S \rangle) + \frac{4\beta^2}{2} E_{\mu_{L,\beta}} (\langle h, S \rangle^2) + \dots \leq 1 + \beta \langle h, G_L h \rangle + \dots \quad (4.28)$$

Since  $E_{\mu_{L,\beta}}(S)$  is constant,  $E_{\mu_{L,\beta}}(\langle h, S \rangle) = \langle h, E_{\mu_{L,\beta}}(S) \rangle = 0$  and we thus get

$$E_{\mu_{L,\beta}}(\langle h, S \rangle^2) \leq \frac{1}{2\beta} \langle h, G_L h \rangle \quad (4.29)$$

Finally, choose  $h_x = v_x \hat{e}_i$ , for some orthonormal basis vectors  $\hat{e}_i$  in  $\mathbb{R}^\nu$ . This singles out the  $i$ -th components of the spins on the left-hand side and has no noticeable effect on the right-hand side (beyond replacing vectors  $h_x$  by scalars  $v_x$ ). Summing the resulting inequality over  $i = 1, \dots, \nu$  we get the dot product of the spins on the left and an extra factor  $\nu$  on the right-hand side.  $\square$

### 4.3 Chessboard estimates

The proof of the infrared bound was based on the bound in Lemma 4.5 which boils down to the Cauchy-Schwarz inequality for the inner product  $f, g \mapsto E_\mu(f \vartheta g)$ . In this section we will systematize the use of this inequality to derive bounds on correlation functions. The key estimate—referred to as the *chessboard estimate*—will turn out to be useful in the proofs of phase coexistence in specific spin systems (even those to which the IRB technology does not apply).

Pick two integers,  $B < L$ , such that  $B$  divides  $L$  and  $L/B$  is even. Fixing the origin of the torus, let  $\Lambda_B$  the block corresponding to  $\{0, 1, \dots, B\}^d$ —i.e., the block of side  $B$  with lower-left corner at the origin. We may cover  $\mathbb{T}_L$  by translates of  $\Lambda_B$ ,

$$\mathbb{T}_L = \bigcup_{t \in \mathbb{T}_{L/B}} (\Lambda_B + Bt), \quad (4.30)$$

noting that the neighboring translates share the vertices on the adjacent sides. (This is a useful fact implied by our choice to work exclusively with reflections through planes of sites.) The translates are indexed by the sites in a “factor torus”  $\mathbb{T}_{L/B}$ .

**Definition 4.7** *A function  $f: \Omega^{\mathbb{T}_L} \rightarrow \mathbb{R}$  is called a  $B$ -block function if it depends only on  $\{S_x: x \in \Lambda_L\}$ . An event  $\mathcal{A} \subset \Omega^{\mathbb{T}_L}$  is called a  $B$ -block event if  $1_{\mathcal{A}}$  is a  $B$ -block function.*

Given a  $B$ -block function  $f$ , and  $t \in \mathbb{T}_{L/B}$ , we define  $\vartheta_t f$  be the reflection of  $f$  “into”  $\Lambda_B + Bt$ . More precisely, for a self-avoiding path on  $\mathbb{T}_{L/B}$  connecting  $\Lambda_B$  to  $\Lambda_B + Bt$ , we may sequentially reflect  $f$  along the planes between the successive blocks in the path. The result is a function that depends only on  $\{S_x: x \in \Lambda_B + Bt\}$ . Due to the commutativity of the reflections, this function does not depend on the choice of the path, so we denote it simply by  $\vartheta_t f$ . Note that since reflections are involutive,  $\vartheta^2 = \text{id}$ , there are only  $2d$  distinct functions one can obtain from  $f$  modulo translations.

**Theorem 4.8 [Chessboard estimate]** *Suppose  $\mu$  is RP with respect to all reflections between the neighboring blocks of the form  $\Lambda_B + Bt$ ,  $t \in \mathbb{T}_{L/B}$ . Then for any  $B$ -block functions  $f_1, \dots, f_m$ , and any distinct  $t_1, \dots, t_m \in \mathbb{T}_{L/B}$ ,*

$$E_\mu\left(\prod_{j=1}^m \vartheta_{t_j} f_j\right) \leq \prod_{j=1}^m \left[ E_\mu\left(\prod_{t \in \mathbb{T}_{L/B}} \vartheta_t f_j\right) \right]^{(B/L)^d} \quad (4.31)$$

Here is a version of this bound for events: If  $\mathcal{A}_1, \dots, \mathcal{A}_m$  are  $B$ -block events and  $t_1, \dots, t_m$  are *distinct* elements of  $\mathbb{T}_{L/B}$ , then

$$\mu\left(\bigcap_{j=1}^m \vartheta_{t_j}(\mathcal{A}_j)\right) \leq \prod_{j=1}^m \left[ \mu\left(\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_j)\right) \right]^{(B/L)^d} \quad (4.32)$$

where  $\vartheta_t(\mathcal{A}) = \{\vartheta_t \mathbf{1}_{\mathcal{A}} = 1\}$ . Note that the exponent  $(B/L)^d$  is the reciprocal volume of the torus  $\mathbb{T}_{L/B}$ . (This is consistent with the fact that both expressions transform homogeneously under the scaling  $f_j \rightarrow \lambda_j f_j$  with  $\lambda_j \geq 0$ .)

*Proof of Theorem 4.8.* We will assume throughout that  $E_\mu(f \vartheta f) = 0$  implies  $f = 0$ . (Otherwise, one has to factor out the ideal of such functions and work on the factor space.) We will first address the 1D case; the general dimensions will be handled by induction.

Abbreviate  $2n = L/B$  and fix a collection of non-zero functions  $f_1, \dots, f_{2n}$ . Define a multilinear functional  $F$  on the set of  $B$ -block functions by

$$F(f_1, \dots, f_{2n}) = E_\mu\left(\prod_{t=1}^{2n} \vartheta_t f_t\right). \quad (4.33)$$

Noting that  $F(f_j, \dots, f_j) > 0$ , we also define

$$G(f_1, \dots, f_{2n}) = \frac{F(f_1, \dots, f_{2n})}{\prod_{j=1}^{2n} F(f_j, \dots, f_j)^{\frac{1}{2n}}} \quad (4.34)$$

These objects enjoy a natural cyclic invariance,  $F(f_1, \dots, f_{2n}) = F(f_{2n}, f_1, \dots, f_{2n-1})$  and, similarly,  $G(f_1, \dots, f_{2n}) = G(f_{2n}, f_1, \dots, f_{2n-1})$ . The definition of  $G$  also implies

$$G(f, \dots, f) = 1. \quad (4.35)$$

Finally, Cauchy-Schwarz along the plane separating  $f_1$  from  $f_{2n}$  and  $f_n$  from  $f_{n+1}$  yields

$$G(f_1, \dots, f_{2n}) \leq G(f_1, \dots, f_n, f_n, \dots, f_1)^{1/2} \times G(f_{2n}, \dots, f_{n+1}, f_{n+1}, \dots, f_{2n})^{1/2}. \quad (4.36)$$

This will of course be the core estimate of the proof.

The desired claim will be proved if we show that

$$G(f_1, \dots, f_{2n}) \leq 1, \quad (4.37)$$

i.e., that  $G$  is maximized by  $2n$ -tuples composed of the same function. We will proceed similarly as in the proof of Gaussian Domination: Given a  $2n$ -tuple of  $B$ -block functions,  $(f_1, \dots, f_{2n})$ , let  $(g_1, \dots, g_{2n})$  be such that

- (1)  $g_i \in \{f_1, \dots, f_{2n}\}$  for each  $i = 1, \dots, 2n$
- (2)  $G(g_1, \dots, g_{2n})$  maximizes  $G$  over all such choices of  $g_1, \dots, g_{2n}$

- (3)  $g_1, \dots, g_{2n}$  is minimal in the sense that it contains the longest block (counted periodically) of the form  $f_i, f_i, \dots, f_i$ , for some  $i \in \{1, \dots, 2n\}$ .

Let  $k$  be the length of this block and, using the cyclic invariance, assume that the block occurs at the beginning of the sequence  $g_1, \dots, g_{2n}$ , i.e.,  $g_1, \dots, g_k = f_i$ .

We claim that  $k = 2n$ . Indeed, in the opposite case,  $k < 2n$ , we have  $g_{2n} \neq f_i$  and so (4.36) and the fact that  $(g_1, \dots, g_{2n})$  is a maximizer of  $G$  yield

$$\begin{aligned} G(g_1, \dots, g_{2n})^2 &\leq G(g_1, \dots, g_n, g_n, \dots, g_1)G(g_{2n}, \dots, g_{n+1}, g_{n+1}, \dots, g_{2n}). \\ &\leq G(g_1, \dots, g_n, g_n, \dots, g_1)G(g_1, \dots, g_{2n}) \end{aligned} \quad (4.38)$$

i.e.,

$$G(g_1, \dots, g_{2n}) \leq G(g_1, \dots, g_n, g_n, \dots, g_1) \quad (4.39)$$

This means that  $(g_1, \dots, g_n, g_n, \dots, g_1)$  is also a legitimate maximizer of  $G$  but it has a longer constant block—namely of length at least  $\min\{2k, 2n\}$ . Hence, we must have  $k = 2n$  as claimed. In light of (4.35–4.37), this proves the claim in  $d = 1$ .

