

---

# Reflection Positivity and Phase Transitions in Lattice Spin Models

Marek Biskup

Department of Mathematics, UCLA, Los Angeles, CA 90095-1555  
biskup@math.ucla.edu

## Contents

<b>1</b>	<b>Introduction</b> . . . . .	2
<b>2</b>	<b>Lattice Spin Models: Crash Course</b> . . . . .	5
2.1	Basic Setup . . . . .	5
2.2	Examples . . . . .	7
2.3	Gibbs Formalism . . . . .	10
2.4	Torus Measures . . . . .	13
2.5	Some Thermodynamics . . . . .	14
2.6	Literature Remarks . . . . .	16
<b>3</b>	<b>Infrared Bound &amp; Spin-wave Condensation</b> . . . . .	17
3.1	Random Walk Connections . . . . .	18
3.2	Infrared Bound . . . . .	20
3.3	Spin-wave Condensation in $O(n)$ -model . . . . .	22
3.4	Literature Remarks . . . . .	27
<b>4</b>	<b>Infrared Bound &amp; Mean-field Theory</b> . . . . .	29
4.1	Mean-field Theory . . . . .	29
4.2	Example: The Potts Model . . . . .	32
4.3	Approximation Theorem & Applications . . . . .	34
4.4	Ideas from the Proofs . . . . .	37
4.5	Literature Remarks . . . . .	41
<b>5</b>	<b>Reflection Positivity</b> . . . . .	43
5.1	Reflection Positive Measures . . . . .	43
5.2	Gaussian Domination . . . . .	47
5.3	Chessboard Estimates . . . . .	49

5.4	Diagonal Reflections, Other Lattices .....	54
5.5	Literature Remarks .....	57
<b>6</b>	<b>Applications of Chessboard Estimates .....</b>	<b>58</b>
6.1	Gaussian Double-well Model .....	59
6.2	Proof of Phase Coexistence .....	61
6.3	Gradient Fields with Non-convex Potential .....	66
6.4	Spin-waves vs Infinite Ground-state Degeneracy .....	70
6.5	Literature Remarks .....	75
<b>7</b>	<b>Topics Not Covered .....</b>	<b>76</b>
<b>8</b>	<b>Three Open Problems .....</b>	<b>78</b>
	<b>References .....</b>	<b>80</b>

## 1 Introduction

Phase transitions are one of the most fascinating, and also most perplexing, phenomena in equilibrium statistical mechanics. On the physics side, many approximate methods to explain or otherwise justify phase transitions are known but a complete mathematical understanding is available only in a handful of simplest of all cases. One set of tractable systems consists of the so called *lattice spin models*. Originally, these came to existence as simplified versions of (somewhat more realistic) models of crystalline materials in solid state physics but their versatile nature earned them a life of their own in many other disciplines where complex systems are of interest.

The present set of notes describes one successful *mathematical* approach to phase transitions in lattice spin models which is based on the technique of *reflection positivity*. This technique was developed in the late 1970s in the groundbreaking works of F. Dyson, J. Fröhlich, R. Israel, E. Lieb, B. Simon and T. Spencer who used it to establish phase transitions in a host of physically-interesting classical and quantum lattice spin models; most notably, the classical Heisenberg ferromagnet and the quantum XY model and Heisenberg antiferromagnet. Other powerful techniques — e.g., Pirogov-Sinai theory, lace expansion or multiscale analysis in field theory — are available at present that can serve a similar purpose in related contexts, but we will leave their review to experts in those areas.

The most attractive feature of reflection positivity — especially, compared to the alternative techniques — is the simplicity of the resulting proofs. There are generally two types of arguments one can

use: The first one is to derive the so called *infrared bound*, which states in quantitative terms that the fluctuations of the spin variables are dominated by those of a lattice Gaussian free field. In systems with an internal symmetry, this yields a proof of a symmetry-breaking phase transition by way of a spin-condensation argument. Another route goes via the so called *chessboard estimates*, which allow one to implement a Peierls-type argument regardless of whether the model exhibits an internal symmetry or not.

Avid users of the alternative techniques are often quick to point out that the simplicity of proofs has its price: As a rather restrictive condition, reflection positivity applies only to a small (in a well defined sense) class of systems. Fortunately for the technique and mathematical physics in general, the models to which it does apply constitute a large portion of what is interesting for physics, *and* to physicists. Thus, unless one is exclusively after 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, R. Dobrushin, R. Kotecký and S. 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 prominently, the  $q$ -state Potts model for  $q \gg 1$ . Another neat application came in the papers of M. 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 L. Chayes, R. Kotecký and S. Shlosman, that the Fisher-renormalization scheme in annealed-diluted systems may be substituted by the emergence of an intermediate phase.

These notes discuss also more recent results where their author had a chance to contribute to the field. The common ground for some of these is the use of reflection positivity to provide mathematical justification of “well-known” conclusions from physics folklore. For instance, in papers by N. Crawford, L. 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 L. Chayes, S. Starr and the present author — asserts that if a reflection-positive quantum

spin system undergoes a phase transition at intermediate temperatures in its classical limit, a similar transition occurs in the quantum system provided the magnitude of the quantum spin is sufficiently large.

There have also been recent 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; it was debated whether a transition can occur already in 2 dimensions. This was settled in recent work of A. van Enter and S. 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 L. Chayes, S. Kivelson, Z. Nussinov 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.

These recent activities show that the full potential of reflection positivity may not yet have been fully exhausted and that the technique will continue to play an important role in mathematical statistical mechanics. It is hoped that the present text will help newcomers to this field learn the essentials of the subject before the need arises to plow through the research papers where the original derivations first appeared.

## Organization

This text began as class notes for nine hours of lectures on reflection positivity at the 2006 Prague School and gradually grew into a survey of (part of) this research area. The presentation opens with a review of basic facts about lattice spin models and then proceeds to study two applications of the infrared bound: a spin-condensation argument and a link to mean-field theory. These are followed by the classical derivation of the infrared bound from reflection positivity. The remainder of the notes is spent on applications of a by-product of this derivation, the chessboard estimate, to proofs of phase coexistence.

The emphasis of the notes is on a *pedagogical* introduction to reflection positivity; for this reason we often sacrifice on generality and rather demonstrate the main ideas on the simplest case of interest. To compensate for the inevitable loss of generality, each chapter is endowed with a section “Literature remarks” where we attempt to list the references deemed most relevant to the topic at hand. The notes are closed with a short section on topics that are not covered as well as some open problems that the author finds worthy of some thought.

## Acknowledgments

A review naturally draws on the work of many authors. We will do our best to give proper credit to their contribution in the sections “Literature remarks” that appear at the end of each chapter and, of course, in the list of references. I have personally learned a great deal of the subject from Lincoln Chayes with whom I have subsequently coauthored more than half-a-dozen papers some of which will be discussed here. During this time I also benefited from collaborations with N. Crawford, S. Kivelson, R. Kotecký, Z. Nussinov and S. Starr, and from various enlightening discussions involving A. van Enter, B. Nachtergaele and S. Shlosman.

The text would presumably never exist were it not for Roman Kotecký’s summer school; I wish to thank him for allowing me to speak on this subject. I am indebted to the participants of the school for comments during the lectures and to T. Bodineau, A. van Enter, E. Lieb and S. Shlosman for suggestions on the first draft of the notes. 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.

## 2 Lattice Spin Models: Crash Course

This section 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 these notes; readers wishing a more comprehensive — and in-depth — treatment should consult classic textbooks on the subject.

### 2.1 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 the 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

$$S := (S_\Lambda, S_{\Lambda^c}) \quad (2.1)$$

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

$$H_\Lambda(S) := \sum_{\substack{A \subset \mathbb{Z}^d \text{ finite} \\ A \cap \Lambda \neq \emptyset}} \Phi_A(S) \quad (2.2)$$

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; in particular, should the need arise to talk about unbounded spins.) It is often more convenient to write the above as a formal sum:

$$H(S) := \sum_A \Phi_A(S) \quad (2.3)$$

with the above specific understanding whenever a precise 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). Before the interaction is “switched on,” the spin configurations will be “distributed” according to the product measure, i.e., 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 (in finite volume) takes the general form  $e^{-\beta H(S)} \prod_x \mu_0(dS_x)$ ; cf Sect. 2.3 for more details.

## 2.2 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 on  $\mathbb{S}^{n-1}$ . The Hamiltonian is

$$H(S) := -J \sum_{\langle x,y \rangle} S_x \cdot S_y \quad (2.4)$$

where the dot denotes the usual (Euclidean) dot-product in  $\mathbb{R}^n$  and  $J \geq 0$ . (Note that this comes at no loss as the sign of  $J$  can be changed 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

$$AS_x \cdot AS_y = S_x \cdot S_y \quad (2.5)$$

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 a simultaneous rotation of all spins. (For  $n = 1$  this reduces to the invariance under the flip  $+1 \leftrightarrow -1$ .)

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

$$H(\sigma) := -J \sum_{\langle x,y \rangle} \sigma_x \sigma_y \quad (2.6)$$

Note that the energy is smaller when the spins at nearest neighbors align and higher when they antialign. (A similar statement holds, of

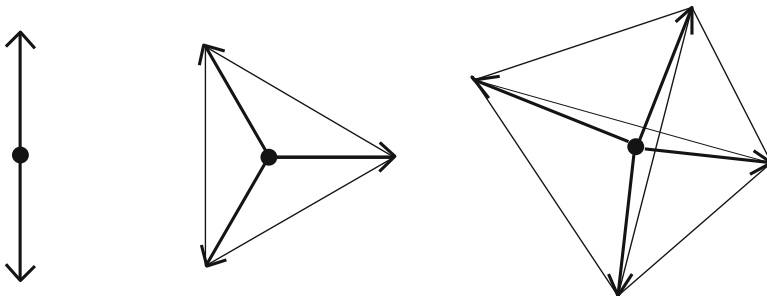
course, for all  $O(n)$  models.) This is due to the choice of the sign  $J \geq 0$  which makes these models *ferromagnets*.

(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 a uniform *a priori* measure). The Hamiltonian is

$$H(\sigma) := -J \sum_{\langle x,y \rangle} \delta_{\sigma_x, \sigma_y} \quad (2.7)$$

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 the so called *random-cluster model*).

It turns out that the Hamiltonian (2.7) can be brought to the form (2.4). 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

$$\hat{v}_\alpha \cdot \hat{v}_\beta = \begin{cases} 1, & \text{if } \alpha = \beta, \\ -\frac{1}{q-1}, & \text{otherwise.} \end{cases} \quad (2.8)$$

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

$$S_x \cdot S_y = \frac{q}{q-1} \delta_{\sigma_x, \sigma_y} - \frac{1}{q-1} \quad (2.9)$$

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

$$H(S) := -\tilde{J} \sum_{\langle x,y \rangle} S_x \cdot S_y \quad (2.10)$$



with  $\tilde{J} := J \frac{q-1}{q}$ . This form, sometimes referred to as *tetrahedral representation*, will be far more useful for our purposes than (2.7).

(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 a uniform *a priori* measure. The Hamiltonian is

$$H(S) := -J \sum_{\langle x,y \rangle} (S_x \cdot S_y)^2 \quad (2.11)$$

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

$$Q_{\alpha\beta} := S^{(\alpha)} S^{(\beta)} - \frac{1}{n} \delta_{\alpha\beta} \quad (2.12)$$

(The subtraction of the identity is rather arbitrary and more or less unnecessary; its goal is to achieve zero trace and thus reduce the number of independent variables characterizing  $Q$  to  $n-1$ ; i.e., 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

$$\text{Tr}(Q\tilde{Q}) = (S \cdot \tilde{S})^2 - \frac{1}{n} \quad (2.13)$$

Since  $Q$  is symmetric, the trace evaluates to

$$\text{Tr}(Q\tilde{Q}) = \sum_{\alpha,\beta} Q_{\alpha\beta} \tilde{Q}_{\alpha\beta} \quad (2.14)$$

which is the canonical scalar product on  $n \times n$  matrices. In this language the Hamiltonian takes again the form we saw in the  $O(n)$  model.

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

$$H(S) = +\frac{1}{2} \sum_{x,y} J_{x,y} |S_x - S_y|^2 \quad (2.15)$$

where  $(J_{xy})$  is a collection of suitable *coupling constants* and  $|\cdot|$  denotes the Euclidean norm in  $\mathbb{R}^n$ . This is possible because, in all cases, the (corresponding) norm of  $S_x$  is constant and so adding it to the Hamiltonian has no effect on the probability measure. The model thus obtained bears striking similarity to our last example:

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

$$H(\phi) := \frac{1}{2} \sum_{x,y} P(x, y) (\phi_y - \phi_x)^2 \quad (2.16)$$

This can be rewritten as

$$H(\phi) = (\phi, (1 - P)\phi)_{L^2(\mathbb{Z}^d)} =: \mathcal{E}_{1-P}(\phi, \phi) \quad (2.17)$$

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 (2.15) and (2.16) 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 formal similarity can nevertheless be exploited to provide information on the models (2.15).

### 2.3 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 probability measure on  $\Omega^\Lambda$  given by

$$\mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\mathrm{d}S_\Lambda) := \frac{e^{-\beta H_\Lambda(S)}}{Z_{\Lambda,\beta}(S_{\Lambda^c})} \prod_{x \in \Lambda} \mu_0(\mathrm{d}S_x) \quad (2.18)$$

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 2.1 (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.  $S_{\Lambda^c}$ ),*

$$\mu_{\Delta,\beta}^{(S_{\Delta^c})}(\cdot | S_{\Lambda^c}) = \mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\cdot) \quad (2.19)$$

*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 2.2 (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.  $S_{\Lambda^c}$ ,*

$$\mu(\cdot | S_{\Lambda^c}) = \mu_{\Lambda,\beta}^{(S_{\Lambda^c})}(\cdot) \quad (2.20)$$

*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) By Lemma 2.1, any weak cluster point of  $(\mu_{\Lambda, \beta}^{(S_{\Lambda^c})})$  belongs to  $\mathfrak{G}_\beta$ .
- (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, which then belongs to  $\mathfrak{G}_\beta$ .
- (3)  $\mathfrak{G}_\beta$  is a convex set (and is closed in the topology of weak convergence). 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 let  $\beta_n$  be a sequence with  $\beta_n \rightarrow \beta < \infty$ . Let  $\mu_n$  be the sequence of the corresponding Gibbs measures. Then every (weak) cluster point of  $(\mu_n)$  is an infinite-volume Gibbs measure for the Hamiltonian  $H$  and inverse temperature  $\beta$ .

Note that the fact that  $\mathfrak{G}_\beta$  is closed and convex ensures that each element can be written as a unique convex combination of extreme points (by the Krein-Millman theorem). The DLR condition permits to extract the corresponding decomposition probabilistically by conditioning on the  $\sigma$ -algebra of tail events.

Now we give a meaning to the terms that are frequently (though sometimes vaguely) employed by physicists:

**Definition 2.3 (Phase coexistence).** *We say that the model is at phase coexistence (or undergoes a first-order phase transition) whenever the parameters are such that  $|\mathfrak{G}_\beta| > 1$ .*

The simplest example where this happens is the Ising model. Let

$$\Lambda_L := \{1, \dots, L\}^d \tag{2.21}$$

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, for the *minus* boundary condition,  $\mu_{\Lambda_L, \beta}^- \xrightarrow{w} \mu^-$ . It turns out that, in dimensions  $d \geq 2$  there exists  $\beta_c(d) \in (0, \infty)$  such that

$$\beta > \beta_c(d) \quad \Rightarrow \quad \mu^+ \neq \mu^- \tag{2.22}$$

i.e., the model is at phase coexistence, while for  $\beta < \beta_c(d)$ , the set of all infinite volume Gibbs measures is a singleton — which means that the model is in the uniqueness regime. One of our goals is to prove similar statements in all of the models introduced above.

## 2.4 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 a boundary — typically a torus. This is the case of *periodic* or *torus boundary conditions*.

Consider the torus  $\mathbb{T}_L$ , which we define as a graph with vertices  $(\mathbb{Z}/L\mathbb{Z})^d$ , endowed with the corresponding (periodic) 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

$$H_L(S) := H_{A_L}(\text{periodic extension of } S \text{ to } \mathbb{Z}^d) \quad (2.23)$$

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

$$H_L(S) = -\frac{1}{2} \sum_{x,y \in \mathbb{T}_L} J_{x,y}^{(L)} S_x \cdot S_y \quad (2.24)$$

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

$$J_{x,y}^{(L)} := \sum_{z \in \mathbb{Z}^d} J_{x,y+Lz} \quad (2.25)$$

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

$$\mu_{L,\beta}(dS) := \frac{e^{-\beta H_L(S)}}{Z_{L,\beta}} \prod_{x \in \mathbb{T}_L} \mu_0(dS_x) \quad (2.26)$$

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

**Lemma 2.4.** *Every (weak) cluster point of  $(\mu_{L,\beta})_{L \geq 1}$  lies in  $\mathfrak{G}_\beta$ .*

Note that there *is* something to prove here because, due to (2.25), the interaction depends on  $L$ .

## 2.5 Some Thermodynamics

Statistical mechanics combines, in its historical development, molecular theory with empirical thermodynamics. Many mathematically rigorous accounts of statistical mechanics thus naturally start the exposition with the notion of the *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 2.5.** *For  $x \in \mathbb{Z}^d$  let  $\tau_x$  be the shift-by- $x$  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 recall that  $\mu_{L,\beta}$  denote the torus Gibbs measures. Then:*

(1) *The limit*

$$f(h) := \lim_{L \rightarrow \infty} \frac{1}{L^d} \log E_{\mu_{L,\beta}} \left\{ \exp \left( h \sum_{x \in \mathbb{T}_L} g \circ \tau_x \right) \right\} \quad (2.27)$$

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

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

$$\left. \frac{\partial f}{\partial h^-} \right|_{h=0} \leq E_\mu(g) \leq \left. \frac{\partial f}{\partial h^+} \right|_{h=0} \quad (2.28)$$

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

$$E_{\mu^\pm}(g) = \left. \frac{\partial f}{\partial h^\pm} \right|_{h=0} \quad (2.29)$$

*Proof of (1), main idea.* For compact state-spaces and absolutely-summable interactions, the existence of the limit follows by standard subadditivity arguments. In fact, the limit will exist and will be the same even if we replace  $\mu_{L,\beta}$  in (2.27) by any sequence of Gibbs measures in  $\Lambda_L$  with (even  $L$ -dependent) boundary conditions. The convexity of  $f$  is a consequence of the Hölder inequality applied to the expectation in (2.27).  $\square$

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

$$Z_L(h) := E_\mu \left\{ \exp \left( h \sum_{x \in \Lambda_L} g \circ \tau_x \right) \right\} \quad (2.30)$$

Since  $\log Z_L$  is convex in  $h$  (again, by Hölder) we have for any  $h > 0$  that

$$\begin{aligned} \log Z_L(h) - \log Z_L(0) &\geq \frac{\partial}{\partial h} \log Z_L(h) \Big|_{h=0} h \\ &= h E_\mu \left( \sum_{x \in \Lambda_L} g \circ \tau_x \right) = h |\Lambda_L| E_\mu(g) \end{aligned} \quad (2.31)$$

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

$$f(h) - f(0) \geq h E_\mu(g) \quad (2.32)$$

Divide by  $h$  and let  $h \downarrow 0$  to get one half of (2.28). 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 the proof of (2) shows that if  $\mu_h \in \mathfrak{G}_{\beta,h}$  is translation-invariant, then

$$\frac{\partial f}{\partial h^-} \leq E_{\mu_h}(g) \leq \frac{\partial f}{\partial h^+} \quad (2.33)$$

In particular, if  $h > 0$  we have

$$E_{\mu_h}(g) \geq \frac{\partial f}{\partial h^-} \underset{h>0}{\geq} \frac{\partial f}{\partial h^+} \Big|_{h=0} \quad (2.34)$$

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

$$E_{\mu^+}(g) \geq \frac{\partial f}{\partial h^+} \Big|_{h=0} \quad (2.35)$$

(The expectations converge because  $g$  is a local — and thus continuous, in the product topology — function.) Applying (2) we verify (2.29) 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

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} g \circ \tau_x \xrightarrow{L \rightarrow \infty} E_{\mu^+}(g), \quad \text{in } \mu^+ \text{-probability} \quad (2.36)$$

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

$$\mu^+ \left( \sum_{x \in \Lambda_L} g \circ \tau_x > (E_{\mu^+}(g) + \epsilon) |\Lambda_L| \right) > \epsilon \quad (2.37)$$

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

$$E_{\mu^+} \left\{ \exp \left( h \sum_{x \in \Lambda_L} g \circ \tau_x \right) \right\} \geq \epsilon e^{|\Lambda_L| h [E_{\mu^+}(g) + \epsilon]} \quad (2.38)$$

In light of the independence of the limit in (1) on the measure we use — as discussed in the sketch of the proof of (1) — this would imply

$$f(h) \geq h(E_{\mu^+}(g) + \epsilon) \quad (2.39)$$

which cannot hold for all  $h > 0$  if the right-derivative of  $f$  at  $h = 0$  is to equal  $E_{\mu^+}(g)$ . Hence (2.36) holds.

By the Pointwise Ergodic Theorem, the convergence in (2.36) actually occurs — and, by (2.36), 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 (2.28) imply that  $f$  is not differentiable at  $h = 0$ . Then one applies (2.29) to infer the existence of two distinct, ergodic Gibbs measures saturating the bounds in (2.28). Examples of this approach will be discussed throughout these notes.

## 2.6 Literature Remarks

This section contains only the minimum necessary to understand the rest of the course. For a comprehensive treatment of Gibbs-measure theory, we refer to classic monographs by Ruelle [88], Israel [66], Simon [97] and Georgii [57]. Further general background on statistical mechanics of such systems can be found in Ruelle's "blue" book [89]. 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. [32].



The proof of Theorem 2.5 touches upon the subject of *large deviation theory* which provides a mathematical framework for many empirical principles underlying classical thermodynamics. The connection of course appears in various disguises in the textbooks [66, 97, 57]; for expositions dealing more exclusively with large-deviation theory we refer to the books by den Hollander [64], Dembo and Zeitouni [29], and Deuschel and Stroock [30]. For the Pointwise Ergodic Theorem and other facts from *ergodic theory* we refer to the textbooks by, e.g., Krengel [73] and Petersen [86].

Stochastic domination and the FKG inequality — dealing with partial ordering of spin configurations, functions thereof and thus also measures — are discussed in, e.g., Georgii [57] or Grimmett [61]. The proof of (2.22) can alternatively be based on Griffiths' correlation inequalities (Griffiths [60]). The phase coexistence in the Ising model at large  $\beta$  was first proved by Peierls via a contour argument that now bears his name (see Griffiths [59]).

Concerning the historical origin of the various model systems; 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 while the Potts model was introduced by Domb and given to Potts as a thesis problem. Ironically, the  $O(1)$ -model bears Ising's name even though his conclusions about it were quite wrong! Apparently, Potts was more deserving.

An excellent reference for mathematical physics of liquid crystals is the monograph by de Gennes and Prost [56]; other, more combinatorial models have been considered by Heilmann and Lieb [62] and Abraham and Heilmann [1]. The tetrahedral representation of the Potts model can be found in Wu's review article [106]; the matrix representation of the liquid-crystal model is an observation of Angelescu and Zagrebnoy [6]. Gradient fields — of which the GFF is the simplest example — have enjoyed considerable attention in recent years; cf the review articles by Funaki [52], Velenik [104] and Sheffield [92]. Another name for the GFF is *harmonic crystal*.

### 3 Infrared Bound & Spin-wave Condensation

The goal of this section is to elucidate the significance of the infrared bound — postponing its proof and connection with reflection positivity until Section 5 — and the use thereof in the proofs of symmetry breaking via the mechanism of *spin-wave condensation*. The presence, and absence, of symmetry breaking in the  $O(n)$ -model with

certain non-negative two-body interactions will be linked to recurrence vs transience of a naturally induced random walk.

### 3.1 Random Walk Connections

Consider the model with the Hamiltonian

$$H = -\frac{1}{2} \sum_{x,y} J_{xy} S_x \cdot S_y \quad (3.1)$$

where the spins  $S_x$  are *a priori* independent and distributed according to a measure  $\mu_0$  which is supported in a compact set  $\Omega \subset \mathbb{R}^\nu$ . Assume that 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 to have unit strength. We will actually always restrict our attention to the following specific examples:

- *Nearest-neighbor interactions:*

$$J_{x,y} = \begin{cases} \frac{1}{2d}, & \text{if } |x - y| = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (3.2)$$

- *Yukawa potentials:*

$$J_{x,y} = C e^{-\mu|x-y|} \quad (3.3)$$

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

- *Power-law decaying potentials:*

$$J_{x,y} = \frac{C}{|x - y|_1^s} \quad (3.4)$$

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

On top of these, we will also permit:

- Any convex combination of the three interactions 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 (see Lemma 5.5). 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 interactions 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

$$P_z(X_{n+1} = y | X_n = x) := J_{xy} \quad (3.5)$$

where  $P_z$  is the law of the chain started at site  $z$ . Of particular interest will be the question whether this random walk is *recurrent* or *transient* — i.e., whether a walk started at the origin returns there infinitely, or only finitely many times. Here is a criterion to this matter:

**Lemma 3.1.** *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*

$$\int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} < \infty \quad (3.6)$$

*Proof.* Recall that a random walk is transient if and only if the first return time to the origin,  $\tau_0 := \inf\{n > 0 : X_n = 0\}$ , is infinite with a positive probability, i.e.,  $P_0(\tau_0 < \infty) < 1$ . By the formula  $E_0 N = [1 - P_0(\tau_0 < \infty)]^{-1}$  — where  $E_0$  is the expectation with respect to  $P_0$  — we thus get that transience is equivalent  $E_0 N < \infty$ . To compute the expectation, we note

$$1_{\{X_n=0\}} = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} e^{ik \cdot X_n} \quad (3.7)$$

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

$$P_0(X_n = 0) = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \hat{J}(k)^n \quad (3.8)$$

Summing over  $n \geq 0$  yields

$$E_0 N = \sum_{n \geq 0} \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \hat{J}(k)^n = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} \quad (3.9)$$

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.  $\mathcal{E}$  Yukawa potentials:* As  $k \rightarrow 0$ ,

$$1 - \hat{J}(k) \sim C|k|^2 \quad (3.10)$$

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

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

$$1 - \hat{J}(k) \sim C \begin{cases} |k|^{s-d}, & \text{if } s < d + 2, \\ |k|^2 \log \frac{1}{|k|}, & \text{if } s = d + 2, \\ |k|^2, & \text{if } s > d + 2. \end{cases} \quad (3.11)$$

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 index of stability  $\alpha = s - d$ .) A convex combination of the three coupling constants will lead to a transient walk provided at least one of the interactions involved therein (with non-zero coefficients) is transient.

### 3.2 Infrared Bound

The principal claim of this section is that the finiteness of the integral in (3.6) is sufficient for the existence of a symmetry-breaking phase transition in many spin systems of the kind (3.1). The reason is the connection of the above random walk to the Gaussian free field (2.16) (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

$$\text{Cov}_\Lambda(\phi_x, \phi_y) = \sum_{n \geq 0} P_x(X_n = y, \tau_{\Lambda^c} = y) =: G_\Lambda(x, y) \quad (3.12)$$

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  $\Lambda$ . In particular, we have

$$\text{Var}_\Lambda(\phi_0) = G_\Lambda(0, 0) \quad (3.13)$$

which, as we will see, tends to the integral (3.6) as  $\Lambda \uparrow \mathbb{Z}^d$ . Since  $E_\Lambda(\phi_0) = 0$  due to our choice of the boundary condition, we conclude

$$\{\text{Law}(\phi_0): \Lambda \subset \mathbb{Z}^d\} \text{ is tight iff } (X_n) \text{ is transient} \quad (3.14)$$

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  $\phi_0$  can be interpreted as the survival of the symmetry breaking in 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 3.2.** *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 3.3 (Infrared Bound).** *Let  $L$  be an even integer and consider the model (3.1) on torus  $\mathbb{T}_L$  with Gibbs measure  $\mu_{L,\beta}$ . Suppose  $(J_{xy})$  is one of the three interactions above and let*

$$c_{L,\beta}(x) := E_{\mu_{L,\beta}}(S_0 \cdot S_x) \quad (3.15)$$

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

$$\hat{c}_{L,\beta}(k) \leq \frac{\nu}{2\beta} \frac{1}{1 - \hat{J}(k)}, \quad k \in \mathbb{T}_L^* \setminus \{0\} \quad (3.16)$$

where  $\nu$  is the dimension of the spin vectors and  $\mathbb{T}_L^*$  is the reciprocal torus,  $\mathbb{T}_L^* := \{\frac{2\pi}{L}(n_1, \dots, n_d) : n_i = 0, \dots, L-1\}$ .

The proof will require developing the technique of reflection positivity and is therefore postponed to Section 5.