To extend the proof to  $d > 1$ , suppose that  $m = (L/B)^d$  and assume, without loss of generality, that we have one function  $f_t$  for each block  $\Lambda_B + Bt$ . Writing

$$\prod_{t \in \mathbb{T}_{L/B}} \vartheta_t f_t = \prod_{j=1}^{2n} \left( \prod_{\substack{t \in \mathbb{T}_{L/B} \\ t_1=j}} \vartheta_t f_t \right) \quad (4.40)$$

we may apply the 1D chessboard estimate along the product over  $j$ . This homogenizes the product over  $f_t$  in the first coordinate direction. Proceeding through all directions we eventually obtain the desired claim.  $\square$

The chessboard estimate allows us to bound the probability of simultaneous occurrence of distinctly-placed  $B$ -block events in terms of their *disseminated versions*  $\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A})$ . The relevant quantities to estimate are thus

$$\mathfrak{z}(\mathcal{A}) = \mu \left( \bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}) \right)^{(B/L)^d} \quad (4.41)$$

Interestingly, as the next lemma shows, the map  $\mathcal{A} \mapsto \mathfrak{z}(\mathcal{A})$  is an outer measure:

**Lemma 4.9 [Subadditivity]** *Let  $\mathcal{A}$  and  $\mathcal{A}_1, \mathcal{A}_2, \dots$ , be  $B$ -block events such that*

$$\mathcal{A} \subset \bigcup_k \mathcal{A}_k \quad (4.42)$$

*Then*

$$\mathfrak{z}(\mathcal{A}) \leq \sum_k \mathfrak{z}(\mathcal{A}_k). \quad (4.43)$$

*Proof.* First we use the subadditivity of  $\mu$  and (4.42) to get

$$\mathfrak{z}(\mathcal{A})^{|\mathbb{T}_{L/B}|} = \mu \left( \bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}) \right) \stackrel{(4.42)}{\leq} \sum_{(k_t)} \mu \left( \bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_{k_t}) \right) \quad (4.44)$$

Next we apply the chessboard estimate

$$\mu\left(\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_{k_t})\right) \leq \prod_{t \in \mathbb{T}_{L/B}} \mathfrak{z}(\mathcal{A}_{k_t}) \quad (4.45)$$

to each term on the right hand side. Finally we apply the distributive law for sums and products with the result

$$\mathfrak{z}(\mathcal{A})^{|\mathbb{T}_{L/B}|} \leq \sum_{(k_t)} \prod_{t \in \mathbb{T}_{L/B}} \mathfrak{z}(\mathcal{A}_{k_t}) = \prod_{t \in \mathbb{T}_{L/B}} \sum_k \mathfrak{z}(\mathcal{A}_k) = \left(\sum_k \mathfrak{z}(\mathcal{A}_k)\right)^{|\mathbb{T}_{L/B}|} \quad (4.46)$$

Taking the  $|\mathbb{T}_{L/B}|$ -th root now yields the desired claim.  $\square$

The subadditivity property of  $\mathfrak{z}$  is extremely useful in applications: In order to estimate the  $\mathfrak{z}$ -value of an event, we may decompose it into smaller—and easier to compute—with—events, compute the  $\mathfrak{z}$ -value for each of them and then just add the results.

## 4.4 Literature remarks

The material of this section is entirely classic; a possible exception is Lemma 4.9 which seems to have been formulated in the present form only relatively recently [10]. The idea of reflection positivity goes to the days of constructive quantum field theory (namely, the Osterwalder-Schrader axioms [56]) where RP was a tool to obtain a sufficiently invariant—and natural—inner product. The use in statistical mechanics was initiated by the work of Fröhlich, Simon and Spencer [37] (infrared bound) and Fröhlich and Lieb [35] (chessboard estimates). The theory was further developed in two papers by Fröhlich, Israel, Lieb and Simon [33, 34]. There have been a couple of nice reviews of this material, e.g., by Shlosman [63] and in Georgii [43].

All use of reflection positivity in this note is restricted to one of the three interactions introduced in Chapter 2. Some generalizations beyond these are possible; e.g., the n.n. interaction of strength  $J$  may be accompanied by a n.n.n. interaction of strength  $\lambda$ —including negative values—and the result is still RP provided  $J \geq 2(d-1)|\lambda|$ . Some other examples are discussed, e.g., in [33, page 32]. Concerning arbitrary couplings, a sufficient conditions for RP is that, apart from translation and reflection invariance,

$$\sum_{\substack{x \in \mathbb{H} \\ y \in \mathbb{Z}^d \setminus \mathbb{H}}} J_{x,y} f(x) f(\vartheta y) \geq 0 \quad (4.47)$$

whenever  $f: \mathbb{H} \rightarrow \mathbb{R}$  is such that

$$\sum_{x \in \mathbb{H}} f(x) = 0 \quad (4.48)$$

Here  $\mathbb{H}$  is a halfspace in  $\mathbb{Z}^d$ —orthogonal to one of the coordinate direction—and  $\vartheta$  is the reflection of  $\mathbb{H}$  onto  $\mathbb{Z}^d \setminus \mathbb{H}$ . This is shown by proving equivalence of this property with the property of *conditional reflection negativity*, cf [33, page 14–17 and Corollary 3.6].

Finally, we wish to comment on RP with respect to *diagonal reflections*. For  $d$ -dimensional torus  $\mathbb{T}_L$  of side  $L = 2n$ , one considers planes through sites of the form

$$\begin{aligned} \{(x_1, \dots, x_d): x_i = x_j \text{ OR } x_i = x_j + n\}, \\ \{(x_1, \dots, x_d): x_i = -x_j \text{ OR } x_i = n - x_j\}, \end{aligned} \tag{4.49}$$

and shifts thereof. Since no edges of  $\mathbb{T}_L$  are bisected by such planes, nearest-neighbor interactions (of any sign) are RP with respect to the corresponding reflections. The whole theory then applies without significant changes. These observations go back to [63].

## Chapter 5

# Applications of chessboard estimates

In this chapter we will apply the technique of chessboard estimates to obtain proofs of phase coexistence in some lattice spin models. The arguments will be carried out in detail only for one rather simple example. For more sophisticated systems we present only the important ideas. Details, anyway, can be found in the corresponding papers.

### 5.1 Gaussian double-well model

Here we will demonstrate the use of chessboard estimates on the model of a Gaussian free-field model in a non-quadratic, double-well on-site potential. The Hamiltonian takes the general form

$$\beta H(\phi) = \beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 + \sum_x V(\phi_x) \quad (5.1)$$

where  $\phi_x \in \mathbb{R}$  with *a priori* measure given by the Lebesgue measure, and  $V$  is a potential. Note that  $\beta$  has been incorporated into the Hamiltonian in such a way that the on-site potential remains independent of it.

The most well known example of such systems is  $V(\phi) = \frac{\kappa}{2}\phi^2$  with  $\kappa > 0$  which is known as the *massive* Gaussian free field. This case can of course be treated completely explicitly; e.g., on the torus the corresponding Gaussian measure on  $(\phi_x)$  is zero mean with covariance

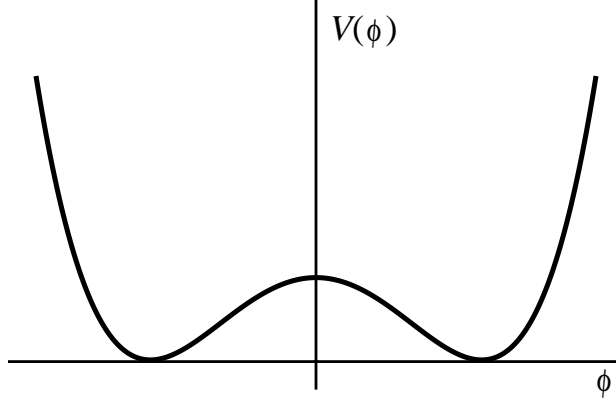
$$\text{Cov}(\phi_x, \phi_y) = \sum_{k \in \mathbb{T}_L^*} \frac{e^{ik \cdot (x-y)}}{\beta \widehat{D}(k) + \kappa} \quad (5.2)$$

where  $\widehat{D}(k)$  is the Fourier transform of the torus (discrete) Laplacian,

$$\widehat{D}(k) = \sum_{j=1}^d |1 - e^{ik_j}|^2 \quad (5.3)$$

Note that the inclusion of the mass,  $\kappa > 0$ —more precisely,  $\kappa$  is the mass squared—makes the covariance regular even for the zero mode  $k = 0$ .

We will look at a modification of this case when  $V$  takes the form



In fact, we will be even more specific and assume that  $V$  is simply given by

$$e^{-V(\phi)} = e^{-\frac{\kappa}{2}(\phi-1)^2} + e^{-\frac{\kappa}{2}(\phi+1)^2} \quad (5.4)$$

It is easy to check that, for  $\kappa$  sufficiently large,  $V$  defined using this formula looks as in the figure. The reason for assuming (5.4) is the possibility of an *Ising-spin representation*. Indeed, we may rewrite (5.4) as

$$e^{-V(\phi)} = \sum_{\sigma=\pm 1} e^{-\frac{\kappa}{2}(\phi-\sigma)^2} = C \sum_{\sigma=\pm 1} e^{-\frac{\kappa}{2}\phi_x - \kappa\phi_x\sigma_x} \quad (5.5)$$

where  $C = e^{-\kappa}$ . A product of such terms is thus proportional to

$$\prod_x e^{-V(\phi_x)} \propto \sum_{(\sigma_x)} \prod_x e^{-\frac{\kappa}{2}\phi_x^2 - \kappa\phi_x\sigma_x} \quad (5.6)$$

This means we can write the Gibbs weight of the model as follows

$$e^{-\beta \sum_{(x,y)} (\phi_x - \phi_y)^2 - \sum_x V(\phi_x)} \propto \sum_{(\sigma_x)} e^{-\beta \sum_{(x,y)} (\phi_x - \phi_y)^2 - \frac{\kappa}{2} \sum_x \phi_x^2} e^{-\kappa \sum_x \phi_x \sigma_x} \quad (5.7)$$

If we elevate  $(\sigma_x)$  to genuine degrees of freedom, we get a model on spins  $S_x = (\phi_x, \sigma_x)$  with *a priori* law Lebesgue on  $\mathbb{R} \times$  counting measure on  $\{-1, 1\}$  and the Hamiltonian

$$\beta H(\phi, \sigma) = \beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 + \frac{\kappa}{2} \sum_x \phi_x^2 + \kappa \sum_x \phi_x \sigma_x \quad (5.8)$$

Notice the first two terms on the right-hand side is the Hamiltonian of the massive (centered) Gaussian free field while the interaction between the  $\phi$ 's and the  $\sigma$ 's has on-site form.

Here are some observations whose (simple) proof we leave to the reader:

**Lemma 5.1** *Let  $\mu$  be a Gibbs measure for Hamiltonian (5.8) and let  $\nu$  be its  $\phi$ -marginal. Then  $\nu$  is a Gibbs measure for the Hamiltonian (5.1) subject to (5.4). The marginal  $\nu$  completely determines  $\mu$ : For any  $f$  depending only on  $\phi$  and  $\sigma$  in a finite set  $\Lambda$ ,*

$$E_\mu(f) = E_\nu \left( \sum_{(\sigma_x)_{x \in \Lambda}} f(\phi, \sigma) \prod_{x \in \Lambda} e^{V(\phi_x) - \frac{\kappa}{2}(\phi_x - \sigma_x)^2} \right) \quad (5.9)$$

We will use  $\mathfrak{G}_{\beta,\kappa}$  to denote the set of all Gibbs measures for the Hamiltonian (5.8) with parameters  $\beta$  and  $\kappa$ . The principal result for this model is as follows:

**Theorem 5.2** *Let  $d \geq 2$ . For all  $\epsilon > 0$  there is a  $a > 0$  such that for all  $\beta, \kappa > a$  there exist  $\mu^+, \mu^- \in \mathfrak{G}_{\beta,\kappa}$  which are translation invariant and obey*

$$\mu^\pm(\sigma_x = \pm 1) \geq 1 - \epsilon \quad (5.10)$$

and

$$|E_{\mu^\pm}(\phi_x) \mp 1| \leq \epsilon. \quad (5.11)$$

In simple terms, at low temperatures and large curvature of the wells of  $V$ , the fields prefer to localize in one of the wells. We remark that, while we chose the model as simple as possible, a similar conclusion would follow for with  $V$  given by

$$e^{-V(\phi)} = e^{-\frac{\kappa_+}{2}(\phi-1)^2+h} + e^{-\frac{\kappa_-}{2}(\phi+1)^2-h}. \quad (5.12)$$

where  $h$  changes the relative weight to the two minima. Indeed, there exists  $h_t$  at which one has two Gibbs measure—the analogues of  $\mu^+$  and  $\mu^-$ . Moreover, if  $\kappa_+ \gg \kappa_-$ , then  $h_t > 0$  because, roughly speaking, the well at  $-1$  offers more room for fluctuations.

## 5.2 Proof of phase coexistence

Here we will prove Theorem 5.2. We will focus on  $d = 2$ ; the proof in general dimension is a straightforward, albeit more involved, generalization.

Let us refer to a face of  $\mathbb{Z}^2$  as a *plaquette* (i.e., a plaquette is a square of side one with a vertex of  $\mathbb{Z}^2$  in each corner). Given a spin configuration  $(\sigma_x)$ , we say that a plaquette is *good* if all four spins take the same value, and *bad* otherwise. Let  $\mathcal{B}$  denote the event that the plaquette with lower-left corner at the origin is bad.

Since the interaction is that of the GFF with a modified single-spin measure, the torus Gibbs measure is RP. The crux of the proof is to show that bad plaquettes are suppressed. Specifically, we want to show that

$$\mathfrak{z}(\mathcal{B}) \ll 1 \quad \text{once} \quad \beta, \kappa \gg 1 \quad (5.13)$$

Appealing to the subadditivity lemma (Lemma 4.9) we only need to estimate the  $\mathfrak{z}$ -value of all possible configurations on the plaquette that constitute  $\mathcal{B}$ . Due to the plus-minus symmetry of the  $\sigma$ 's, it suffices to examine three patterns:

$$\begin{array}{ccc} - & + & + & - \\ + & - & + & - \\ + & + & + & + \end{array} \quad (5.14)$$

We begin with the most interesting of the three:

**Lemma 5.3** *For any  $\beta, \kappa > 0$ ,*

$$\mathfrak{z}\left(\begin{array}{cc} - & + \\ + & - \end{array}\right) \leq e^{-\frac{4\beta\kappa}{8\beta+\kappa}} \quad (5.15)$$

*Proof.* Let  $Z_L = \sum_{\sigma} \int e^{-\beta H_L(\phi, \sigma)} \prod_{x \in \mathbb{T}_L} d\phi_x$  be the torus partition function. Given a plaquette spin pattern, let  $Z_L(\text{pattern})$  denote the same object with  $\sigma$  fixed to the disseminated pattern—the sole element of  $\bigcap_{t \in \mathbb{T}_L} \vartheta_t(\text{pattern})$ . (We are working with  $B = 1$ .) By the definition of  $\mathfrak{z}$  we have

$$\mathfrak{z}_{(+\pm)}^{(-+)^{|\mathbb{T}_L|}} = \frac{Z_L(-+)}{Z_L(+\pm)} \leq \frac{Z_L(-+)}{Z_L(+\pm)}. \quad (5.16)$$

Now the partition function with all  $\sigma$ 's restricted to  $+$  is given by

$$\begin{aligned} Z_L(+\pm) &= \int e^{-\beta \sum_{(x,y)} (\phi_x - \phi_y)^2 - \frac{\kappa}{2} \sum_x \phi_x^2} e^{-\kappa \sum_x \phi_x} \prod_{x \in \mathbb{T}_L} d\phi_x \\ &= (\dots) E_{\text{GFF}}(e^{-\kappa \sum_x \phi_x}) \end{aligned} \quad (5.17)$$

where the expectation is with respect to the massive Gaussian free field and the prefactor denotes the integral of the Gaussian kernel over all  $\phi_x$ . Similarly we obtain

$$Z_L(-+) = (\dots) E_{\text{GFF}}(e^{-\kappa \sum_x \phi_x (-1)^{|x|}}) \quad (5.18)$$

where we noticed that by disseminating the pattern  $\bar{+}$  we obtain a configuration which is one at even parity  $x$  and minus on at odd parity  $x$ . Thus we conclude

$$\mathfrak{z}_{(+\pm)}^{(-+)^{|\mathbb{T}_L|}} \leq \frac{E_{\text{GFF}}(e^{-\kappa \sum_x \phi_x})}{E_{\text{GFF}}(e^{-\kappa \sum_x \phi_x (-1)^{|x|}})} \quad (5.19)$$

i.e., we only need to compute the ratio of the Gaussian expectations, and not the prefactors. Next we recall a standard formula for Gaussian moment generating functions: If  $X$  is a multivariate Gaussian, then

$$E(e^{\lambda \cdot X}) = e^{\lambda \cdot EX + \frac{1}{2} \text{Var}(\lambda \cdot X)} \quad (5.20)$$

Since  $E_{\text{GFF}}(\phi_x) = 0$ , we only need to compute the diagonal matrix element of  $\text{Cov}(\phi_x, \phi_y)$  for vectors  $\mathbf{1} = (1, 1, \dots)$  and  $(-1)^{|x|}$ . However, a quick look at (5.2) will convince us that these functions are eigenvectors of the covariance matrix corresponding to  $k = 0$  and  $k = (\pi, \pi)$ , respectively. Since  $\widehat{D}(0) = 0$  while  $\widehat{D}(\pi, \pi) = 8$ , we get

$$\text{Var}_{\text{GFF}}\left(\sum_x \phi_x\right) = \frac{|\mathbb{T}_L|}{\kappa} \quad (5.21)$$

$$\text{Var}_{\text{GFF}}\left(\sum_x \phi_x (-1)^{|x|}\right) = \frac{|\mathbb{T}_L|}{8\beta + \kappa} \quad (5.22)$$

where the factor  $|\mathbb{T}_L|$  is the (square of) the  $L^2(\mathbb{T}_L)$ -norm of the functions under consideration. Plugging this in (5.19) we conclude

$$\mathfrak{z}_{(+\pm)}^{(-+)^{|\mathbb{T}_L|}} \leq \exp\left\{\frac{1}{2} |\mathbb{T}_L| \kappa^2 \left(\frac{1}{8\beta + \kappa} - \frac{1}{\kappa}\right)\right\} \quad (5.23)$$

from which the claim readily follows.  $\square$

Next we attend to the other patterns:

**Lemma 5.4** For any  $\beta, \kappa > 0$ ,

$$\mathfrak{z}_{(+\bar{-})} \leq e^{-\frac{2\beta\kappa}{4\beta+\kappa}} \quad (5.24)$$

and

$$\mathfrak{z}_{(+\bar{+})} \leq e^{-\frac{2\beta\kappa}{8\beta+\kappa}} \quad (5.25)$$

*Proof.* As for (5.24), dissemination of  $+\bar{-}$  leads to alternating stripes of pluses and minuses, i.e.,  $\sigma_x = (-1)^{|x_1|}$ . Again, this is an eigenvector of the covariance matrix (5.2) with  $k = (\pi, 0)$ . The corresponding  $\widehat{D}$  equals 4 and so

$$\mathfrak{z}_{(+\bar{-})}^{|\mathbb{T}_L|} \leq \exp\left\{\frac{1}{2}|\mathbb{T}_L|\kappa^2\left(\frac{1}{4\beta+\kappa} - \frac{1}{\kappa}\right)\right\} \quad (5.26)$$

yielding (5.24).