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 spatial displacement of the two spins. The result has the following equivalent formulation: For all  $(v_x) \in \mathbb{C}^{\mathbb{T}_L}$  with  $\sum_x v_x = 0$ ,

$$\sum_{x,y \in \mathbb{T}_L} v_x \bar{v}_y E_{\mu_{L,\beta}}(S_0 \cdot S_x) \leq \frac{\nu}{2\beta} \sum_{x,y \in \mathbb{T}_L} v_x \bar{v}_y G_L(x,y) \quad (3.17)$$

where

$$G_L(x, y) := \frac{1}{L^d} \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \frac{e^{ik \cdot (x-y)}}{1 - \hat{J}(k)} \quad (3.18)$$

Observe that the latter is the covariance matrix of the GFF on  $\mathbb{T}_L$ , projected on the set of 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. (These differences are orthogonal to constant functions, of course.) A short formulation of the infrared bound is thus:

*The correlation of the spins in models (3.1) 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 a GFF.*

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

### 3.3 Spin-wave Condensation in $O(n)$ -model

Having temporarily dispensed with the IRB, we will continue in our original line of thought. Theorem 3.3 implies:

**Corollary 3.4 (Spin-wave Condensation).** *Suppose  $|S_x| = 1$ . Then*

$$E_{\mu_{L,\beta}} \left( \left| \frac{1}{L^d} \sum_{x \in \mathbb{T}_L} S_x \right|^2 \right) \geq 1 - \frac{\nu}{2\beta} G_L(0, 0) \quad (3.19)$$

*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

$$E_{\mu_{L,\beta}} |\hat{S}_k|^2 \leq \frac{\nu}{2\beta} \frac{L^d}{1 - \hat{J}(k)}, \quad k \in \mathbb{T}_L^* \setminus \{0\} \quad (3.20)$$

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

$$\sum_{k \in \mathbb{T}_L^*} |\hat{S}_k|^2 = L^d \sum_{x \in \mathbb{T}_L} |S_x|^2 = L^{2d} \quad (3.21)$$

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

$$\frac{1}{L^{2d}} |\hat{S}_0|^2 = 1 - \frac{1}{L^{2d}} \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} |\hat{S}_k|^2 \quad (3.22)$$

Now take expectation and apply (3.20):

$$E_{\mu_{L,\beta}} \left( \frac{1}{L^{2d}} |\hat{S}_0|^2 \right) \geq 1 - \frac{\nu}{2\beta} \frac{1}{L^d} \sum_{k \in \mathbb{T}_L \setminus \{0\}} \frac{1}{1 - \hat{J}(k)} \quad (3.23)$$

In light of (3.18), this is (3.19).  $\square$

With (3.19) 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 (3.6) and so the right-hand side has a finite limit. By taking  $\beta$  sufficiently large, the limit is actually strictly positive. This in turn implies that the zero mode of the spin-wave decomposition is *macroscopically* populated — very much like 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 3.5 (Phase Coexistence in  $O(n)$ -model).** *Consider the  $O(n)$ -model with  $n \geq 1$  and one of the three interactions above. Let*

$$\beta_0 := \frac{n}{2} \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{1}{1 - \hat{J}(k)} \quad (3.24)$$

*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*

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} S_x \xrightarrow{L \rightarrow \infty} m_\star \theta, \quad \mu_\theta\text{-a.s.} \quad (3.25)$$

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

Note that (3.25) implies that the measures  $\mu_\theta$  are mutually singular with respect to one another. Note also that  $\beta_0$  is finite — and the statement is not vacuous — if and only if the associated random walk is transient.

*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 (3.19) to show that the free energy is not differentiable in an appropriately-chosen external field when this field is set to zero. Then we apply Theorem 2.5 to conclude the existence of the required distinct, ergodic Gibbs measures.

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

$$f(h) := \lim_{L \rightarrow \infty} \frac{1}{L^d} \log E_{\mu_{L,\beta}} (e^{h\theta \cdot \hat{S}_0}) \quad (3.26)$$

(The limit exists by Theorem 2.5.) We want to show that  $\frac{\partial}{\partial h^+} f(0) > 0$  (and thus, by symmetry,  $\frac{\partial}{\partial h^-} f(0) < 0$ ). Corollary 3.4 yields

$$E_{\mu_{L,\beta}}(L^{-2d}|\hat{S}_0|^2) \geq \frac{\beta - \beta_0}{\beta} + o(1) \quad (3.27)$$

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

$$E_{\mu_{L,\beta}}(L^{-2d}|\hat{S}_0|^2) \leq \epsilon + \mu_{L,\beta}(|\hat{S}_0| \geq \epsilon L^d) \quad (3.28)$$

and so

$$\mu_{L,\beta}(|\hat{S}_0| \geq \frac{1}{2} \frac{\beta - \beta_0}{\beta} L^d) \geq \frac{1}{2} \frac{\beta - \beta_0}{\beta} + o(1) \quad (3.29)$$

By the  $O(n)$  symmetry of the torus measures  $\mu_{L,\beta}$ , the law of  $\hat{S}_0/L^d$  is rotationally invariant with non-degenerate ‘‘radius’’ distribution. This implies

$$\mu_{L,\beta}(\theta \cdot \hat{S}_0 \geq \frac{1}{4} \frac{\beta - \beta_0}{\beta} L^d) \geq C_n \frac{\beta - \beta_0}{\beta} + o(1) \quad (3.30)$$

where  $C_2 := 1/6$  and, in general,  $C_n > 0$  is an explicitly obtainable constant. 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

$$\frac{\partial f}{\partial h^+} \Big|_{h=0} \geq \frac{\beta - \beta_0}{4\beta} \quad (3.31)$$

Applying Theorem 2.5, 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

$$E_{\mu_\theta}(\theta \cdot S_x) = \frac{\partial f}{\partial h^+} \Big|_{h=0} > 0 \quad (3.32)$$

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

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} S_x \xrightarrow{L \rightarrow \infty} m_\star \tilde{\theta}, \quad \mu_\theta\text{-a.s.} \quad (3.33)$$

where  $\tilde{\theta} \in \mathbb{S}^{n-1}$  and where  $m_\star > 0$  is the magnitude of the derivative. Note that, in light of (3.32) and the translation invariance of  $\mu_\theta$ ,

$$m_\star \theta \cdot \tilde{\theta} = \frac{\partial f}{\partial h^+} \Big|_{h=0} \quad (3.34)$$



The distinctness of  $\mu_\theta$  will follow once we prove (3.25), i.e.,  $\theta = \tilde{\theta}$ . (This is, of course, intuitively obvious because the way we constructed  $\mu_\theta$  indicates that the law of  $S_x$  under  $\mu_\theta$  should be biased in the direction of  $\theta$ .)

Suppose  $\tilde{\theta} \neq \theta$ . Find a rotation  $A \in O(n)$  such that  $A\tilde{\theta} = \theta$ . Let  $\tilde{\mu}$  be the measure such that  $E_{\tilde{\mu}}(f(S)) := E_{\mu_\theta}(f(AS))$  for all local functions  $f$ . (The existence of such a measure follows from the Kolmogorov Extension Theorem.) Since both the Hamiltonian and the *a priori* measure are  $O(n)$ -invariant, we have  $\tilde{\mu} \in \mathfrak{G}_\beta$ . But (3.33) implies  $E_{\mu_\theta}(S_x) = m_\star \tilde{\theta}$ , and so from (3.34) we have

$$E_{\tilde{\mu}}(\theta \cdot S_x) = E_{\mu_\theta}(\theta \cdot AS_x) = m_\star |\theta|^2 \underset{\tilde{\theta} \neq \theta}{>} m_\star \theta \cdot \tilde{\theta} = \left. \frac{\partial f}{\partial h^+} \right|_{h=0} \quad (3.35)$$

As  $\tilde{\mu}$  is a Gibbs measure, this contradicts the general bounds in Theorem 2.5. Hence, we must have  $\theta = \tilde{\theta}$  after all.  $\square$

The above statement and proof are formulated for the specific case of the  $O(n)$  model. A similar proof will apply the existence of a symmetry-breaking phase transition at low temperatures in the Ising, Potts and the liquid-crystal models in all transient dimensions. As the Ising and Potts model have only a *discrete* set of spin states, a symmetry-breaking transition will occur generally in all dimensions  $d \geq 2$ . However, this has to be proved by different methods than those employed above (e.g., by invoking chessboard estimates).

Our next goal is to establish the complementary part of Theorem 3.2, i.e., the implication  $\implies$ , which asserts the *absence* of symmetry breaking in the recurrent cases. This argument predates the other direction by 20 years and bears the name of its discoverers:

**Theorem 3.6 (Mermin-Wagner Theorem).** *Let  $n \geq 2$  and consider the  $O(n)$ -model with non-negative interactions constants  $(J_{x,y})$  satisfying the conditions (I1,I2) from Sect. 3.1. Suppose the corresponding random walk is recurrent. Then every  $\mu \in \mathfrak{G}_\beta$  is invariant under any simultaneous (i.e., homogeneous) rotation of all spins.*

*Proof.* We will show that the spins can be arbitrarily rotated at an arbitrary small cost of the total energy. (This is why we need  $n \geq 2$ .) 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

$$\omega_\varphi(S_x) := e^{i\varphi_x R} S_x, \quad x \in \mathbb{Z}^d \quad (3.36)$$

To investigate the effect of such an inhomogeneous rotation on the Hamiltonian, note that

$$\begin{aligned} \omega_\varphi(S_x) \cdot \omega_\varphi(S_y) &= S_x \cdot e^{i(\varphi_y - \varphi_x)R} S_y \\ &= S_x \cdot S_y - S_x \cdot [1 - e^{i(\varphi_y - \varphi_x)R}] S_y \end{aligned} \quad (3.37)$$

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

$$H_\Lambda(\omega_\varphi(S)) = H_\Lambda(S) + \Delta H \quad (3.38)$$

where

$$\Delta H := \frac{1}{2} \sum_{x,y} J_{xy} S_x \cdot [1 - e^{i(\varphi_y - \varphi_x)R}] S_y \quad (3.39)$$

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  $f$ ,

$$E_\mu(f \circ \omega_\varphi) = E_\mu(f e^{-\beta \Delta H}) \quad (3.40)$$

We will now let  $\varphi_x \rightarrow \alpha$  in a specific way that ensures  $\Delta H \rightarrow 0$ ; this will permit us to extract the desired conclusion by limiting arguments.

First we will need to control the  $\varphi$ -dependence of  $\Delta H$ , so we expand the exponential:

$$\begin{aligned} \Delta H &= -\frac{i}{2} \sum_{x,y} J_{xy} (S_x \cdot R S_y)(\varphi_y - \varphi_x) \\ &\quad + \frac{1}{4} \sum_{x,y} J_{xy} (R S_x \cdot R S_y)(\varphi_y - \varphi_x)^2 + \dots \end{aligned} \quad (3.41)$$

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  $J_{xy} (S_x \cdot R S_y)$  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

$$|\Delta H| \leq C \sum_{x,y} J_{xy} (\varphi_y - \varphi_x)^2 = 2C \mathcal{E}_{1-J}(\varphi, \varphi) \quad (3.42)$$

for some constant  $C < \infty$ . Here we used that  $(R S_x \cdot R S_y)$  is bounded and recalled the definition of the Dirichlet form  $\mathcal{E}_{1-J}(\cdot, \cdot)$  of the random walk driven by the  $(J_{x,y})$ 's.

Our next task will be to control the Dirichlet form under the condition that  $\varphi$  tends to  $\alpha$  in every finite set. To that end we fix  $0 < R < \infty$  and set

$$\varphi_x := \alpha \mathbf{P}_x(\tau_0 < \tau_{\Lambda_R^c}) \quad (3.43)$$

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

$$\begin{aligned} \mathcal{E}_{1-J}(\varphi, \varphi) &= \sum_x \varphi_x \sum_y J_{x,y}(\varphi_x - \varphi_y) \\ &= \underset{\substack{\text{harmonic or} \\ \text{zero in } \{0\}^c}}{\alpha} \sum_y J_{0,y}(\alpha - \varphi_y) \end{aligned} \quad (3.44)$$

But the recurrence of the associated random walk 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. Thus  $\Delta H \rightarrow 0$  as  $R \rightarrow \infty$  and so applying  $R \rightarrow \infty$  to (3.40) with the choice (3.43) yields

$$E_\mu(f \circ \omega_\alpha) = E_\mu(f) \quad (3.45)$$

for every continuous local function  $f$ . Thereby we conclude that  $\mu$  is invariant under simultaneous rotation of all spins.  $\square$

### 3.4 Literature Remarks

The content of the entire section is very classical. The Infrared Bound (and its proof based on reflection positivity) was discovered in the seminal work of Fröhlich, Simon and Spencer [50] from 1976 where it was also applied to prove a phase transition in the  $O(n)$ -model (as well as the isotropic Heisenberg and other models). Dyson, Lieb and Simon [38] showed how to adapt the method to a (somewhat more limited) class of quantum spin models. The technique was further developed and its applications extended in two papers of Fröhlich, Israel, Lieb and Simon [46, 47].

Thanks to the representation (2.12), the proof of a long-range order in the liquid-crystal model, derived by Angelescu and Zagrebnov [6], follows the same route as for the  $O(n)$  model. However, the type of long-range order that is concluded for the actual spin system is different. Indeed, let  $\mu$  be a (weak) cluster point of the torus states. Then

$$\lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} [(S_0 \cdot S_x)^2 - 1/n] > 0 \quad (3.46)$$

with a positive probability under  $\mu$ . (The limit exists by the Pointwise Ergodic Theorem.) As  $\mu$  is  $O(n)$ -invariant, if  $S_x$  were asymptotically independent of  $S_0$  for large  $x$ , we would expect  $E_\mu(S_0 \cdot S_x)^2 \rightarrow 1/n$  as  $|x| \rightarrow \infty$ . Apparently, this is not the case, the direction of  $S_x$  remains heavily correlated with the direction of  $S_0$  for arbitrary  $x$ , i.e., there is an *orientational* long range order.

Whether or not the  $O(n)$  symmetry of the law of  $S_x$  is broken is an open (and important) question. (The law of each individual  $S_x$  is invariant under the flip  $S_x \leftrightarrow -S_x$  and so the magnetization is zero in all states.) As noted before, other models of liquid crystals based on dimers on  $\mathbb{Z}^2$  were considered by Heilmann and Lieb [62] and Abraham and Heilmann [1] prior to the work [6]. There an orientational long-range order was proved using chessboard estimates; the question of absence of *complete translational ordering* (i.e., breakdown of translation invariance) remained open.

The Mermin-Wagner theorem goes back to 1966 [82]. Various interesting mathematical treatments and extensions followed [34, 87, 49]; the argument presented here is inspired by the exposition in Simon's book [97]. A fully probabilistic approach to this result, discovered by Dobrushin and Shlosman [34], 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 [65]. 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 [103].

The basis of the Mermin-Wagner theorem, as well as its extension, is the *continuum* nature of the spin space. Indeed, in the Ising (and also Potts) model, a low-temperature symmetry breaking occurs even in some recurrent dimensions; e.g., in  $d = 2$  for the nearest-neighbor interactions. For what determines the presence and absence of symmetry breaking in  $d = 1$ , see the work of Aizenman, Chayes, Chayes and Newman [3] and references therein.

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 section hold without reference to random walks. For detailed expositions of the theory of random walks we recommend the monographs by Spitzer [100] and Lawler [76]; the material naturally appears in most graduate probability textbooks (e.g., Durrett [36]).

It is interesting to note that even in  $d = 2$ , the nearest-neighbor  $O(n)$  model exhibits a phase transition when  $n = 2$ . Namely, while

the Gibbs state is unique at all  $\beta < \infty$ , for large  $\beta$  it exhibits power law decay of correlations with  $\beta$ -dependent exponents. This regime is (again, after its discoverers) referred to as the *Kosterlitz-Thouless phase* [70]. A rigorous treatment exists, based on renormalization theory and connection with Coulomb gas, thanks to the pioneering work of Fröhlich and Spencer [51]; see also more recent papers by Dimock and Hurd [31]. This is of much interest in light of recent discovery of new conformally-invariant planar processes — the Schramm-Loewner evolution (a.k.a. SLE). No such phenomenon is expected when  $n \geq 3$  though there is a minor opposition to this (e.g., Patrasciou and Seiler [85]).

## 4 Infrared Bound & Mean-field Theory

In this chapter we will discuss how the infrared bound can be used to control the error in so-called mean-field approximation. Unlike the spin-wave condensation, which is concerned primarily with 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 (3.6) is still a prerequisite.)

### 4.1 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 a lattice spin model with the usual Hamiltonian (3.1). 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 all spins outside the origin. Indeed, the one-spin Gibbs measure is determined by the (one-spin) Hamiltonian

$$H_{\{0\}}(S) = - \sum_x J_{0,x} S_0 \cdot S_x = -S_0 \cdot \sum_x J_{0,x} S_x = -S_0 \cdot M_0 \quad (4.1)$$

where we introduced the shorthand

$$M_0 := \sum_x J_{0,x} S_x \quad (4.2)$$

Thus, by the DLR,

$$E_\mu(S_0) = E_\mu\left(\frac{E_{\mu_0}(S_0 e^{\beta S_0 \cdot M_0})}{E_{\mu_0}(e^{\beta S_0 \cdot M_0})}\right) \quad (4.3)$$

where, abusing the notation slightly, the “inner” expectations are only over  $S_0$  —  $M_0$  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

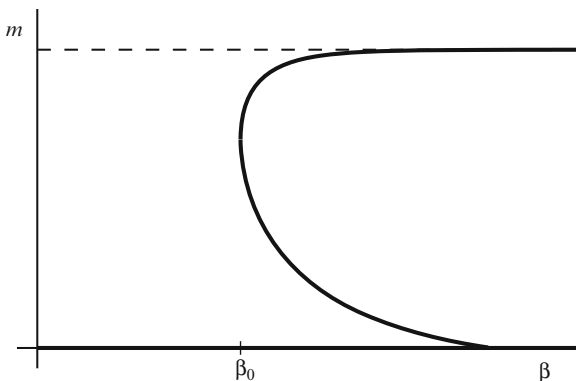
$$m := E_\mu(S_0) \quad (4.4)$$

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

$$m = \frac{E_{\mu_0}(S e^{\beta S \cdot m})}{E_{\mu_0}(e^{\beta S \cdot m})} \quad (4.5)$$

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

Besides the unjustified step in the derivation, a serious practical problem with (4.5) is that it often has multiple solutions. Indeed, for the set of points  $(\beta, m)$  that obey this equation, one typically gets a picture like this:



Here, for  $\beta < \beta_0$ , the only solution is  $m = 0$  — this is always a solution whenever  $E_{\mu_0}(S) = 0$  — but at  $\beta = \beta_0$ , two new branches appear and coexist over an interval of  $\beta$ 's. 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 (4.5) alone. For that one has to go beyond the heuristic derivation presented above.

As is standard, one comes up with an additional “selection” principle that determines which solution is “physical.” At the level of classical thermodynamics, this is done by postulating that the solution must minimize an appropriate free energy function. In the choice of this function we will be guided by the fact that there is a proper statistical-mechanical system for which the above derivations can be explicitly validated by way of large-deviation theory. This system is the corresponding model on the *complete graph*.

Consider a graph on  $N$  vertices with each pair of vertices joined by an undirected edge. At each vertex  $x = 1, \dots, N$  we have a spin  $S_x$  with i.i.d. *a priori* law  $\mu_0$ . Each spin interacts with every other spin; the interaction Hamiltonian is given by

$$H_N(S) := -\frac{1}{2N} \sum_{x,y=1}^N S_x \cdot S_y \quad (4.6)$$

The normalization by  $1/N$  ensures that the energy grows proportionally to  $N$ ; the “2” in the denominator compensates for counting each pair of spins twice.

To derive the formula for the free energy function, consider first the cumulant generating function of the measure  $\mu_0$ ,

$$G(h) := \log E_{\mu_0}(e^{h \cdot S}), \quad h \in \mathbb{R}^\nu \quad (4.7)$$

Its Legendre transform,

$$\mathcal{S}(m) := \inf_{h \in \mathbb{R}^\nu} [G(h) - h \cdot m] \quad (4.8)$$

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

$$\mu_0 \left( \sum_{x=1}^N S_x \approx mN \right) = e^{-N\mathcal{S}(m) + o(N)} \quad (4.9)$$

(The function is infinite outside  $\text{Conv}(\Omega)$ , the convex hull of  $\Omega$  and the set of possible values of the magnetization.) 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

$$\mu_0 \left( e^{\frac{\beta}{2N} \sum_{x,y=1}^N S_x \cdot S_y} \mathbf{1}_{\{\sum_x S_x \approx mN\}} \right) = e^{-N\Phi_\beta(m) + o(N)} \quad (4.10)$$

Here the rate function

$$\Phi_\beta(m) := -\frac{\beta}{2}|m|^2 - \mathcal{S}(m) \quad (4.11)$$

is the desired *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 (4.5):

**Lemma 4.1.** *We have*

$$\nabla\Phi_\beta(m) = 0 \quad \Leftrightarrow \quad m = \nabla G(\beta m) \quad (4.12)$$

*Explicitly, the solutions to (4.5) 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

$$\nabla\Phi_\beta(m) = 0 \quad \Leftrightarrow \quad \beta m = h_m \quad \Leftrightarrow \quad m = \nabla G(\beta m) \quad (4.13)$$

It remains to observe that  $m = \nabla G(\beta m)$  is a concise way to write (4.5).  $\square$

Lemma 4.1 shows that the appearance of multiple solutions to (4.5) coincides with the emergence of secondary local maxima/minima.

## 4.2 Example: The Potts Model

It is worthwhile to demonstrate the above general formalism on the explicit example of the Potts model. We will work with the tetrahedral representation, i.e., on the spin space  $\Omega := \{\hat{v}_1, \dots, \hat{v}_q\}$ . The mean-field free energy function is best expressed in the parametrization using the *mole fractions*,  $x_1, \dots, x_q$ , which on the complete graph represent the fractions of all vertices with spins pointing in the directions  $\hat{v}_1, \dots, \hat{v}_q$ , respectively. Clearly,

$$\sum_{i=1}^q x_i = 1 \quad (4.14)$$

The corresponding magnetization vector is

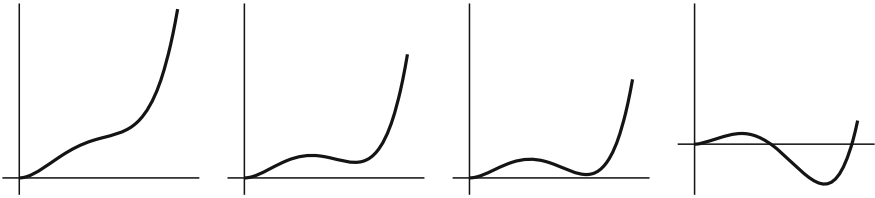
$$m = x_1 \hat{v}_1 + \dots + x_q \hat{v}_q. \quad (4.15)$$



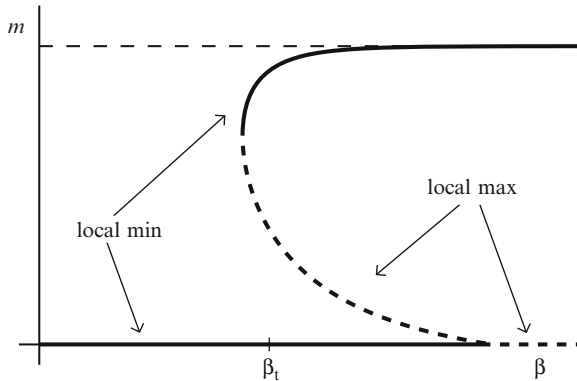
In this notation we have

$$\Phi_\beta(m) = \sum_{k=1}^q \left( -\frac{\beta}{2} x_k^2 + x_k \log x_k \right). \quad (4.16)$$

It is not surprising, but somewhat non-trivial to prove (see [9, Lemma 4.4]) that all interesting behavior of  $\Phi_\beta$  occurs “on-axes” that is, the absolute minimizers — and, in fact, all local extrema — of  $\Phi_\beta$  occur in the directions of one of the spin states. (Which direction we choose is immaterial as they are related by symmetry.) The following picture shows the qualitative look of the function  $\mathbf{m} \mapsto \Phi_\beta(\mathbf{m}\hat{v}_1)$  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 point  $\beta_0$  where the secondary minima/maxima first appear.

### 4.3 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 (3.1) and the coupling constants being one of the 3 types above.

**Definition 4.2.** *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 4.3.** *Suppose  $|S_x| \leq 1$ . Let  $\mu \in \mathfrak{G}_\beta$  be a translation-invariant, ergodic, torus state and define*

$$m_\star := E_\mu(S_0). \quad (4.17)$$

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

$$\Phi_\beta(m_\star) \leq \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + \frac{\nu\beta}{2} \mathcal{I}_d \quad (4.18)$$

where

$$\mathcal{I}_d := \int_{[-\pi,\pi]^d} \frac{dk}{(2\pi)^d} \frac{\hat{J}(k)^2}{1 - \hat{J}(k)} \quad (4.19)$$

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. Thus, in strongly transient situations one should expect that  $\mathcal{I}_d$  is fairly small. And, indeed, we have the following asymptotics:

- *n.n. interactions:*

$$\mathcal{I}_d \sim \frac{1}{2d}, \quad d \rightarrow \infty. \quad (4.20)$$

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

$$\mathcal{I}_d \leq C\mu^d. \quad (4.21)$$

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

$$\mathcal{I}_d \leq C(s - d). \quad (4.22)$$

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 4.3 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 4.4.** *Let  $q \geq 3$  and suppose that  $\mathcal{I}_d \ll 1/q$ . Then there is  $\beta_t \in (0, \infty)$  and translation-invariant, ergodic measures  $\nu_0, \nu_1, \dots, \nu_q \in \mathfrak{G}_{\beta_t}$  such that*

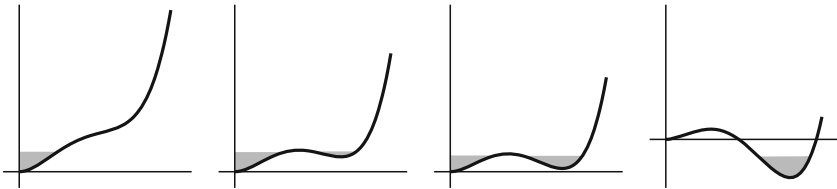
$$|E_{\nu_0}(S_x)| \ll 1 \quad (4.23)$$

and

$$E_{\nu_j}(S_x) = m_\star \hat{\nu}_j, \quad j = 1, \dots, q, \quad (4.24)$$

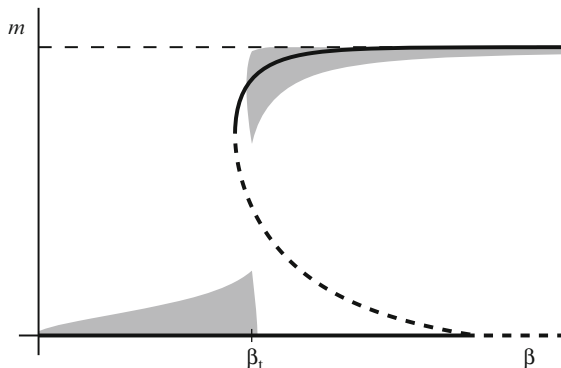
where  $m_\star \geq 1/2$ . In particular, the 3-state Potts model undergoes a first-order phase transition provided the spatial dimension is sufficiently large.

This result is pretty much the consequence of the pictures in Sect. 4.1. Indeed, including the error bound (4.18), 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 once we mark directly into the mean-field magnetization plot the set of values of the magnetization allowed by the inequality (4.18):



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

An additional argument is actually needed to provide a full proof of (4.24). Indeed, we claim that the symmetry breaking happens *exactly* in the direction of one of the spin states while the approximation by mean-field theory only guarantees that the expectation is *near* one of these directions.

*Proof of (4.24), sketch.* Consider an ergodic Gibbs state  $\mu$  with  $m_\star := E_\mu(S_x) \neq 0$  at inverse temperature  $\beta$ . Given a sample  $\sigma = (\sigma_x)$  from  $\mu$ , at each *unordered* pair  $\langle x, y \rangle$  of vertices from  $\mathbb{Z}^d$  let

$$\eta_{xy} := 1_{\{\sigma_x = \sigma_y\}} Z_{xy} \quad (4.25)$$

where  $(Z_{xy})$  are *a priori* independent, zero-one valued random variables with

$$\mathbb{P}(Z_{xy} = 1) = 1 - \mathbb{P}(Z_{xy} = 0) := 1 - e^{-\beta J_{xy}} \quad (4.26)$$

This defines a coupling of  $\mu$  with a *random cluster measure* — the distribution of the  $\eta$ 's — which, by the fact that the extension comes from i.i.d. random variables, is also ergodic.

When  $m_\star \neq 0$ , the  $\eta$ -marginal features a unique infinite connected component of edges  $\langle x, y \rangle$  with  $\eta_{xy} = 1$  whose (site) density is proportional to  $|m_\star|$ . By the construction, the spin variables take a (constant) value on each connected component, which is a.s. unique (by ergodicity) on the infinite one and uniform on the finite ones. Thus, the bias

of the spin distribution comes only from the infinite component and so it points in one of the  $q$  spin directions. The claim thus follows.  $\square$

#### 4.4 Ideas from the Proofs

A fundamental technical ingredient of the proof is again provided by the IRB, so throughout we will assume one of the three interactions discussed above. However, we will need the following enhanced version:

**Lemma 4.5 (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 us denote  $m_\star := E_\mu(S_0)$ . Then for all  $(v_x)_{x \in \mathbb{Z}^d} \in \mathbb{C}^{\mathbb{Z}^d}$  with finite support,*

$$\sum_{x,y} v_x \bar{v}_y E_\mu((S_x - m_\star) \cdot (S_y - m_\star)) \leq \frac{\nu}{2\beta} \sum_{x,y} v_x \bar{v}_y G(x, y). \quad (4.27)$$

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

$$\sum_{x,y} w_x \bar{w}_y E_\mu(S_x \cdot S_y) \leq \frac{\nu}{2\beta} \sum_{x,y} w_x \bar{w}_y G(x, y) \quad (4.28)$$

where

$$G(x, y) := \lim_{L \rightarrow \infty} G_L(x, y) = \int_{[-\pi, \pi]^d} \frac{dk}{(2\pi)^d} \frac{e^{ik \cdot (x-y)}}{1 - \hat{J}(k)} \quad (4.29)$$

What separates (4.28) from (4.27) 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  $(v_x)$  has finite support but let now  $\sum_x v_x$  be arbitrary. Let  $\Lambda_L \subset \mathbb{Z}^d$  contain the support of  $(v_x)$ . To convert to the previous argument, let

$$a_L := \frac{1}{|\Lambda_L|} \sum_x v_x \quad (4.30)$$

and

$$w_x := v_x - a_L 1_{\Lambda_L}(x) \quad (4.31)$$

Note that  $\sum_x w_x = 0$ . Then

$$\begin{aligned} \sum_{x,y} w_x \bar{w}_y E_\mu((S_x - m_\star) \cdot (S_y - m_\star)) &= \sum_{x,y} v_x \bar{v}_y E_\mu((S_x - m_\star) \cdot (S_y - m_\star)) \\ &\quad - 2E_\mu\left(\left[a_L \sum_{x \in \Lambda_L} (S_x - m_\star)\right] \cdot \left[\sum_y v_y (S_y - m_\star)\right]\right) \\ &\quad + E_\mu\left(\left|a_L \sum_{x \in \Lambda_L} (S_x - m_\star)\right|^2\right) \end{aligned} \quad (4.32)$$