The pattern  $\bar{+}$  is more complex because its dissemination will not lead to an eigenvector of the covariance matrix. However, we circumvent this problem by proving that

$$\mathfrak{z}_{(\bar{+}\bar{+})} \leq \mathfrak{z}_{(\bar{+}\bar{-})}^{1/2} \quad (5.27)$$

Indeed, since the torus comes from a square, and the interaction is nearest-neighbor, we may also use the diagonal reflections; cf the remarks in Sect. 4.4. Thus reflecting along the diagonal containing two +'s, the pattern  $\bar{+}\bar{+}$  yields  $\bar{+}\bar{-}$  and  $\bar{+}\bar{+}$ . This implies

$$\mathfrak{z}_{(\bar{+}\bar{+})} \leq \mathfrak{z}_{(\bar{+}\bar{-})}^{1/2} \mathfrak{z}_{(\bar{+}\bar{+})}^{1/2} \quad (5.28)$$

But  $\mathfrak{z}_{(\bar{+}\bar{+})} \leq 1$  and so (5.27) follows. By Lemma 5.3, (5.27) implies (5.25).  $\square$

**Corollary 5.5** For each  $\epsilon > 0$  there exists  $a > 0$  such that if  $\beta, \kappa > a$ , then  $\mathfrak{z}(\mathcal{B}) \leq \epsilon$ .

*Proof.* The event  $\mathcal{B}$  can be written as the union over a finite number of bad patterns. On the basis of Lemmas 5.3–5.4 the claim holds for  $\mathcal{B}$  replaced by any fixed bad pattern. The desired bound now follows—with slightly worse constants—by invoking Lemma 4.9.  $\square$

Next we explain our focus on the bad event:

**Lemma 5.6** There exists a constant  $c \in (1, \infty)$  such that if  $c\mathfrak{z}(\mathcal{B}) < 1/2$  then for any  $x, y \in \mathbb{T}_L$ ,

$$\mu_L(\sigma_x = 1, \sigma_y = -1) \leq 2c\mathfrak{z}(\mathcal{B}). \quad (5.29)$$

*Proof.* This is a consequence of a simple Peierls' estimate. Indeed, if  $\sigma_x = 1$  and  $\sigma_y = -1$ , then  $x$  is separated from  $y$  by a “circuit” of bad plaquettes. (Formally, either all plaquettes containing  $x$  are bad or there exists a non-trivial connected component of good—i.e., *not* bad—plaquettes containing  $x$ . This component cannot cover the whole torus because  $\sigma_y = -1$ ; the above “circuit” is then comprised of the bad plaquettes on the boundary of this component.) This means that

$$\mu_L(\sigma_x = 1, \sigma_y = -1) \leq \sum_{\gamma} \mu_L\left(\bigcap_{t \in \gamma} \vartheta_t(\mathcal{B})\right) \leq \sum_{\gamma} \mathfrak{z}(\mathcal{B})^{|\gamma|} \quad (5.30)$$

where  $|\gamma|$  denotes the maximal number of disjoint bad plaquettes in  $\gamma$  and where we used the chessboard estimates to derive the second bound. By standard arguments, the number of circuits of “length”  $n$  surrounding  $x$  or winding around  $\mathbb{T}_L$  at least once is bounded by  $c^n$ , for some constant  $c > 1$ . It follows that

$$\mu_L(\sigma_x = 1, \sigma_y = -1) \leq \sum_{n \geq 1} c^n \mathfrak{z}(\mathcal{B})^n \quad (5.31)$$

Under the condition  $c\mathfrak{z}(\mathcal{B}) < 1/2$  this sum is less than twice its first term.  $\square$

Finally, we can assemble the ingredients into the proof of phase coexistence:

*Proof of Theorem 5.2.* By symmetry of the torus measure, we have

$$\mu_L(\sigma_x = 1) = 1/2 = \mu_L(\sigma_x = -1). \quad (5.32)$$

Let  $z$  be a site at the back of the torus—that is distant at least  $L/2$  from the origin—and define

$$\mu_L^\pm(-) = \mu_L(-|\sigma_z = \pm 1). \quad (5.33)$$

These measures satisfy the DLR condition with respect to any function that depends only on the “front” of the torus and so any weak cluster point of these measures will be an infinite-volume Gibbs measure. Extract such measures by subsequential limits and call them  $\mu^+$  and  $\mu^-$ , respectively.

We claim that  $\mu^+ \neq \mu^-$ . Indeed, by Lemma 5.6 we have

$$\mu_L^+(\sigma_x = -1) \leq 2c\mathfrak{z}(\mathcal{B}) \quad (5.34)$$

once  $\mathfrak{z}(\mathcal{B}) \ll 1$  and, by Corollary 5.5, this actually happens once  $\beta, \kappa \gg 1$ . Thus if, say,  $2c\mathfrak{z}(\mathcal{B}) \leq 1/4$ , then  $\mu_L^+(\sigma_x = -1) \leq 1/4$  and, at the same time,  $\mu_L^-(\sigma_x = +1) \leq 1/4$ . The same holds for the limiting objects and so  $\mu^+ \neq \mu^-$ . Note that the measures can be averaged over shifts so that they become translation invariant.  $\square$

In the last step of the proof we used, rather conveniently, the plus-minus symmetry of the torus measure. In the asymmetric cases, e.g., (5.12), one can either invoke a continuity argument—choose  $h = h_L$  such that (5.32) holds—or turn (5.29) into the proof that  $|\mathbb{T}_L|^{-1} \sum_{x \in \mathbb{T}_L} \sigma_x$  will take values in  $[-1, -1 + \epsilon] \cup [1 - \epsilon, 1]$  with probability tending to one as  $L \rightarrow \infty$ . The latter “forbidden-gap” argument is rather robust and extends, with appropriate modifications, to all shift-ergodic infinite-volume Gibbs measures. Hence, the empirical magnetization in ergodic measures cannot change continuously with  $h$ .

### 5.3 Gradient fields with non-convex potential

Having demonstrated the use of chessboard estimates on a toy model, we will proceed to discuss more complicated systems. We begin with an example which is somewhat similar to the Gaussian double-well model.

A natural generalization of the massless GFF is obtained by replacing the quadratic gradient interaction by a general, even function of the gradients. The relevant Hamiltonian (again

with temperature incorporated in it) is

$$\beta H(\phi) = \sum_{\langle x, y \rangle} V(\phi_x - \phi_y) \quad (5.35)$$

The requirements that we generally put on  $V$  are continuity, evenness and quadratic decay at infinity. Under these conditions one can always define finite-volume Gibbs measures.

As to the measures in infinite volume, the massless nature of the model may prevent existence of a meaningful thermodynamic limit in low dimensions; however, if one restricts attention to *gradient variables*,

$$\eta_b = \phi_y - \phi_x \quad \text{if } b \text{ is the oriented edge } (x, y), \quad (5.36)$$

then the infinite-volume Gibbs measures exist, and may be characterized by a DLR condition, in all  $d \geq 1$ . We call these *gradient Gibbs measures* (GGM). A non-trivial feature of the GGM is that they obey a host of constraints. Indeed, almost every  $\eta$  is such that

$$\eta_{b_1} + \eta_{b_2} + \eta_{b_3} + \eta_{b_4} = 0 \quad (5.37)$$

for any plaquette  $(b_1, \dots, b_4)$  with bonds listed (and oriented) counterclockwise.

Surprisingly, the classification of all possible translation-invariant, infinite-volume GGMs can be achieved under the condition that  $V$  is strictly convex:

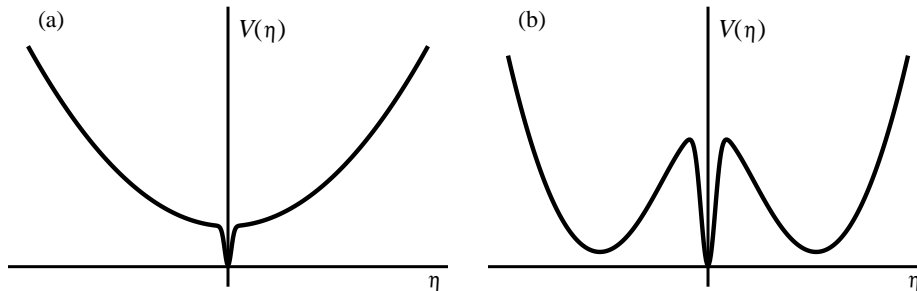
**Theorem 5.7** *Suppose  $V$  is convex, twice continuously differentiable with  $V''$  bounded away from zero and infinity. Then the shift-ergodic GGMs  $\mu$  are in one-to-one correspondence with their tilt, which is a vector  $a \in \mathbb{R}^d$  such that*

$$E_\mu(\eta_b) = a \cdot b \quad (5.38)$$

for every (oriented) bond  $b$  (we regard  $b$  as a unit vector for this purpose).