But ergodicity of  $\mu$  implies that

$$E_\mu\left(\left|\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} (S_x - m_\star)\right|^2\right) \xrightarrow{L \rightarrow \infty} 0 \quad (4.33)$$

and so, by Cauchy-Schwarz, the last two terms in (4.32) converge to zero as  $L \rightarrow \infty$ . Now apply (4.28) and pass to the limit  $L \rightarrow \infty$  there. A direct calculation (and the Riemann-Lebesgue lemma) shows that

$$\frac{1}{|\Lambda_L|} \sum_{x \in \Lambda_L} G(x, y) \xrightarrow{L \rightarrow \infty} 0 \quad (4.34)$$

and so the terms involving  $a_L$  on the right-hand side of (4.27) suffer a similar fate. This means that the left-hand sides of (4.27–4.28) tend to each other, and same for the right-hand sides. The desired bound (4.27) is thus a limiting version of (4.28).  $\square$

Clearly, the restriction to finitely-supported  $(v_x)$  is not necessary; instead, one can consider completions of this set in various reasonable norms. The above formulation has an immediate, but rather fundamental, consequence:

**Corollary 4.6 (Key Estimate).** *Let  $\mu \in \mathfrak{G}_\beta$  be an ergodic torus state and let  $m_\star := E_\mu(S_x)$ . Then we have*

$$E_\mu\left(\left|\sum_x J_{0,x} S_x - m_\star\right|^2\right) \leq \frac{\nu}{2\beta} \mathcal{I}_d. \quad (4.35)$$

*Proof.* Choose  $v_x := J_{0,x}$  and note that with this choice the left-hand side of (4.27) becomes the left-hand side of (4.35). As to the right-hand side of (4.27), we get

$$\frac{\nu}{2\beta} \sum_{x,y} \int_{[-\pi,\pi]^d} \frac{dk}{(2\pi)^d} \frac{e^{ik \cdot (x-y)}}{1 - \hat{J}(k)} J_{0,x} J_{0,y} \quad (4.36)$$

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

This corollary provides a 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 of Theorem 4.3 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. 3.1. The relevant observations are as follows:

**Proposition 4.7.** *Let  $\mu \in \mathfrak{G}_\beta$  be translation invariant and let  $m_\star := E_\mu(S_x)$ .*

(1) *We have*

$$\Phi_\beta(m_\star) \leq \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + \frac{\beta}{2} \sum_{x \in \mathbb{Z}^d} J_{0,x} [E_\mu(S_0 \cdot S_x) - |m_\star|^2] \quad (4.37)$$

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

$$\sum_{x \in \mathbb{Z}^d} J_{0,x} [E_\mu(S_0 \cdot S_x) - |m_\star|^2] \leq \beta E_\mu \left( \left| \sum_x J_{0,x} S_x - m_\star \right|^2 \right) \quad (4.38)$$

*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

$$Z_\Lambda \geq \exp \left\{ -|\Lambda| \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + O(\partial\Lambda) \right\}. \quad (4.39)$$

To prove this we pick  $m$  in the (relative) interior of  $\text{Conv}(\Omega)$  and define a *tilted* measure

$$\mu_h(dS) := e^{h \cdot S - G(h)} \mu_0(dS) \quad (4.40)$$

with  $h$  adjusted so that  $E_{\mu_h}(S) = m$ . (Such  $h$  exists for each  $m$  in the relative interior of  $\text{Conv}(\Omega)$ , by standard arguments for the Legendre transform.) We then get

$$Z_\Lambda = E_{\otimes \mu_h} \left( e^{-\beta H_\Lambda(S) - h \cdot M_\Lambda + |\Lambda| G(h)} \right) \quad (4.41)$$

where we introduced the shorthand

$$M_\Lambda := \sum_{x \in \Lambda} S_x \quad (4.42)$$

Now apply Jensen to get the expectation into the exponent; the product nature of  $\otimes\mu_h$  implies that  $E_{\otimes\mu_h}(H_\Lambda(S)) = -|\Lambda|\frac{1}{2}|m|^2 + O(\partial\Lambda)$  and so (4.39) follows by noting that  $G(h) - h \cdot m = \mathcal{S}(m)$  due to our choice of  $h$ , and subsequently optimizing over all admissible  $m$ .

Now fix a general  $h \in \mathbb{R}^\nu$  and let  $\mu$  be a Gibbs measure as specified in the claim. First we note that the DLR condition implies

$$E_\mu(e^{+\beta H_\Lambda + h \cdot M_\Lambda} Z_\Lambda) = E_{\otimes\mu_0}(e^{h \cdot M_\Lambda}) = e^{|\Lambda|G(h)} \quad (4.43)$$

The  $Z_\Lambda$  term can be bounded away via (4.39); Jensen's inequality then gives

$$\beta E_\mu(H_\Lambda) + |\Lambda| h \cdot m_\star - |\Lambda| \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) + O(\partial\Lambda) \leq |\Lambda|G(h) \quad (4.44)$$

Next, translation invariance of  $\mu$  yields

$$E_\mu(H_\Lambda) = -|\Lambda| \frac{1}{2} \sum_x J_{0,x} E_\mu(S_0 \cdot S_x) + O(\partial\Lambda) \quad (4.45)$$

and so dividing by  $\Lambda$  and taking  $\Lambda \uparrow \mathbb{Z}^d$  along cubes gets us

$$-\frac{\beta}{2} \sum_x J_{0,x} E_\mu(S_0 \cdot S_x) - \inf_{m \in \text{Conv}(\Omega)} \Phi_\beta(m) \leq G(h) - h \cdot m_\star \quad (4.46)$$

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 (4.11) now proves the claim.  $\square$

*Proof of (2).* Let us return to the notation  $M_0 := \sum_x J_{0,x} S_x$ . The left-hand side of (4.38) can then 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

$$E_\mu(M_0 \cdot S_0) = E_\mu(M_0 \cdot \nabla G(\beta M_0)) \quad (4.47)$$

The DLR condition also implies

$$m_\star = E_\mu(M_0) = E_\mu[\nabla G(\beta M_0)] \quad (4.48)$$

and so we have

$$\begin{aligned} E_\mu(S_0 \cdot M_0) - |m_\star|^2 \\ = E_\mu\left((M_0 - m_\star) \cdot (\nabla G(\beta M_0) - \nabla G(\beta m_\star))\right) \end{aligned} \quad (4.49)$$

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



$$(M_0 - m_\star) \cdot (\nabla G(\beta M_0) - \nabla G(\beta m_\star)) \leq \beta |M_0 - m_\star|^2 \quad (4.50)$$

by the Mean-Value Theorem. Taking expectations proves (4.38).  $\square$

Theorem 4.3 now follows by combining Proposition 4.7 with Corollary 4.6. Interestingly, (4.37) gives

$$\sum_{x \in \mathbb{Z}^d} J_{0,x} E_\mu(S_0 \cdot S_x) \geq |m_\star|^2 \quad (4.51)$$

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

## 4.5 Literature Remarks

The inception of mean-field theory goes back to Curie [28] and Weiss [105]. One of the early connections to the models on the complete graph appears in Ellis' textbook on large-deviation theory [40]. Most of this section is based on the papers of Biskup and Chayes [9] and Biskup, Chayes and Crawford [10]. The Key Estimate had been used before in some specific cases; e.g., for the Ising model in the paper by Bricmont, Kesten, Lebowitz and Schonmann [22] and for the  $q$ -state Potts model in the paper by Kesten and Schonmann [69]. 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 [69] 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 4.4 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 [19] — but this is so far known only in  $d = 2$  (or for  $q$  very large). The coupling in the proof of (4.24) is due to Edwards and Sokal [39]; for further properties see Grimmett [61] or Biskup, Borgs, Chayes and Kotecky [8]. The uniqueness of the infinite connected component is well known in the nearest-neighbor case from a beautiful argument of Burton and Keane [23]; for the long-range models it has to be supplied by a percolation bound dominating the number of edges connecting a box of side  $L$  to its complement.

The requirement  $\mathcal{I}_d \ll 1/q$  is actually an embarrassment of the theory as the transition should become more pronounced, and thus easier to control, with increasing  $q$ . Thus, even for nearest-neighbor case, we still do not have a dimension in which all  $q \geq 3$  state Potts models go first order. (It is expected that this happens already in  $d = 3$ .) The restriction to transient dimensions is actually not absolutely necessary; cf recent work Chayes [24].

It is natural to ask whether one can say anything about the continuum- $q$  extension of the Potts model, the random cluster model; see Grimmett [61]. Unfortunately, the main condition for proving the IRB, reflection positivity, holds if *and only if*  $q$  is integer (Biskup [7]).

Another model for which this method yields a novel result is the liquid-crystal model discussed in Sect. 2.2. Here Angelescu and Zagrebnov [6] 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 remarks at the end of Chapter 3. In [9] it has been shown that, for  $n \geq 3$ , the order parameter undergoes a discontinuous transition at intermediate temperatures; van Enter and Shlosman [41, 42] later proved such transitions in highly non-linear cases. Similar “mean-field driven” first order phase transitions have also been proved for the cubic model [9] and the Blume-Capel model [10].

Once the general theory is in place, the proof of a phase transition for a specific model boils down to the analysis of the mean-field free energy function. While in principle always doable, in practice this may be quite a challenge even in some relatively simple examples. See, e.g., [9, Sect. 4.4] what this requires in the context of the liquid-crystal model.

Finally, we note that the IRB has been connected to mean-field theory before; namely, in the work of Aizenman [2] (cf also Fröhlich [45] and Sokal [98]) in the context of lattice field theories and that of Aizenman and Fernández in the context of Ising systems in either high spatial dimensions [4] or for spread-out interactions [5]. A representative result from these papers is that *the critical exponents in the Ising model take mean-field values 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 extension was obtained recently by Sakai [90] who proved the IRB — and the corresponding conclusions about the critical exponents — directly via a version of the lace expansion.

## 5 Reflection Positivity

In the last two sections 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 estimate which we will use extensively in Chapter 6.

### 5.1 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

$$\vartheta(S_x) := S_{\vartheta(x)} \quad (5.1)$$

Clearly,  $\vartheta$  is a morphism of algebra  $\mathfrak{A}^+$  onto  $\mathfrak{A}^-$  and  $\vartheta^2 = \text{id}$ .

**Definition 5.1 (Reflection Positivity).** *A measure  $\mu$  on  $\Omega^{\mathbb{T}_L}$  is reflection positive (RP) with respect to  $\vartheta$  if*

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

$$E_\mu(f \vartheta g) = E_\mu(g \vartheta f) \quad (5.2)$$

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

$$E_\mu(f \vartheta f) \geq 0 \quad (5.3)$$

Note that the above implies that  $f, g \mapsto E_\mu(f \vartheta g)$  is a positive-semidefinite symmetric bilinear form. Condition (5.2) is usually automatically true — it requires only  $\vartheta$ -invariance of  $\mu$  — so it is the second condition that makes this concept non-trivial (hence also the name). Here we first note that the concept is not entirely vacuous:

**Lemma 5.2.** *The product measure,  $\mu = \bigotimes \mu_0$ , 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,

$$E_\mu(f \vartheta g) = E_\mu(f)E_\mu(\vartheta g) = E_\mu(f)E_\mu(g) \quad (5.4)$$

whereby both conditions in Definition 5.1 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

$$E_\mu(f \vartheta g|S_P) = E_\mu(f|S_P)E_\mu(\vartheta g|S_P) = E_\mu(f|S_P)E_\mu(g|S_P) \quad (5.5)$$

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

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

$$[E_\mu(f \vartheta g)]^2 \leq E_\mu(f \vartheta f)E_\mu(g \vartheta g) \quad (5.6)$$

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

**Lemma 5.3.** *Let  $\mu$  be RP with respect to  $\vartheta$  and let  $A, B, C_\alpha, D_\alpha \in \mathfrak{A}^+$ . Then*

$$\begin{aligned} & [E_\mu(e^{A+\vartheta B+\sum_\alpha C_\alpha \vartheta D_\alpha})]^2 \\ & \leq [E_\mu(e^{A+\vartheta A+\sum_\alpha C_\alpha \vartheta C_\alpha})] [E_\mu(e^{B+\vartheta B+\sum_\alpha D_\alpha \vartheta D_\alpha})] \end{aligned} \quad (5.7)$$

*Proof.* Clearly, in the absence of the  $C_\alpha \vartheta D_\alpha$  terms, this simply reduces to (5.6). To include these terms we use expansion into Taylor series:

$$\begin{aligned} & E_\mu(e^{A+\vartheta B+\sum_\alpha C_\alpha \vartheta D_\alpha}) \\ & = \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_\mu((e^A C_{\alpha_1} \dots C_{\alpha_n}) \vartheta (e^B D_{\alpha_1} \dots D_{\alpha_n})) \end{aligned} \quad (5.8)$$

Now we apply (5.6) to the expectation on the right-hand side and then one more time to the resulting sum:

$$\begin{aligned}
& E_\mu(e^{A+\vartheta B+\sum_\alpha C_\alpha \vartheta D_\alpha}) \\
& \leq \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} \left[ E_\mu((e^A C_{\alpha_1} \dots C_{\alpha_n}) \vartheta(e^A C_{\alpha_1} \dots C_{\alpha_n}))^{1/2} \right. \\
& \quad \left. \times E_\mu((e^B D_{\alpha_1} \dots D_{\alpha_n}) \vartheta(e^B D_{\alpha_1} \dots D_{\alpha_n}))^{1/2} \right] \\
& \leq \left( \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_\mu((e^A C_{\alpha_1} \dots C_{\alpha_n}) \vartheta(e^A C_{\alpha_1} \dots C_{\alpha_n})) \right)^{1/2} \\
& \quad \times \left( \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_\mu((e^B D_{\alpha_1} \dots D_{\alpha_n}) \vartheta(e^B D_{\alpha_1} \dots D_{\alpha_n})) \right)^{1/2}
\end{aligned} \tag{5.9}$$

Resummation via (5.8) now yields the desired expression.  $\square$

The argument we just saw yields a fundamental criterion for proving reflection positivity:

**Corollary 5.4.** *Fix a plane of reflection  $P$  and let  $\vartheta$  be the corresponding reflection operator. Suppose that the torus Hamiltonian takes the form*

$$-H_L = A + \vartheta A + \sum_\alpha C_\alpha \vartheta C_\alpha \tag{5.10}$$

with  $A, C_\alpha \in \mathfrak{A}^+$ . Then for all  $\beta \geq 0$  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 5.3: Fix  $f, g \in \mathfrak{A}^+$ . Expansion of the exponential term in  $\sum_\alpha C_\alpha \vartheta C_\alpha$  yields

$$\begin{aligned}
E_{\mu_{L,\beta}}(f \vartheta g) &= \frac{1}{Z_L} E_{\otimes \mu_0}(f(\vartheta g) e^{\beta(A+\vartheta A+\sum_\alpha C_\alpha \vartheta C_\alpha)}) \\
&= \frac{1}{Z_L} \sum_{n \geq 0} \frac{1}{n!} \sum_{\alpha_1, \dots, \alpha_n} E_{\otimes \mu_0}((f e^{\beta A} C_{\alpha_1} \dots C_{\alpha_n}) \vartheta(g e^{\beta A} C_{\alpha_1} \dots C_{\alpha_n}))
\end{aligned} \tag{5.11}$$

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 5.3).  $\square$

Now we are ready to check that all 3 interactions that we focused our attention on in previous lectures are of the form in Lemma 5.3 and thus lead to RP torus Gibbs measures:

**Lemma 5.5.** *For any plane  $P$ , the n.n. (ferromagnet) interaction, Yukawa potentials and the power-law decaying potentials, the torus Hamiltonian can be written in the form (5.10) for some  $A, C_\alpha \in \mathfrak{A}^+$ .*

*Proof.* We focus on reflections through bonds; the case of reflections through sites is analogous. 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:

$$-H_L = \underbrace{\frac{1}{2} \sum_{x,y \in \mathbb{T}_L^+} J_{xy}^{(L)} S_x \cdot S_y}_A + \underbrace{\frac{1}{2} \sum_{x,y \in \mathbb{T}_L^-} J_{xy}^{(L)} S_x \cdot S_y}_{\vartheta A} + \underbrace{\sum_{i=1}^d \sum_{\substack{x \in \mathbb{T}_L^+ \\ y \in \mathbb{T}_L^-}} J_{xy}^{(L)} S_x^{(i)} S_y^{(i)}}_{R_i} \quad (5.12)$$

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 a case-by-case basis:

*Nearest-neighbor interactions:* Here

$$R_i = \frac{1}{2d} \sum_{\substack{\langle x,y \rangle \\ x \in \mathbb{T}_L^+ \\ y \in \mathbb{T}_L^-}} S_x^{(i)} S_y^{(i)} \quad (5.13)$$

which is of the desired form since  $S_y = \vartheta(S_x)$  whenever  $x$  and  $y$  contribute to the above sum.