The word *tilt* comes from the interpretation of  $a$  as the slope or the incline of the interface whose height-gradient along bond  $b$  is given by  $\eta_b$ . The proof of this result—which is due to Funaki and Spohn—is based on the use of Brascamp-Lieb inequality through which the convexity assumption enters in an essential way.

A natural question to ask is what happens when  $V$  is *not* convex. Specific examples of interest might be  $V$  taking the form of a double-well potential—kind of like for the Gaussian double-well model—or  $V$ 's as in the figure:



As it turns out, the double-well case is not quite tractable at the moment—and most likely behaves like a massless GFF on large scales—but the other two cases are within reach. We will focus on the case (a) and, as for the Gaussian double-well model, assume a particular form of the potential:

$$e^{-V(\eta)} = p e^{-\kappa_0 \eta^2/2} + (1-p) e^{-\kappa_D \eta^2/2} \quad (5.39)$$

where  $\kappa_0$  and  $\kappa_D$  are positive numbers and  $p \in [0, 1]$  is a parameter to be varied. For this system one can prove the following result:

**Theorem 5.8** *Suppose  $d = 2$  and  $\kappa_0 \gg \kappa_D$ . Then there is  $p_t \in (0, 1)$  and, for  $V$  with  $p = p_t$ , there are two distinct, infinite-volume, shift-ergodic GGMs  $\mu_{\text{ord}}$  and  $\mu_{\text{dis}}$  that are invariant with respect to lattice rotations and have the following properties:*

(1) *zero tilt:*

$$\frac{1}{|\Lambda_L|} \sum_{\substack{b=(x,y) \\ x,y \in \Lambda_L}} \eta_b \xrightarrow{L \rightarrow \infty} 0, \quad \mu_{\text{ord}}, \mu_{\text{dis}}\text{-a.s.} \quad (5.40)$$

(2) *distinct fluctuation size:*

$$E_{\mu_{\text{ord}}}(\eta_b^2) \ll E_{\mu_{\text{dis}}}(\eta_b^2) \quad (5.41)$$

(3) *GFF as the scaling limit: For any smooth  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  with compact support,*

$$\sum_{b=(x,y)} \eta_b [f(\epsilon y) - f(\epsilon x)] \xrightarrow[\epsilon \downarrow 0]{\mathcal{D}} \mathcal{N}(0, C \|\nabla f\|_2^2) \quad (5.42)$$

where  $\mathcal{N}(0, \sigma^2)$  denotes a normal mean-zero variance- $\sigma^2$  random variable and  $C$  is a positive constant.

The upshot of this result is that, once the convexity of  $V$  is strongly violated, the conclusions of Theorem 5.7 do not apply. While the example is restricted to  $d = 2$ , and to potentials of the form (5.39), generalizations to  $d \geq 2$  and other potentials as in the above figure are possible and fairly straightforward.

Here are the main steps of the proof. First, as for the Gaussian double-well model, we use (5.39) to expand the Gibbs weight according to whether the first or the second term in (5.39) applies. This gives rise to a configuration of coupling strengths ( $\kappa_b$ ), one for each bond  $b$ , which take values in  $\{\kappa_0, \kappa_D\}$ . The joint Hamiltonian of the  $\eta$ 's and the  $\kappa$ 's is

$$\beta H(\eta, \kappa) = \sum_b \frac{\kappa_b}{2} \eta_b^2 \quad (5.43)$$

The joint measure is RP with respect to reflections through bonds and sites and, given ( $\kappa_b$ ), the  $\eta$ 's are Gaussian.

Next we focus on lattice plaquettes and divide these into good and bad according to whether the  $\kappa$ 's are all the same or not, respectively. The dissemination of bad patterns again leads to Gaussian integrals, but this time for GFF with inhomogeneous—yet periodically varying—couplings. The periodic nature allows the use of Fourier modes to diagonalize the requisite

covariance matrices. Performing the integrals we then check that all patterns designated as bad are suppressed (this is where  $\kappa_0 \gg \kappa_D$  is needed). The rest of the proof is pretty much as for the double-well model.

Full details of the proof of (1-2) are to be found in the paper by Kotecký and the present author; the proof of (3) is a consequence of some basic arguments in homogenization theory.

## 5.4 Spin-waves vs infinite ground-state degeneracy

Next we will discuss a couple of spin models whose distinctive feature is a high degeneracy of their ground state which is removed, at positive temperature, by soft-mode spin-wave fluctuations. The simplest example is as follows:

Orbital compass model: Here  $S_x \in \mathbb{S}^{d-1}$  with  $x \in \mathbb{Z}^d$ . The Hamiltonian is

$$H(S) = \sum_x \sum_{\alpha=1}^d (S_x^{(\alpha)} - S_{x+\hat{e}_\alpha}^{(\alpha)})^2 \quad (5.44)$$

where  $S_x^{(\alpha)}$  denotes the  $\alpha$ -th Cartesian component of the spin and  $\hat{e}_\alpha$  is the unit vector in the  $\alpha$ -th coordinate direction.

Note that *every* constant configuration is a minimum-energy state of (5.44). Other ground states may be obtained from the constant ones by picking a coordinate direction  $\alpha$  and changing the sign of the  $\alpha$ -th component of all spins in some of the “lines” parallel with  $\hat{e}_\alpha$ . In  $d = 2$  these are all ground states but in  $d \geq 3$  other complicated operations are possible that preserve the minimum-energy property.

Here is a theorem one can prove about the 2D system:

**Theorem 5.9** *For each  $\epsilon > 0$  there exist  $\beta_0 > 0$  and, for each  $\beta \geq \beta_0$ , there exist two distinct, shift-ergodic Gibbs measures  $\mu_1, \mu_2 \in \mathfrak{G}_\beta$  such that*

$$E_{\mu_j}(|S_x \cdot \hat{e}_j|) \geq 1 - \epsilon, \quad j = 1, 2 \quad (5.45)$$

*Moreover, for any  $\mu \in \mathfrak{G}_\beta$  we have*

$$E_\mu(S_x) = 0 \quad (5.46)$$

*and there are no shift-ergodic  $\mu \in \mathfrak{G}_\beta$ ,  $\beta \geq \beta_0$ , for which  $\max_{j=1,2} E_\mu(|S_x \cdot \hat{e}_j|) < 1 - \epsilon$ .*

The main idea underlying the proof is the evaluation of the free energy associated with spin-wave perturbations of the constant ground states; it is expected that only the states with the largest contribution of these fluctuations survive at positive temperatures. Specifically, we need to quantify the growth rate of the torus partition function with all spins constrained to lie within  $\Delta$  of a given direction:

**Lemma 5.10** *For each  $\epsilon > 0$  there is  $\delta > 0$  such that if  $\beta, \Delta$  obey*

$$\beta \Delta^2 > \frac{1}{\delta} \quad \text{and} \quad \beta \Delta^3 < \delta \quad (5.47)$$

then for every  $\hat{v}_\theta = (\cos \theta, \sin \theta) \in \mathbb{S}^1$ ,

$$E_{\otimes \mu_0} \left( e^{-\beta H_L(S)} \prod_{x \in \mathbb{T}_L} 1_{\{|S_x - \hat{v}_\theta| < \Delta\}} \right) = \left( \frac{2\pi}{\beta} \right)^{L^2/2} e^{-L^2[F(\theta) + o(\epsilon)]}, \quad (5.48)$$

where

$$F(\theta) = \frac{1}{2} \int \frac{dk}{(2\pi)^2} \log \{ \sin^2(\theta) |1 - e^{ik_1}|^2 + \cos^2(\theta) |1 - e^{ik_2}|^2 \}. \quad (5.49)$$

The quantity  $F$  may be interpreted as *spin-wave free energy*. A convexity argument—based on the fact  $\sin^2(\theta) + \cos^2(\theta) = 1$ —now shows that  $F$  is minimized by  $\theta = 0, \pi/2, \pi, 3\pi/2$ , i.e., exactly in one of the coordinate directions. We will fix  $\kappa > 0$  and let  $\Delta = \beta^{-\frac{5}{12}}$  and  $B = \log \beta$  and let  $\mathcal{B}_E$  and  $\mathcal{B}_{SW}$  be the following events:

- (1)  $\mathcal{B}_E = \{ \text{a pair of neighboring spins in } \Lambda_B \text{ make angle larger than } \Delta \text{ with each other} \}$
- (2)  $\mathcal{B}_{SW}$  is the event that the block  $\Lambda_B$  has all neighboring spins within  $\Delta$  of each other with at least  $\kappa \gg \Delta$  from one of the four coordinate direction

The calculation in Lemma 5.10 and a simple use of the subadditivity lemma show

$$\mathfrak{z}(\mathcal{B}_{SW}) \leq \frac{c_1}{\Delta} e^{-c_2 B^3 \kappa^2} \quad (5.50)$$

for some constants  $c_1, c_2 > 0$ . Thus, once  $\beta \gg 1$ , the density of such blocks in any typical configuration from  $\mu_{L,\beta}$  will be rather small. Similarly we prove that

$$\mathfrak{z}(\mathcal{B}_E) \leq B^3 e^{-c_3 \beta \Delta^2} \quad (5.51)$$

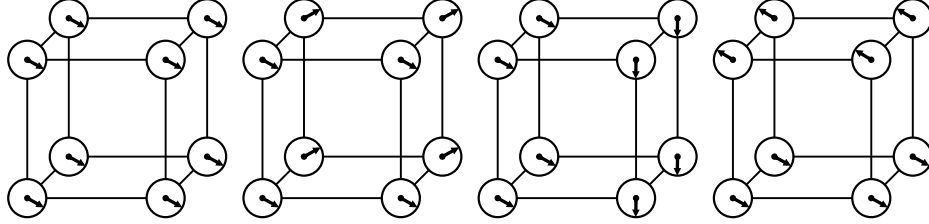
which for  $\beta \gg 1$  is also very small. Thus the *bad* blocks—i.e., those where  $\mathcal{B}_E \cup \mathcal{B}_{SW}$  occurs—are unlikely. However, if a block is aligned in one coordinate direction and another block is aligned in a different direction, they must be separated by a “circuit” of bad blocks. Such circuits are improbable which leads to phase separation. Details of these calculations—which extend even to quantum setting—can be found in a paper by Chayes, Starr and the present author.