*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  $y \in \mathbb{T}_L^-$ ,

$$J_{xy}^{(L)} = C \sum_{n \geq 0} e^{-\mu(|x|+|y|+nL)} \quad (5.14)$$

Hence,

$$R_i = C \sum_{n \geq 0} e^{-\mu n L} \left( \sum_{x \in \mathbb{T}_L^+} e^{-\mu|x|} S_x^{(i)} \right) \left( \sum_{y \in \mathbb{T}_L^-} e^{-\mu|y|} S_y^{(i)} \right) \quad (5.15)$$

which is of the desired form.

*Power-law potentials:* Here we note

$$\frac{1}{|x-y|_1^s} = \int_0^\infty d\mu \mu^{s-1} e^{-\mu|x-y|_1} \quad (5.16)$$

which reduces the problem to the Yukawa case.  $\square$

We remark that Corollary 5.3 allows a minor generalization: if a torus measure  $\mu$  is RP, and a torus Hamiltonian  $H_L$  takes the form (5.10), then also the measure  $e^{-\beta H_L} d\mu$  is RP. This may seem to be 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.

## 5.2 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 5.6 (Gaussian Domination).** *Let  $(J_{xy})$  be one of the three interactions above. Fix  $\beta \geq 0$  and for  $h = (h_x)_{x \in \mathbb{T}_L} \in (\mathbb{R}^\nu)^{\mathbb{T}_L}$  define*

$$Z_L(h) := E_{\otimes \mu_0} \left( \exp \left\{ -\beta \sum_{x,y \in \mathbb{T}_L} J_{xy}^{(L)} |S_x - S_y + h_x - h_y|^2 \right\} \right) \quad (5.17)$$

Then

$$Z_L(h) \leq Z_L(0) \quad (5.18)$$

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

$$-H_L = A + \vartheta B + \sum_{\alpha} C_{\alpha} \vartheta D_{\alpha} \quad (5.19)$$

Indeed, for  $h \equiv 0$  this is simply Lemma 5.5 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 5.3 which yields

$$Z_L(h)^2 \leq Z_L(h_+) Z_L(h_-) \quad (5.20)$$

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 (5.18): 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

$$N(h) := \#\{ \langle x, y \rangle : h_x \neq h_y \} \quad (5.21)$$

is the smallest 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

$$\min\{N(h_+^*), N(h_-^*)\} < N(h^*) \quad (5.22)$$

Suppose without loss of generality that  $N(h_+^*) < N(h^*)$ . Then the fact that  $h^*$  was a maximizer implies

$$Z_L(h^*)^2 \leq Z_L(h_+^*)Z_L(h_-^*) \leq Z_L(h_+^*)Z_L(h^*) \quad (5.23)$$

which means

$$Z_L(h^*) \leq Z_L(h_+^*) \quad (5.24)$$

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  $c$ , (5.18) follows.  $\square$

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

*Proof of Theorem 3.3.* To ease the notation, we will write throughout

$$\langle \eta, \zeta \rangle := \sum_{x \in \mathbb{T}_L} \eta_x \zeta_x \quad (5.25)$$

to denote the natural inner product on  $L^2(\mathbb{T}_L)$ . First we note that for any  $(\eta_x) \in (\mathbb{R}^\nu)^{\mathbb{T}_L}$ ,

$$\sum_{x, y \in \mathbb{T}_L} J_{xy}^{(L)} |\eta_x - \eta_y|^2 = \langle \eta, G_L^{-1} \eta \rangle \quad (5.26)$$

where  $G_L$  is as in (3.18). (Indeed, in Fourier components,  $\hat{G}_L^{-1}(k) = 1 - \hat{J}(k)$ .) As is easy to check,

$$\begin{aligned} Z_L(h) &= E_{\otimes \mu_0} \left( e^{-\beta \langle S+h, G_L^{-1}(S+h) \rangle} \right) \\ &= Z_L(0) E_{\mu_{L, \beta}} \left( e^{-2\beta \langle h, G_L^{-1} S \rangle - \beta \langle h, G_L^{-1} h \rangle} \right) \end{aligned} \quad (5.27)$$

where  $\mu_{L, \beta}$  is the torus Gibbs measure. The statement of Gaussian domination (5.18) is thus equivalent to

$$E_{\mu_{L, \beta}} \left( e^{-2\beta \langle h, G_L^{-1} S \rangle} \right) \leq e^{\beta \langle h, G_L^{-1} h \rangle} \quad (5.28)$$

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



$$E_{\mu_{L,\beta}}(e^{-2\beta\langle h,S\rangle}) \leq e^{\beta\langle h,G_L h\rangle} \quad \text{whenever} \quad \sum_{x \in \mathbb{T}_L} h_x = 0 \quad (5.29)$$

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

$$\begin{aligned} 1 - 2\beta E_{\mu_{L,\beta}}(\langle h,S\rangle) + \frac{4\beta^2}{2} E_{\mu_{L,\beta}}(\langle h,S\rangle^2) + \dots \\ \leq 1 + \beta\langle h,G_L h\rangle + \dots \end{aligned} \quad (5.30)$$

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

$$E_{\mu_{L,\beta}}(\langle h,S\rangle^2) \leq \frac{1}{2\beta} \langle h, G_L h \rangle \quad (5.31)$$

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 result 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.  $\square$

### 5.3 Chessboard Estimates

The proof of the infrared bound was based on Lemma 5.3 which boils down to the Cauchy-Schwarz inequality for the inner product

$$f, g \mapsto E_\mu(f \vartheta g) \quad (5.32)$$

In this section we will systematize the use of the Cauchy-Schwarz inequality to derive bounds on correlation functions. The key inequality — 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).

Throughout we will restrict attention to reflections through planes of sites as this is somewhat more useful in applications (except for quantum systems). 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  $\Lambda_B$ ,

$$\mathbb{T}_L = \bigcup_{t \in \mathbb{T}_{L/B}} (\Lambda_B + Bt) \quad (5.33)$$

noting that the neighboring translates share the vertices on the adjacent sides. (This is the specific feature of the setup based on reflections through planes of sites.) The translates are indexed by the sites in a “factor torus”  $\mathbb{T}_{L/B}$ .

**Definition 5.7.** A function  $f: \Omega^{\mathbb{T}L} \rightarrow \mathbb{R}$  is called a  $B$ -block function if it depends only on  $\{S_x: x \in \Lambda_B\}$ . 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 5.8 (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  $t_1, \dots, t_m \in \mathbb{T}_{L/B}$ ,

$$E_{\mu} \left( \prod_{j=1}^m \vartheta_{t_j} f_j \right) \leq \prod_{j=1}^m \left[ E_{\mu} \left( \prod_{t \in \mathbb{T}_{L/B}} \vartheta_t f_j \right) \right]^{(B/L)^d} \quad (5.34)$$

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

$$\mu \left( \bigcap_{j=1}^m \vartheta_{t_j} (\mathcal{A}_j) \right) \leq \prod_{j=1}^m \left[ \mu \left( \bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t (\mathcal{A}_j) \right) \right]^{(B/L)^d} \quad (5.35)$$

where

$$\vartheta_t (\mathcal{A}) := \{\vartheta_t 1_{\mathcal{A}} = 1\} \quad (5.36)$$

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 5.8.* 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

$$F(f_1, \dots, f_{2n}) := E_\mu \left( \prod_{t=1}^{2n} \vartheta_t f_t \right) \quad (5.37)$$

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

$$G(f_1, \dots, f_{2n}) := \frac{F(f_1, \dots, f_{2n})}{\prod_{j=1}^{2n} F(f_j, \dots, f_j)^{\frac{1}{2n}}} \quad (5.38)$$

These objects enjoy a natural cyclic invariance,

$$F(f_1, \dots, f_{2n}) = F(f_{2n}, f_1, \dots, f_{2n-1}) \quad (5.39)$$

and, similarly,

$$G(f_1, \dots, f_{2n}) = G(f_{2n}, f_1, \dots, f_{2n-1}) \quad (5.40)$$

The definition of  $G$  also implies

$$G(f, \dots, f) = 1 \quad (5.41)$$

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

$$\begin{aligned} G(f_1, \dots, f_{2n}) &\leq G(f_1, \dots, f_n, f_n, \dots, f_1)^{1/2} \\ &\quad \times G(f_{2n}, \dots, f_{n+1}, f_{n+1}, \dots, f_{2n})^{1/2} \end{aligned} \quad (5.42)$$

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

The desired claim will be proved if we show that

$$G(f_1, \dots, f_{2n}) \leq 1 \quad (5.43)$$

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., we have  $g_1, \dots, g_k = f_i$  (with  $g_{k+1} \neq f_i$  unless  $k = 2n$ ).

We claim that  $k = 2n$ . Indeed, in the opposite case,  $k < 2n$ , we must have  $g_{2n} \neq f_i$  whereby (5.42) combined with the fact that  $(g_1, \dots, g_{2n})$  is a maximizer of  $G$  imply

$$\begin{aligned} G(g_1, \dots, g_{2n})^2 &\leq G(g_1, \dots, g_n, g_n, \dots, g_1)G(g_{2n}, \dots, g_{n+1}, g_{n+1}, \dots, g_{2n}) \\ &\leq G(g_1, \dots, g_n, g_n, \dots, g_1)G(g_1, \dots, g_{2n}) \end{aligned} \tag{5.44}$$

i.e.,

$$G(g_1, \dots, g_{2n}) \leq G(g_1, \dots, g_n, g_n, \dots, g_1) \tag{5.45}$$

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  $\min\{2k, 2n\}$ . This is a contradiction and so we must have  $k = 2n$  after all. In light of (5.41–5.43), 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  $A_B + Bt$ . Writing

$$\prod_{t \in \mathbb{T}_{L/B}} \vartheta_t f_t = \prod_{j=1}^{2n} \left( \prod_{\substack{t \in \mathbb{T}_{L/B} \\ t_1=j}} \vartheta_t f_t \right) \tag{5.46}$$

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

$$\mathfrak{z}_L(\mathcal{A}) := \mu \left( \bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}) \right)^{(B/L)^d} \tag{5.47}$$

The set function  $\mathcal{A} \mapsto \mathfrak{z}_L(\mathcal{A})$  is not generally additive. However, what matters for applications is that it is subadditive:

**Lemma 5.9 (Subadditivity).** *Let  $\mathcal{A}$  and  $\mathcal{A}_1, \mathcal{A}_2, \dots$ , be a collection of  $B$ -block events such that*

$$\mathcal{A} \subset \bigcup_k \mathcal{A}_k \quad (5.48)$$

Then

$$\mathfrak{z}_L(\mathcal{A}) \leq \sum_k \mathfrak{z}_L(\mathcal{A}_k) \quad (5.49)$$

*Proof.* First we use the subadditivity of  $\mu$  and (5.48) to get

$$\mathfrak{z}_L(\mathcal{A})^{|\mathbb{T}_{L/B}|} = \mu\left(\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A})\right) \stackrel{(5.48)}{\leq} \sum_{(k_t)} \mu\left(\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_{k_t})\right) \quad (5.50)$$

Next we apply the chessboard estimate

$$\mu\left(\bigcap_{t \in \mathbb{T}_{L/B}} \vartheta_t(\mathcal{A}_{k_t})\right) \leq \prod_{t \in \mathbb{T}_{L/B}} \mathfrak{z}_L(\mathcal{A}_{k_t}) \quad (5.51)$$

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

$$\begin{aligned} \mathfrak{z}_L(\mathcal{A})^{|\mathbb{T}_{L/B}|} &\leq \sum_{(k_t)} \prod_{t \in \mathbb{T}_{L/B}} \mathfrak{z}_L(\mathcal{A}_{k_t}) \\ &= \prod_{t \in \mathbb{T}_{L/B}} \sum_k \mathfrak{z}_L(\mathcal{A}_k) = \left(\sum_k \mathfrak{z}_L(\mathcal{A}_k)\right)^{|\mathbb{T}_{L/B}|} \end{aligned} \quad (5.52)$$

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

Here is how subadditivity  $\mathfrak{z}_L$  is generally used in computations: In order to estimate the  $\mathfrak{z}_L$ -value of an event, we first cover it by the union of a collection of smaller — and, as desired, easier to compute-with — events, then evaluate the  $\mathfrak{z}_L$ -value for each of them and, finally, add the results.

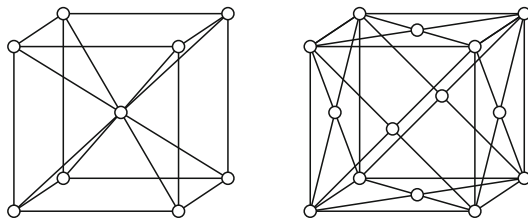
In estimates, we often work with the limiting version,

$$\mathfrak{z}(\mathcal{A}) := \lim_{L \rightarrow \infty} \mathfrak{z}_L(\mathcal{A}) \quad (5.53)$$

of this quantity. We may interpret this as a *partition function per site* restricted to event  $\mathcal{A}$  on each  $B$ -block. The advantage of taking the limit is that it often washes out some annoying finite-size factors and thus provides a more tractable expression to work with. In addition, the limit can be computed using arbitrary — not just periodic — boundary conditions.

### 5.4 Diagonal Reflections, Other Lattices

The above proof of the chessboard estimate is tailored to the underlying setting of the hypercubic lattice, primarily because of its use of the orthogonality between the principal lattice directions. However, some practical problems may lead us to the consideration of other lattices. Some cases generalize directly, e.g., certain instances of the *body-centered cubic* (BCC) or *face-centered cubic* (FCC) lattices, whose unit cells look respectively as follows:



Both of these are decorations of the cubic lattice in which an extra vertex placed in the center of each unit cube (BCC) or a face (FCC) and is attached by edges to the vertices in its ultimate vicinity.

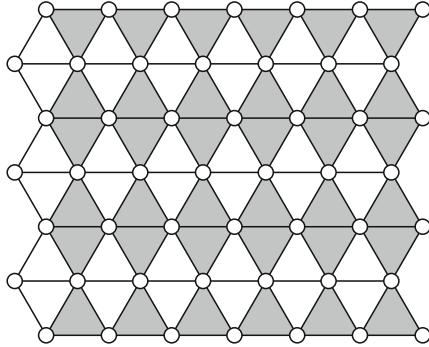
Assuming the interaction (2.15) with  $J_{xy}$  non-zero and positive only for adjacent (i.e., nearest-neighbor) pairs of vertices, the torus Gibbs measure is reflection positive for reflections both through and between the planes of sites of  $\mathbb{Z}^3$ . (A key observation is that the planes between sites of  $\mathbb{Z}^3$  contain some of the added vertices but bisect no additional edges.) The strengths of the interactions across the “old” and “new” edges may not even be the same.