120-degree model: Here we focus on  $d = 3$ . The spins,  $S_x$ , take values in  $\mathbb{S}^1$ ; the Hamiltonian looks similar to the orbital compass model except that  $S_x^{(\alpha)}$  no longer represents a Cartesian component. Explicitly, let  $\hat{b}_1, \hat{b}_2, \hat{b}_3$  denote the vectors in  $\mathbb{S}^1$  representing the three complex roots of unity. Then

$$H(S) = \sum_x \sum_{\alpha=1,2,3} (S_x \cdot \hat{b}_\alpha - S_{x+\hat{e}_\alpha} \cdot \hat{b}_\alpha)^2. \quad (5.52)$$

Again, all constant configurations are ground states; and further ground states may again be obtained by judicious reflections. Fortunately, even in  $d = 3$  the number of energy-preserving operations one can perform on ground states is much smaller than for the orbital

compass model, and all ground states can thus be classified. Indeed, given a ground state configuration, every unit cube in  $\mathbb{Z}^3$  looks as one of the four cubes in the picture



modulo, of course, a simultaneous rotation of all spins. Here is what we can say rigorously about this model:

**Theorem 5.11** *Let  $\hat{w}_1, \dots, \hat{w}_6 \in \mathbb{S}^1$  be the six sixth roots of unity. For each  $\epsilon > 0$  there exist  $\beta_0 > 0$  and, for each  $\beta \geq \beta_0$ , there exist six distinct, shift-ergodic Gibbs measures  $\mu_1, \dots, \mu_6 \in \mathfrak{G}_\beta$  such that*

$$E_{\mu_j}(S_x \cdot \hat{w}_j) \geq 1 - \epsilon, \quad j = 1, \dots, 6. \quad (5.53)$$

*There are no shift-ergodic  $\mu \in \mathfrak{G}_\beta$ ,  $\beta \geq \beta_0$ , for which  $\max_{j=1, \dots, 6} E_\mu(S_x \cdot \hat{w}_j) < 1 - \epsilon$ .*

The idea underlying this theorem is quite similar to the orbital compass model; e.g., the analogue of the spin-wave free energy (5.49) is

$$F(\theta) = \frac{1}{2} \int \frac{dk}{(2\pi)^3} \left[ \log \sum_{\alpha=1,2,3} q_\alpha(\theta) |1 - e^{ik_\alpha}|^2 \right] \quad (5.54)$$

where  $q_1 = \sin^2(\theta)$ ,  $q_2 = \sin^2(\theta - 120^\circ)$  and  $q_3 = \sin^2(\theta + 120^\circ)$ . Again, a rather sophisticated argument shows that  $F$  is minimal only for  $\theta$  of the form  $\frac{\pi}{3}j$ ,  $j = 1, \dots, 6$ . The rest of the argument is as for the orbital-compass model. See the paper of Chayes, Nussinov and the present author for further details.

n.n. and n.n.n. antiferromagnet: Finally, we consider a toy model that exemplifies the features of both systems above. Here  $d = 2$  and the spins take again values in  $\mathbb{S}^1$ , but the interaction is *antiferromagnetic*—that is, with a preference for antialignment—for both nearest and next-nearest neighbors:

$$H(S) = \gamma \sum_x [S_x \cdot S_{x+\hat{e}_1} + S_x \cdot S_{x+\hat{e}_2}] + \sum_x [S_x \cdot S_{x+\hat{e}_1+\hat{e}_2} + S_x \cdot S_{x+\hat{e}_1-\hat{e}_2}] \quad (5.55)$$

Assuming  $|\gamma| < 2$ , we obtain the minimum energy state by first enforcing the n.n.n. constraints—there is an antiferromagnetic, or Néel, order on both even and odd sublattice—and only then worrying about how to satiate the n.n. constraint. But once the sublattices are ordered antiferromagnetically, the net interaction between the sublattices is zero—and so each of the sublattices can be rotated independently! A configuration of this form is depicted in the figure below. We can prove the following theorem:

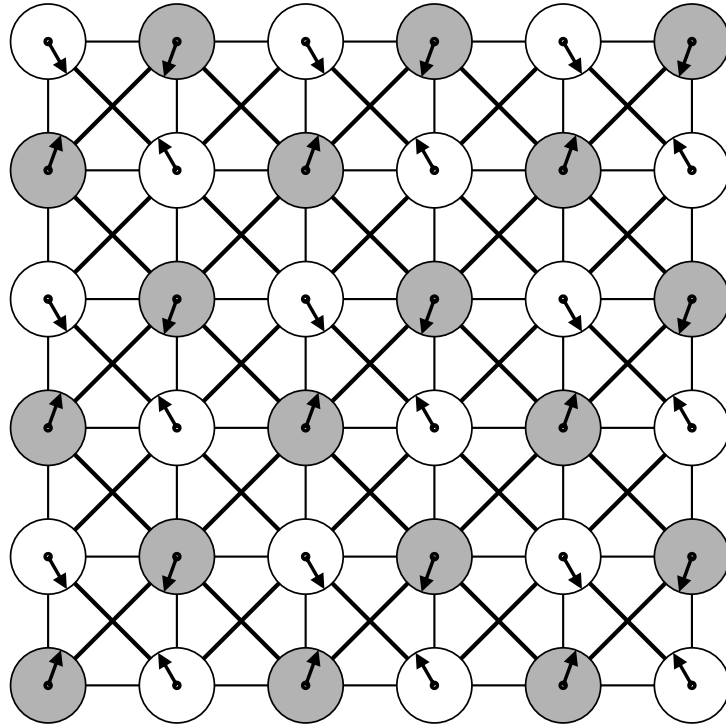
**Theorem 5.12** *For each  $\epsilon > 0$  there exist  $\beta_0 > 0$  and, for each  $\beta \geq \beta_0$ , there exist two distinct, shift-ergodic Gibbs measures  $\mu_1, \mu_2 \in \mathfrak{G}_\beta$  such that*

$$-E_{\mu_j}(S_x \cdot S_{x+\hat{e}_1 \pm \hat{e}_2}) \geq 1 - \epsilon \quad (5.56)$$

and

$$E_{\mu_j}(S_x \cdot S_{x+\hat{e}_j}) \geq 1 - \epsilon, \quad j = 1, 2. \quad (5.57)$$

*There are no shift-ergodic  $\mu \in \mathfrak{G}_\beta$ ,  $\beta \geq \beta_0$ , for which either (5.56) or at least one of (5.57) does not hold.*



As for the two models above, everything boils down to a spin-wave calculations in the end. Here the relevant parameter is the relative orientation  $\theta$  of the two antiferromagnetically ordered sublattices. The spin-wave free energy is then

$$F(\theta) = \frac{1}{2} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log D_{\mathbf{k}}(\theta), \quad (5.58)$$

where

$$D_{\mathbf{k}}(\theta) = |1 - e^{i(k_1+k_2)}|^2 + |1 - e^{i(k_1-k_2)}|^2 + \gamma \cos(\theta)(|1 - e^{ik_1}|^2 - |1 - e^{ik_1}|^2). \quad (5.59)$$

As it turns out,  $F$  is minimized by  $\theta = 0$  or  $\theta = \pi$ . In terms of the typical spin configurations, the former corresponds to horizontal alignment and vertical antialignment of nearest neighbors, and the latter to horizontal antialignment and vertical alignment, i.e., stripe states. See the paper by Chayes, Kivelson and the present author for further details.

## 5.5 Literature remarks

The Gaussian double-well model is a standard example which can be treated either by methods of reflection positivity, or by Pirogov-Sinai theory [24]. Representations (5.4) have been used already by Külske [52, 53] and Zahradník [71]. The method of proof presented here draws on the work of Dobrushin, Kotecký and Shlosman [22, 49, 50] which was used to control order-disorder transitions in a number of systems; most notably, the  $q$ -state Potts model with  $q \gg 1$  [49]. These methods can be combined with graphical representations of Edwards-Sokal [27] (or Fortuin-Kasteleyn [31]) to establish rather complicated phase diagrams, e.g., [9, 19]. Recently, the method has been used to resolve a controversy about a transition can occur in 2D non-linear vector models [29, 30].

Theorem 5.7 has been proved by Funaki and Spohn [39]. As already mentioned, their proof is based on convexity properties of the potential  $V$ —by invoking the Brascamp-Lieb inequality as well as certain coupling argument to the natural dynamical version of the model—and so it does not extend beyond the convex case. (A review of the gradient measures, and further intriguing results, can be found in Funaki [38], Velenik [68] or Sheffield [60].) Theorem 5.8 has been proved by Biskup and Kotecký [14] except for the convergence to the GFF in each of the coexisting phases whose proof in the process of writing. (I would like to thank Herbert Spohn for bringing this question to my attention.)