In  $d = 2$ , a corresponding graph is the lattice with a vertex placed in the middle of each square of  $\mathbb{Z}^2$  and edges from it to each of the four corners thereof. By the same reasoning, the nearest-neighbor ferromagnetic interaction leads to a reflection positive torus Gibbs measure.

The situation becomes more involved for the *triangular* (two-dimensional) lattice, whose standard embedding into the complex plane  $\mathbb{C}$  has vertices

$$m + n e^{i\pi/3}, \quad m, n \in \mathbb{Z} \tag{5.54}$$

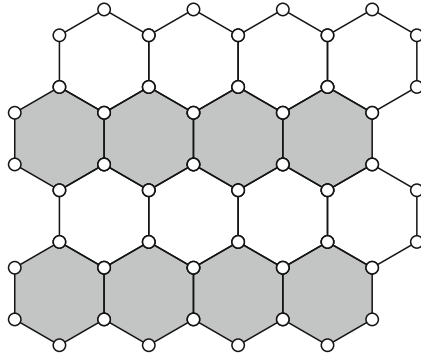
and an edge between any pair of such vertices that differ by a number in the set  $\{1, e^{i\pi/3}, e^{i2\pi/3}\}$ . The principal problem with such graphs is how to place a finite piece of this lattice on a torus in a way that gives rise to reflection positive measures. Here is a convenient choice:



with the torus obtained by identifying the vertices on the opposite sides.

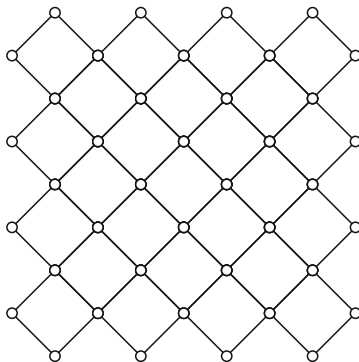
The allowed planes of reflection are all horizontal lines (reflections through sites) and the vertical lines (reflections through both sites and bonds). Again, for ferromagnetic nearest-neighbor interactions, the Gibbs measure with interaction (2.15) is reflection positive. A minor, though annoying, problem occurs in the application of chessboard estimates because the vertical lines of reflections actually cut through triangles. A solution is to focus only on those events that lie either on white or on gray triangles in the above picture and use reflection only with respect to vertical lines that do not cut through the chosen triangles.

A completely analogous situation occurs for the *honeycomb* lattice. Here we consider the domain of the form



and wrap it into a torus by identifying the vertices on the opposite side. Again, for nearest-neighbor ferromagnetic interactions, the resulting Gibbs measure is reflection positive with respect to reflections in vertical lines on sites and horizontal lines between sites. In the application of chessboard estimates to a collection of “hexagon events,” we only use every other horizontal and vertical reflections to a corresponding subset of these events; e.g., those sitting on the shaded hexagons.

A final case of interest is that of *diagonal* reflections in  $\mathbb{Z}^d$ . In  $d = 2$ , this is achieved by wrapping the domain of the form



periodically into a torus. Reflections in the horizontal and vertical lines of sites — the diagonals — are now symmetries of this graph; for nearest-neighbor interactions (of any sign) the corresponding torus Gibbs measure is reflection positive.

The advantage of the diagonal torus is that it permits the use of reflection positivity on collections of “bond events,” i.e., those associated with pairs of nearest-neighbor spins. Subsequent applications of chessboard estimates disseminate a single-bond event over the entire torus. This, in turn, helps in estimates of the quantity  $\mathfrak{z}(\mathcal{A})$  whenever  $\mathcal{A}$  is an event depending on a single square that is itself an intersection of bond events:

**Lemma 5.10.** *Given a unit cube in  $\mathbb{Z}^d$ , let  $\mathcal{A}_b$ , with  $b$  running over all of the  $c_d := d2^{d-1}$  edges in this cube, be a collection of bond events. Then*

$$\mathfrak{z}\left(\bigcap_b \mathcal{A}_b\right) \leq \prod_b \mathfrak{z}(\mathcal{A}_b)^{1/c_d} \quad (5.55)$$

Here  $\mathfrak{z}(\mathcal{A}_b)$  is the partition function per site restricted to configurations such that  $\mathcal{A}_b$ , or its corresponding reflection, occurs at all edges of  $\mathbb{Z}^d$ .

*Proof.* Let us first focus on  $d = 2$ . The key fact is that the partition function per site,  $\mathfrak{z}(\mathcal{A})$ , does not depend on what boundary conditions were used to define it. So, in order to compute  $\mathfrak{z}$  of the intersection event, we may first wrap the square lattice into the diagonal torus, and disseminate the bond events before passing to the  $L \rightarrow \infty$ . As there are  $c_2 = 4$  edges in each lattice square, there is an extra power of  $1/4$ .

In  $d > 2$ , we perform the same by singling out two lattice directions and wrapping the torus diagonally in these, and regularly in the



remaining ones. This homogenizes the event in two lattice directions. Proceeding by induction, the claim follows.  $\square$

## 5.5 Literature Remarks

The material of this section is entirely classical; a possible exception is Lemma 5.9 which seems to have been formulated in the present form only relatively recently [13]. The use of reflection positivity goes back to the days of constructive quantum field theory (namely, the Osterwalder-Schrader axioms [84]) 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 [50] (infrared bound) and Fröhlich and Lieb [48] (chessboard estimates). The theory was further developed in two papers by Fröhlich, Israel, Lieb and Simon [46, 47]. There have been a couple of nice reviews of this material, e.g., by Shlosman [95] and in Georgii [57].

All use of reflection positivity in these notes is restricted to one of the three interactions introduced in Chapter 3. Various generalizations beyond these are possible. For instance, 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  $J \geq 2(d-1)|\lambda|$ . For reflections through planes of sites, we may even allow any sort of interactions involving the spins in a given lattice cube. (This exhausts all finite range interactions; any longer range RP interactions are automatically infinite range.) Many other examples are discussed, e.g., in [46, page 32].

Notwithstanding our decision to restrict attention only to three specific interactions, the set of reflection positive interactions is not so small as it may appear. Indeed, in the class of translation and rotation invariant coupling constants, letting

$$F(x_1, \dots, x_d) := J_{0,x} \tag{5.56}$$

we check that a sufficient conditions for RP is that the matrix

$$(x, y) \mapsto F(x_1 + y_1, x_2 - y_2, \dots, x_d - y_d) 1_{\{x_1 > 0\}} 1_{\{y_1 > 0\}} \tag{5.57}$$

is positive semidefinite. (See (5.15) for a specific case of this.) By Shur's Theorem — namely that if  $(a_{ij})$  and  $(b_{ij})$  are positive semidefinite matrices, then so is  $(a_{ij}b_{ij})$  — we thus know that if  $J^{(1)}$  and  $J^{(2)}$  are two collections of RP couplings, then also the collection  $J_{xy}^{(1)} J_{xy}^{(2)}$  is RP. In particular, the set of RP couplings is closed under taking products.

The situation on other lattices is discussed in [47]; the use of diagonal reflections goes back to [95]. We caution the reader that it is rather easy to make a mistake in this context. For instance, the regularly-wrapped  $(L \times L)$  torus in  $\mathbb{Z}^2$  is also symmetric with respect to all of the diagonal reflections. However, for diagonal reflection on direct torus it is not possible to define *two* components of the “plane of reflection” so that the reflection in one leaves the other intact. So we cannot simultaneously use both direct and diagonal reflections, and this prevents a direct proof of (5.55) in finite volume. (This error appeared in [17, eq. 4.39] though, as shown in Lemma 5.10, all differences get washed out in the thermodynamic limit.)

Gaussian domination appears in a rather different context as the celebrated *Brascamp-Lieb inequality*. Consider the measure on  $\mathbb{R}^n$  of the form

$$\mu(dx) := Z^{-1}e^{-V(x)}dx \quad (5.58)$$

with  $V$  smooth and strictly convex. Let  $V''(x)$  be the Hessian, i.e., an  $n \times n$  matrix of all second derivatives of  $V$ . Then for each smooth  $f$  with compact support,

$$E_\mu(f^2) - (E_\mu f)^2 \leq \int dx \langle (V'')^{-1} \nabla f(x), \nabla f(x) \rangle \quad (5.59)$$

where  $\langle \cdot, \cdot \rangle$  denote the  $n$ -dimensional Euclidean inner product. In particular, if  $Q$  is a positive definite  $n \times n$  matrix that dominates the Hessian from below at all  $x$ , then the correlations of  $\mu$  are dominated by those of the Gaussian measure with covariance  $2Q^{-1}$ . This is, unfortunately, not very useful in the analysis of the Gibbs measures for general lattice spin systems as these are generally not of the required form — e.g., because the restriction to a specific spin-space (a unit sphere for the Heisenberg model) cannot be approximated by convex functions.

## 6 Applications of Chessboard Estimates

In this section 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.

### 6.1 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

$$\beta H(\phi) := \beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 + \sum_x V(\phi_x) \quad (6.1)$$

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

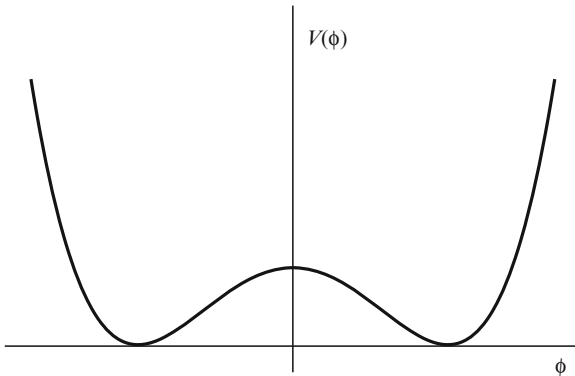
$$\text{Cov}(\phi_x, \phi_y) = \sum_{k \in \mathbb{T}_L^*} \frac{e^{ik \cdot (x-y)}}{\beta \widehat{D}(k) + \kappa} \quad (6.2)$$

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

$$\widehat{D}(k) := \sum_{j=1}^d |1 - e^{ik_j}|^2 \quad (6.3)$$

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

$$e^{-V(\phi)} := e^{-\frac{\kappa}{2}(\phi-1)^2} + e^{-\frac{\kappa}{2}(\phi+1)^2} \quad (6.4)$$

It is easy to check that, for  $\kappa$  sufficiently large,  $V$  defined using this formula looks as in the figure. The reason for assuming (6.4) is the possibility of an *Ising-spin representation*. Indeed, we may rewrite (6.4) as

$$e^{-V(\phi)} = \sum_{\sigma=\pm 1} e^{-\frac{\kappa}{2}(\phi-\sigma)^2} = C \sum_{\sigma=\pm 1} e^{-\frac{\kappa}{2}\phi_x - \kappa\phi_x\sigma_x} \quad (6.5)$$

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

$$\prod_x e^{-V(\phi_x)} \propto \sum_{(\sigma_x)} \prod_x e^{-\frac{\kappa}{2}\phi_x^2 - \kappa\phi_x\sigma_x} \quad (6.6)$$

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

$$\begin{aligned} & e^{-\beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 - \sum_x V(\phi_x)} \\ & \propto \sum_{(\sigma_x)} e^{-\beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 - \frac{\kappa}{2} \sum_x \phi_x^2} e^{-\kappa \sum_x \phi_x \sigma_x} \end{aligned} \quad (6.7)$$

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

$$\beta H(\phi, \sigma) := \beta \sum_{\langle x,y \rangle} (\phi_x - \phi_y)^2 + \frac{\kappa}{2} \sum_x \phi_x^2 + \kappa \sum_x \phi_x \sigma_x \quad (6.8)$$

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 6.1.** *Let  $\mu$  be a Gibbs measure for Hamiltonian (6.8) and let  $\nu$  be its  $\phi$ -marginal. Then  $\nu$  is a Gibbs measure for the Hamiltonian (6.1) subject to (6.4). The marginal  $\nu$  completely determines  $\mu$ : For any  $f$  depending only on  $\phi$  and  $\sigma$  in a finite set  $\Lambda$ ,*

$$E_\mu(f) = E_\nu \left( \sum_{(\sigma_x)_{x \in \Lambda}} f(\phi, \sigma) \prod_{x \in \Lambda} e^{V(\phi_x) - \frac{\kappa}{2}(\phi_x - \sigma_x)^2} \right) \quad (6.9)$$

We will use  $\mathfrak{G}_{\beta,\kappa}$  to denote the set of all Gibbs measures for the Hamiltonian (6.8) with parameters  $\beta$  and  $\kappa$ . The principal result for this model is as follows:

**Theorem 6.2.** *Let  $d \geq 2$ . For each  $\epsilon > 0$  there is  $c > 0$  such that if  $\kappa, \kappa/\beta > c$ , then there exist  $\mu^+, \mu^- \in \mathfrak{G}_{\beta,\kappa}$  which are translation invariant and obey*

$$\mu^\pm(\sigma_x = \pm 1) \geq 1 - \epsilon \quad (6.10)$$

and

$$E_{\mu^\pm}((\phi_x \mp 1)^2) \leq \epsilon \quad (6.11)$$

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  $V$  given by

$$e^{-V(\phi)} = e^{-\frac{\kappa_+}{2}(\phi-1)^2+h} + e^{-\frac{\kappa_-}{2}(\phi+1)^2-h}. \quad (6.12)$$

where  $h$  changes the relative weight of 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.