The interest in models in Sect. 5.4 came from a physics controversy about whether orbital ordering in transition-metal oxides exists at low temperatures. On the basis of rigorous work by Biskup, Chayes and Nussinov [10] (120-degree model) Biskup, Chayes, Nussinov and Starr [11, 12] (2D and 3D orbital compass model), it was demonstrated that, at least at the level of classical models, spin-wave fluctuations stabilize certain ground states [55]. The conclusions hold also the 2D quantum orbital-compass model with large quantum spins [12]. The mechanism of entropic stabilization—or, in physics jargon, *order by disorder*—is most clearly demonstrated in the n.n. & n.n.n. antiferromagnet studied by Biskup, Chayes and Kivelson [8]. This model actually goes back to the original order-by-disorder physics arguments by Shender [61] and Henley [46].

All three theorems in Sect. 5.4 have, apart from an existence clause, also a clause on the *absence* of ergodic states whose local properties deviate from those whose existence was asserted. Actually, these were not the content of the original work [8, 10] because, at that time, the focus on torus measures dictated by reflection positivity was deemed to make it impossible to *rule out* the occurrence of some exotic measures. A passage to such statements was opened by the work of Biskup and Kotecký [13]; the non-existence clauses in Theorems 5.9, 5.11 and 5.12 are direct consequences of the main result of [13] and the method of proof of the existence part. This technique does not quite apply in the setting of gradient models due to the strong role the boundary conditions play in this case.



# Three open problems

We finish with a brief discussion of three general open problems of the subject covered by these notes which the present author finds worthy of significant research effort.

In Chapters 2 and 3 we have shown how useful the infrared bound is in proofs of symmetry breaking and control of the mean-field approximation. Unfortunately, the only way we currently have for proving the IRB is reflection positivity. So our first problem is:

**Problem 1** *Consider models with the Hamiltonian  $H = -\sum_{\langle x,y \rangle} S_x \cdot S_y$ . Prove the IRB directly without appeal to RP.*

As already mentioned, a successful attempt in this direction has been made by Sakai [59], who managed to apply the lace expansion to a modified random current representation of the Ising model used previously by Aizenman [1]. Notwithstanding, here we have in mind something perhaps more robust which addresses directly the principal reason why we need RP, which is that

the spins  $(S_x)$  are not *a priori* independent Gaussian

A step in this direction is the *spherical approximation* for the  $O(n)$  model (see [64, Section II.11]) in which the constraint  $|S_x| = 1$  at every spin is replaced by a (global) constraint on  $\sum_x |S_x|^2$ . This approximation is asymptotic as  $n \rightarrow \infty$ . A paper by Burioni, Cassi and Vezzani [17] claims progress on Problem 1 on general graphs but the level of control of certain steps in their argument seems, at this point, not sufficient.

The IRB is often viewed as a rigorous version of *spin-wave theory*. This theory, initiated in the work of Dyson [25] and others, describes continuous deformations of the lowest energy states by means of an appropriate Gaussian field theory. In Chapter 5 we saw that chessboard estimates may be applied *in conjunction* with spin-wave calculations—which are generally deemed to be the realm of the IRB—to prove phase transitions. This was possible because spin-waves disqualified all but a finite number of ground states from candidacy for low-temperature states. Notwithstanding, one might be able to do the same even in the presence of infinitely many low-temperature states:

**Problem 2** *Prove symmetry breaking at low temperatures in systems with continuous internal symmetry—e.g., the  $O(2)$ -model—without the use of the IRB. The use of chessboard estimates is allowed.*

Further motivation to look at this problem comes from quantum theory: The quantum Heisenberg ferromagnet is not RP (see Speer [66]) and so there is no proof of the IRB

and, consequently, no proof of low-temperature symmetry breaking. On the other hand, the classical Heisenberg ferromagnet is RP and so the spin-condensation argument applies. If we had a more robust proof of symmetry breaking in the classical model, e.g., using chess-board estimates, one might hope to extend the techniques of Biskup, Chayes and Starr [12] to include also the quantum system.

While the theory described in these notes is not restricted exclusively to ferromagnetic systems, in order to have the IRB one needs a good deal of attractivity in the system. It is actually clear that the IRB cannot hold as stated for genuine antiferromagnets—those which cannot be turned to ferromagnets by reversing every other spin—e.g., hard core lattice gas, which is a model with variables  $n_x \in \{0, 1\}$  and the “Gibbs” weight proportional to

$$\lambda^{\sum_x n_x} \prod_{\langle x, y \rangle} (1 - n_x n_y), \quad (5.60)$$

or the  $q$ -state Potts antiferromagnet, which is the model in (1.7) with  $J < 0$ . Indeed, the *staggered* long-range order, which is known to occur in the hard core lattice gas once  $\lambda \gg 1$ , implies that the macroscopically occupied mode is  $k = (\pi, \dots, \pi)$  rather than  $k = 0$ . Nevertheless, we hope that some progress can be made and so we pose:

**Problem 3** *Derive a version of the IRB for the hard-core lattice gas and/or the  $q$ -state Potts antiferromagnet at zero temperature.*

Solving this problem would, hopefully, also provide an easier passage to the proof that the critical  $\lambda$  for the appearance of staggered order tends to zero as  $d \rightarrow \infty$ —in fact, if the mean-field theory is right then one should have  $\lambda_c \sim c/d$ —and that the 3-coloring of  $\mathbb{Z}^2$  exhibits six distinct extremal measures of maximal entropy. These results have recently been obtained by sophisticated contour-counting arguments [40, 41].

# Bibliography

- [1] M. Aizenman, *Geometric analysis of  $\phi^4$  fields and Ising models. I, II.*, Commun. Math. Phys. **86** (1982) 1–48.
- [2] M. Aizenman and R. Fernández, *On the critical behavior of the magnetization in high-dimensional Ising models*, J. Statist. Phys. **44** (1986) 393–454.
- [3] M. Aizenman and R. Fernández, *Critical exponents for long-range interactions*, Lett. Math. Phys. **16** (1988), no. 1, 39–49.
- [4] N. Angelescu and V.A. Zagrebnov, *A lattice model of liquid crystals with matrix order parameter*, J. Phys. A **15** (1982), no. 11, L639–L643.
- [5] M. Biskup, *Reflection positivity of the random-cluster measure invalidated for non-integer  $q$* , J. Statist. Phys. **92** (1998) 369–375.
- [6] M. Biskup and L. Chayes, *Rigorous analysis of discontinuous phase transitions via mean-field bounds*, Commun. Math. Phys. **238** (2003), no. 1-2, 53–93.
- [7] M. Biskup, L. Chayes and N. Crawford, *Mean-field driven first-order phase transitions in systems with long-range interactions*, J. Statist. Phys. **122** (2006), no. 6, 1139–1193.
- [8] M. Biskup, L. Chayes, and S.A. Kivelson, *Order by disorder, without order, in a two-dimensional spin system with  $O(2)$ -symmetry*, Ann. Henri Poincaré **5** (2004), no. 6, 1181–1205.
- [9] M. Biskup, L. Chayes, and R. Kotecký, *Coexistence of partially disordered/ordered phases in an extended Potts model*, J. Statist. Phys. **99** (2000), no. 5/6, 1169–1206.
- [10] M. Biskup, L. Chayes, and Z. Nussinov, *Orbital ordering in transition-metal compounds: I. The 120-degree model*, Commun. Math. Phys. **255** (2005) 253–292.
- [11] M. Biskup, L. Chayes and Z. Nussinov, *Orbital ordering in transition-metal compounds: II. The orbital compass model*, in preparation.
- [12] M. Biskup, L. Chayes and S. Starr, *Quantum spin systems at positive temperature*, Commun. Math. Phys. (to appear)
- [13] M. Biskup and R. Kotecký, *Forbidden gap argument for phase transitions proved by means of chessboard estimates*, Commun. Math. Phys. **264** (2006), no. 3, 631–656.

- [14] M. Biskup and R. Kotecký, *Phase coexistence of gradient Gibbs states*, Probab. Theory Rel. Fields (to appear)
- [15] T. Bodineau, *Translation invariant Gibbs states for the Ising model*, Probab. Theory Related Fields **135** (2006), no. 2, 153–168.
- [16] J. Bricmont, H. Kesten, J.L. Lebowitz and R.H. Schonmann, *A note on the Ising model in high dimensions*, Commun. Math. Phys. **122** (1989) 597–607.
- [17] R. Burioni, D. Cassi and A. Vezzani, *Transience on the average and spontaneous symmetry breaking on graphs*, J. Phys. A **32** (1999), no. 30, 5539–5550.
- [18] L. Chayes, R. Kotecký, and S. B. Shlosman. *Staggered phases in diluted systems with continuous spins*, Commun. Math. Phys. **189** (1997) 631–640.
- [19] L. Chayes and J. Machta, *Graphical representations and cluster algorithms. Part I: Discrete spin systems*, Physica A **239** (1997) 542–601.
- [20] P. Curie, *Propriétés magnétiques des corps a diverses températures*, Ann. de Chimie et Physique **5** (1885) 289; reprinted in *Œuvres de Pierre Curie* (Gauthier-Villars, Paris, 1908) pp. 232–334.
- [21] R. Dobrushin, *The description of a random field by means of conditional probabilities and conditions of its regularity*, Theor. Prob. Appl. **13** (1968) 197–224.
- [22] R.L. Dobrushin and S.B. Shlosman, *Phases corresponding to minima of the local energy*, Selecta Math. Soviet. **1** (1981), no. 4, 317–338.
- [23] R.L. Dobrushin and S. Shlosman, *Absence of breakdown of continuous symmetry in two-dimensional models of statistical physics*, Commun. Math. Phys. **42** (1975), no. 1, 31–40.
- [24] R.L. Dobrushin and M. Zahradník, *Phase diagrams for continuous-spin models: an extension of the Pirogov-Sinai theory*, In: R.L. Dobrushin (ed.), *Mathematical problems of statistical mechanics and dynamics*, pp. 1–123, Math. Appl. (Soviet Ser.), vol. 6, Reidel, Dordrecht, 1986.
- [25] F.J. Dyson, *General theory of spin-wave interactions*, Phys. Rev. **102** (1956), no. 5, 1217–1230.
- [26] F.J. Dyson, E.H. Lieb and B. Simon, *Phase transitions in quantum spin systems with isotropic and nonisotropic interactions*, J. Statist. Phys. **18** (1978) 335–383.
- [27] R.G. Edwards and A.D. Sokal, *Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988) 2009–2012.
- [28] R.S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Grundlehren der Mathematischen Wissenschaften, vol. 271 (Springer-Verlag, New York, 1985).
- [29] A.C.D. van Enter and S.B. Shlosman, *First-order transitions for n-vector models in two and more dimensions: Rigorous proof*, Phys. Rev. Lett. **89** (2002) 285702.

- [30] A.C.D. van Enter and S.B. Shlosman, *Provable first-order transitions for nonlinear vector and gauge models with continuous symmetries*, Commun. Math. Phys. **255** (2005) 21–32.
- [31] C.M. Fortuin and P.W. Kasteleyn, *On the random cluster model. I. Introduction and relation to other models*, Physica (Amsterdam) **57** (1972) 536–564.
- [32] J. Fröhlich, *On the triviality of  $\lambda\phi_d^4$  theories and the approach to the critical point in  $d \geq 4$  dimensions*, Nucl. Phys. B **200** (1982), no. 2, 281–296.
- [33] J. Fröhlich, R. Israel, E.H. Lieb and B. Simon, *Phase transitions and reflection positivity. I. General theory and long-range lattice models*, Commun. Math. Phys. **62** (1978), no. 1, 1–34.
- [34] J. Fröhlich, R. Israel, E.H. Lieb and B. Simon, *Phase transitions and reflection positivity. II. Lattice systems with short-range and Coulomb interactions*, J. Statist. Phys. **22** (1980), no. 3, 297–347.
- [35] J. Fröhlich and E.H. Lieb, *Phase transitions in anisotropic lattice spin systems*, Commun. Math. Phys. **60** (1978), no. 3, 233–267.
- [36] J. Fröhlich and C.-E. Pfister, *On the absence of spontaneous symmetry breaking and of crystalline ordering in two-dimensional systems*, Commun. Math. Phys. **81** (1981), no. 2, 277–298.
- [37] J. Fröhlich, B. Simon and T. Spencer, *Infrared bounds, phase transitions and continuous symmetry breaking*, Commun. Math. Phys. **50** (1976) 79–95.
- [38] T. Funaki, *Stochastic Interface Models*, Lecture Notes for the International Probability School at Saint-Flour, Lecture Notes in Math., 1869, Springer, Berlin, 2005.
- [39] T. Funaki and H. Spohn, *Motion by mean curvature from the Ginzburg-Landau  $\nabla\phi$  interface model*, Commun. Math. Phys. **185** (1997) 1–36.
- [40] D. Galvin and J. Kahn, *On phase transition in the hard-core model on  $\mathbb{Z}^d$* , Combin. Probab. Comput. **13** (2004), no. 2, 137–164.
- [41] D. Galvin and D. Randall, *Torpid mixing of local markov chains on 3-Colorings of the discrete torus*, Proceedings of the Eighteenth Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2007), to appear.
- [42] P.G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Oxford University Press, New York, 1993.
- [43] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter Studies in Mathematics, vol. 9 (Walter de Gruyter & Co., Berlin, 1988).
- [44] T. Gobron and I. Merola, *First-order phase transition in Potts models with finite-range interactions*, J. Statist. Phys. (to appear)

- [45] G. Grimmett, *The Random-Cluster Model*, Probability on discrete structures, 73–123, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004.
- [46] C.L. Henley, *Ordering due to disorder in a frustrated vector antiferromagnet*, Phys. Rev. Lett. **62** (1989) 2056–2059.
- [47] D. Ioffe, S. Shlosman and Y. Velenik, *2D models of statistical physics with continuous symmetry: The case of singular interactions*, Commun. Math. Phys. **226** (2002), no. 2, 433–454.
- [48] R.B. Israel, *Convexity in the theory of lattice gases*, With an introduction by Arthur S. Wightman. Princeton University Press, Princeton, N.J., 1979.
- [49] H. Kesten and R. Schonmann, *Behavior in large dimensions of the Potts and Heisenberg models*, Rev. Math. Phys. **1** (1990) 147–182.
- [50] R. Kotecký and S.B. Shlosman, *Existence of first-order transitions for Potts models*, In: S. Albeverio, Ph. Combe, M. Sirigue-Collins (eds.), Proc. of the International Workshop—Stochastic Processes in Quantum Theory and Statistical Physics, Lecture Notes in Physics **173**, 248–253, Springer-Verlag, Berlin-Heidelberg-New York, 1982.
- [51] R. Kotecký and S.B. Shlosman, *First-order phase transitions in large entropy lattice models*, Commun. Math. Phys. **83** (1982), no. 4, 493–515.
- [52] C. Külske, *The continuous spin random field model: Ferromagnetic ordering in  $d \geq 3$* , Rev. Math. Phys. **11** (1999), no.10, 1269–1314.
- [53] C. Külske, *Stability for a continuous SOS-interface model in a randomly perturbed periodic potential*, WIAS Preprint no. 466 (1998); cf Christof’s website: <http://www.math.rug.nl/~kuelske/publications.html>
- [54] N.D. Mermin and H. Wagner, *Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models*, Phys. Rev. Lett. **17** (1966), no. 22, 1133–1136.
- [55] Z. Nussinov, M. Biskup, L. Chayes and J. van den Brink, *Orbital order in classical models of transition-metal compounds*, Europhys. Lett. **67** (2004), no. 6, 990–996.
- [56] K. Osterwalder and R. Schrader, *Axioms for Euclidean Green’s functions*, Commun. Math. Phys. **31** (1973), 83–112.
- [57] C.-E. Pfister, *On the symmetry of the Gibbs states in two-dimensional lattice systems*, Commun. Math. Phys. **79** (1981), no. 2, 181–188.
- [58] D. Ruelle, *Statistical Mechanics. Rigorous Results*, Reprint of the 1989 edition, World Scientific Publishing Co., Inc., River Edge, NJ; Imperial College Press, London, 1999.
- [59] A. Sakai, *Lace expansion for the Ising model*, arxiv:math-ph/0510093.
- [60] S. Sheffield, *Random Surfaces*, Asterisque 2005, No. 304, 177pp.

- [61] E.F. Shender, *Antiferromagnetic garnets with fluctuationally interacting sublattices*, Sov. Phys. JETP **56** (1982) 178–184.
- [62] S. Shlosman, *Phase transitions for two-dimensional models with isotropic short range interactions and continuous symmetry*, Commun. Math. Phys. **71** (1980) 207–212.
- [63] S.B. Shlosman, *The method of reflective positivity in the mathematical theory of phase transitions of the first kind* (Russian), Uspekhi Mat. Nauk **41** (1986), no. 3(249), 69–111, 240.
- [64] B. Simon, *The Statistical Mechanics of Lattice Fases*, Vol. I., Princeton Series in Physics (Princeton University Press, Princeton, NJ, 1993).
- [65] A.D. Sokal, *An alternate constructive approach to the  $\phi_3^4$  quantum field theory, and a possible destructive approach to  $\phi_4^4$* , Ann. Inst. H. Poincaré Sect. A **37** (1982), no. 4, 317–398.
- [66] E.R. Speer, *Failure of reflection positivity in the quantum Heisenberg ferromagnet*, Lett. Math. Phys. **10** (1985), no. 1, 41–47.
- [67] B. Tóth's website: [http://www.math.bme.hu/~balint/oktatas/statisztikus\\_fizika/jegyzet/](http://www.math.bme.hu/~balint/oktatas/statisztikus_fizika/jegyzet/)
- [68] Y. Velenik, *Localization and delocalization of random interfaces*, Probab. Surveys **3** (2006) 112–169.
- [69] P. Weiss, *L'Hypothèse du champ moléculaire et la proprié ferromagnétique*, J. de Physique **6** (1907) 661.
- [70] F.Y. Wu, *The Potts model*, Rev. Modern Phys. **54** (1982) 235–268.
- [71] M. Zahradník, *Contour methods and Pirogov-Sinai theory for continuous spin lattice models*, In: R.A. Minlos, S. Shlosman and Yu.M. Suhov (eds.), *On Dobrushin's way. From probability theory to statistical physics*, pp. 197–220, Amer. Math. Soc. Transl. Ser. 2, vol. 198, Amer. Math. Soc., Providence, RI, 2000.