## 6.2 Proof of Phase Coexistence

Here we will prove Theorem 6.2. 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

$$\mathfrak{z}(\mathcal{B}) \ll 1 \quad \text{once } \beta, \kappa \gg 1 \quad (6.13)$$

Appealing to the subadditivity lemma (Lemma 5.9) 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:

$$\begin{array}{ccc} + & - & + \\ + & - & + \end{array} \quad (6.14)$$

We begin with the most interesting of the three:

**Lemma 6.3.** *For any  $\beta, \kappa > 0$ ,*

$$\mathfrak{z}_{\left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)} \leq e^{-\frac{4\beta\kappa}{8\beta+\kappa}} \quad (6.15)$$

*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

$$\mathfrak{z}_L \left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)^{|\mathbb{T}_L|} := \frac{Z_L \left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)}{Z_L} \leq \frac{Z_L \left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)}{Z_L \left(\begin{smallmatrix} ++ \\ ++ \end{smallmatrix}\right)} \quad (6.16)$$

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

$$\begin{aligned} Z_L \left(\begin{smallmatrix} ++ \\ ++ \end{smallmatrix}\right) &= \int e^{-\beta \sum_{\langle x, y \rangle} (\phi_x - \phi_y)^2 - \frac{\kappa}{2} \sum_x \phi_x^2} e^{-\kappa \sum_x \phi_x} \prod_{x \in \mathbb{T}_L} d\phi_x \\ &= \left( \dots \right) E_{\text{GFF}} \left( e^{-\kappa \sum_x \phi_x} \right) \end{aligned} \quad (6.17)$$

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

$$Z_L \left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right) = \left( \dots \right) E_{\text{GFF}} \left( e^{-\kappa \sum_x \phi_x (-1)^{|x|}} \right) \quad (6.18)$$

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

$$\mathfrak{z}_L \left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)^{|\mathbb{T}_L|} \leq \frac{E_{\text{GFF}} \left( e^{-\kappa \sum_x \phi_x} \right)}{E_{\text{GFF}} \left( e^{-\kappa \sum_x \phi_x (-1)^{|x|}} \right)} \quad (6.19)$$

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

$$E(e^{\lambda \cdot X}) = e^{\lambda \cdot EX + \frac{1}{2} \text{Var}(\lambda \cdot X)} \tag{6.20}$$

Since  $E_{\text{GFF}}(\phi_x) = 0$ , we only need to compute the (diagonal) matrix element of  $\text{Cov}(\phi_x, \phi_y)$  against vectors  $1 = (1, 1, \dots)$  and  $(-1)^{|x|}$ . However, a quick look at (6.2) 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

$$\text{Var}_{\text{GFF}}\left(\sum_x \phi_x\right) = \frac{|\mathbb{T}_L|}{\kappa} \tag{6.21}$$

$$\text{Var}_{\text{GFF}}\left(\sum_x \phi_x (-1)^{|x|}\right) = \frac{|\mathbb{T}_L|}{8\beta + \kappa} \tag{6.22}$$

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 (6.19) we conclude

$$\mathfrak{z}_L\left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)^{|\mathbb{T}_L|} \leq \exp\left\{\frac{1}{2}|\mathbb{T}_L|\kappa^2\left(\frac{1}{8\beta + \kappa} - \frac{1}{\kappa}\right)\right\} \tag{6.23}$$

from which the claim readily follows.  $\square$

Next we attend to the other patterns:

**Lemma 6.4.** *For any  $\beta, \kappa > 0$ ,*

$$\mathfrak{z}\left(\begin{smallmatrix} +- \\ +- \end{smallmatrix}\right) \leq e^{-\frac{2\beta\kappa}{4\beta + \kappa}} \tag{6.24}$$

and

$$\mathfrak{z}\left(\begin{smallmatrix} -+ \\ ++ \end{smallmatrix}\right) \leq e^{-\frac{2\beta\kappa}{8\beta + \kappa}} \tag{6.25}$$

*Proof.* As for (6.24), dissemination of  $\begin{smallmatrix} +- \\ +- \end{smallmatrix}$  leads to alternating stripes of plusses and minuses, i.e.,  $\sigma_x = (-1)^{|x_1|}$ . Again, this is an eigenvector of the covariance matrix (6.2) with  $k = (\pi, 0)$ . The corresponding  $\widehat{D}$  equals 4 and so

$$\mathfrak{z}_L\left(\begin{smallmatrix} +- \\ +- \end{smallmatrix}\right)^{|\mathbb{T}_L|} \leq \exp\left\{\frac{1}{2}|\mathbb{T}_L|\kappa^2\left(\frac{1}{4\beta + \kappa} - \frac{1}{\kappa}\right)\right\} \tag{6.26}$$

yielding (6.24).

The pattern  $\begin{smallmatrix} -+ \\ ++ \end{smallmatrix}$  is more complex because its dissemination will not lead to an eigenvector of the covariance matrix. However, we circumvent this problem by noting that Lemma 5.10 implies

$$\mathfrak{z}\left(\begin{smallmatrix} -+ \\ ++ \end{smallmatrix}\right) \leq \mathfrak{z}\left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)^{1/2} \mathfrak{z}\left(\begin{smallmatrix} ++ \\ ++ \end{smallmatrix}\right)^{1/2} \leq \mathfrak{z}\left(\begin{smallmatrix} -+ \\ +- \end{smallmatrix}\right)^{1/2} \tag{6.27}$$

where we used  $\mathfrak{z}\left(\begin{smallmatrix} ++ \\ ++ \end{smallmatrix}\right) \leq 1$ . Now (6.25) follows from Lemma 6.3.  $\square$

**Corollary 6.5.** *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 6.3–6.4 the claim holds for  $\mathcal{B}$  replaced by any fixed bad pattern. The desired bound now follows — with slightly worse constants — by invoking Lemma 5.9.  $\square$

Next we explain our focus on the bad event:

**Lemma 6.6.** *There exists a constant  $c \in (1, \infty)$  such that if  $c\mathfrak{z}(\mathcal{B}) < 1/2$  then for any  $x, y \in \mathbb{T}_L$ ,*

$$\mu_L(\sigma_x = 1, \sigma_y = -1) \leq 2c\mathfrak{z}(\mathcal{B}). \quad (6.28)$$

*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

$$\mu_L(\sigma_x = 1, \sigma_y = -1) \leq \sum_{\gamma} \mu_L\left(\bigcap_{t \in \gamma} \vartheta_t(\mathcal{B})\right) \leq \sum_{\gamma} \mathfrak{z}(\mathcal{B})^{|\gamma|} \quad (6.29)$$

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

$$\mu_L(\sigma_x = 1, \sigma_y = -1) \leq \sum_{n \geq 1} c^n \mathfrak{z}(\mathcal{B})^n \quad (6.30)$$

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 desired proof of phase coexistence:

*Proof of Theorem 6.2.* By symmetry of the torus measure, we have

$$\mu_L(\sigma_x = 1) = 1/2 = \mu_L(\sigma_x = -1). \quad (6.31)$$

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



$$\mu_L^\pm(-) := \mu_L(-|\sigma_z = \pm 1). \quad (6.32)$$

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 6.6 we have

$$\mu_L^+(\sigma_x = -1) \leq 2c\mathfrak{z}(\mathcal{B}) \quad (6.33)$$

once  $\mathfrak{z}(\mathcal{B}) \ll 1$  and, by Corollary 6.5, 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.  $\square$

Notice that 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., (6.12), one can either invoke a continuity argument — choose  $h = h_L$  such that (6.31) holds — or turn (6.28) 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  $h$ .

To prove Theorem 6.2, it remains to show the concentration of the  $\phi$ 's around the  $\sigma$ 's:

*Proof of (6.11).* Let  $\mu$  be a Gibbs measure for parameters  $\beta$  and  $\kappa$ . Then (6.8) shows that, conditional on the  $\sigma$ 's, the  $\phi$ 's are Gaussian with mean

$$E_\mu(\phi_x|\sigma) = \kappa((2\beta\Delta + \kappa)^{-1}\sigma)_x \quad (6.34)$$

and covariance  $(2\beta\Delta + \kappa)^{-1}$ , where  $\Delta$  is the lattice Laplacian. Now, once  $\beta/\kappa \ll 1$  we may expand the inverse operator into a power series to get

$$E_\mu(\phi_x|\sigma) - \sigma_x = \sum_{n \geq 1} \left(\frac{2\beta}{\kappa}\right)^n (\Delta^n \sigma)_x \quad (6.35)$$

which by the fact that  $|\sigma_z| = 1$  is  $O(\beta/\kappa)$  independently of  $x$ . Since the conditional variance of  $\phi_x$  is  $O(1/\kappa)$ , we obtain

$$\begin{aligned} E_\mu((\phi_x - \sigma_x)^2|\sigma) &\leq 2E_\mu([\phi_x - E(\phi_x|\sigma)]^2|\sigma) + 2(E(\phi_x|\sigma) - \sigma_x)^2 \\ &= O((\beta/\kappa)^2) + O(1/\kappa) \end{aligned} \quad (6.36)$$

with the constants implicit in the  $O$ 's independent of the  $\sigma$ 's and  $x$ . Thus, if  $\kappa \gg 1$  and  $\kappa/\beta \gg 1$ , then (6.11) follows by the fact that  $\mu^\pm$  put most of the mass on  $\sigma_x = \pm 1$ .  $\square$

### 6.3 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

$$\beta H(\phi) := \sum_{\langle x,y \rangle} V(\phi_x - \phi_y) \quad (6.37)$$

The requirements that we generally put on  $V$  are continuity, evenness and quadratic growth 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*,

$$\eta_b := \phi_y - \phi_x \quad \text{if } b \text{ is the oriented edge } (x, y), \quad (6.38)$$

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

$$\eta_{b_1} + \eta_{b_2} + \eta_{b_3} + \eta_{b_4} = 0 \quad (6.39)$$

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

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

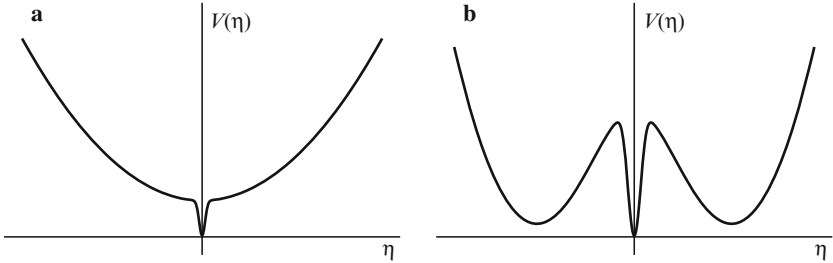
**Theorem 6.7.** *Suppose  $V$  is convex, twice continuously differentiable with  $V''$  bounded away from zero and infinity. Then the shift-ergodic GGMs  $\mu$  are in one-to-one correspondence with their tilt, which is a vector  $a \in \mathbb{R}^d$  such that*

$$E_\mu(\eta_b) = a \cdot b \tag{6.40}$$

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 the Brascamp-Lieb inequality through which the convexity assumption enters in an essential way. It is also known that the large-scale fluctuation structure of the  $\eta$ 's is that of a Gaussian Free Field.

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:

$$e^{-V(\eta)} := p e^{-\kappa_O \eta^2/2} + (1 - p) e^{-\kappa_D \eta^2/2} \tag{6.41}$$

where  $\kappa_O$  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 6.8.** *Suppose  $d = 2$  and  $\kappa_O \gg \kappa_D$ . 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:

$$\frac{1}{|\Lambda_L|} \sum_{\substack{b=(x,y) \\ x,y \in \Lambda_L}} \eta_b \xrightarrow{L \rightarrow \infty} 0, \quad \mu_{\text{ord}}, \mu_{\text{dis}}\text{-a.s.} \tag{6.42}$$

(2) *distinct fluctuation size:*

$$E_{\mu_{\text{ord}}}(\eta_b^2) \ll E_{\mu_{\text{dis}}}(\eta_b^2) \quad (6.43)$$

The upshot of this result is that, once the convexity of  $V$  is strongly violated, the conclusions of Theorem 6.7 do not apply. While the example is restricted to  $d = 2$ , and to potentials of the form (6.41), generalizations to  $d \geq 2$  and other potentials as in the above figure are possible and reasonably straightforward.

Here are the main steps of the proof. First, as for the Gaussian double-well model, we use (6.41) to expand the Gibbs weight according to whether the first or the second term in (6.41) applies. This gives rise to a configuration of coupling strengths  $(\kappa_b)$ , one for each bond  $b$ , which take values in  $\{\kappa_{\text{O}}, \kappa_{\text{D}}\}$ . The joint Hamiltonian of the  $\eta$ 's and the  $\kappa$ 's is

$$\beta H(\eta, \kappa) := \sum_b \frac{\kappa_b}{2} \eta_b^2 \quad (6.44)$$

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

For the proof of phase coexistence, we focus on lattice plaquettes and divide these into good and bad according to whether all of the edges have the same coupling  $\kappa$  or not. The dissemination of each bad patterns leads to a Gaussian integral but this time for GFF with inhomogeneous — yet periodically varying — couplings. For instance, the pattern with three bonds of type  $\kappa_{\text{O}}$  and one of type  $\kappa_{\text{D}}$  disseminates into periodic configuration where the edges on every other vertical line is of type  $\kappa_{\text{D}}$  and all other edges are of type  $\kappa_{\text{O}}$ . Similarly for all other bad patterns.

The periodic nature of the disseminated events allows the use of Fourier modes — i.e., pass to the reciprocal torus — to diagonalize the requisite covariance matrices. For instance, the aforementioned pattern with three  $\kappa_{\text{O}}$ 's and one  $\kappa_{\text{D}}$  leads to a configuration which is periodic with period two. A calculation shows that the covariance is block diagonal with  $2 \times 2$  blocks of the form

$$\Pi(k) := \begin{pmatrix} \kappa_{\text{O}}|a_-|^2 + \frac{1}{2}(\kappa_{\text{O}} + \kappa_{\text{D}})|b_-|^2 & \frac{1}{2}(\kappa_{\text{O}} - \kappa_{\text{D}})|b_-|^2 \\ \frac{1}{2}(\kappa_{\text{O}} - \kappa_{\text{D}})|b_-|^2 & \kappa_{\text{O}}|a_+|^2 + \frac{1}{2}(\kappa_{\text{O}} + \kappa_{\text{D}})|b_-|^2 \end{pmatrix} \quad (6.45)$$

where  $a_{\pm}$  and  $b_{\pm}$  are defined by

$$a_{\pm} = 1 \pm e^{ik_1} \quad \text{and} \quad b_{\pm} = 1 \pm e^{ik_2} \quad (6.46)$$

with  $k := (k_1, k_2)$  varying through one half of the reciprocal torus  $\mathbb{T}_L^*$ . (The block combines the contribution of both  $k$  and  $k + \pi\hat{e}_1$ , and so we only need half of all  $k$ 's.) The requisite Gaussian integral then reduces to  $\prod_{k \in \mathbb{T}_L^* \setminus \{0\}} [\det \Pi(k)]^{-1/4}$  where in the exponent we get  $1/4$  instead of the expected  $1/2$  to account for double counting of the  $k$ 's. To estimate the growth rate of this product, we note that

$$\begin{aligned} & \prod_{k \in \mathbb{T}_L^* \setminus \{0\}} [\det \Pi(k)]^{-1/4} \\ &= \exp \left\{ -|\mathbb{T}_L| \frac{1}{4} \int \frac{dk}{(2\pi)^2} \log \det \Pi(k) + o(|\mathbb{T}_L|) \right\} \end{aligned} \quad (6.47)$$

The integral plays the role of the free energy associated with the Gaussian variables on the background of the specific periodic configuration of the  $\kappa$ 's. A similar expression — with different integrand — applies to each pattern.

Comparing the integrals for all possible arrangements of the two types of bonds around a plaquette, we find that under the condition  $\kappa_O \gg \kappa_D$ , the bad patterns are heavily suppressed. Thus bad plaquettes are infrequent and can be regarded as parts of a contour. As it is not possible to pass from all- $\kappa_O$  pattern to all- $\kappa_D$  pattern without crossing a bad plaquette, the coexistence follows — as for the double-well model — by a standard Peierls' argument and chessboard estimates. Full details of the proof are to be found in a paper by Kotecký and the present author.

The two-dimensional model has the special feature that we can actually compute  $p_t$ :

**Theorem 6.9.** *Let  $d = 2$ . If  $\kappa_O/\kappa_D \gg 1$ , then  $p_t$  is given by*

$$\frac{p_t}{1 - p_t} = \left( \frac{\kappa_D}{\kappa_O} \right)^{1/4}. \quad (6.48)$$

This is a consequence of a duality relation that can be used to exchange the roles of  $\kappa_O$  and  $\kappa_D$ . It is also interesting to note that, while the one-to-one correspondence between the Gibbs measures and their tilt is violated for non-convex potentials, the large-scale fluctuation structure remains that of a Gaussian Free Field. Indeed, we have:

**Theorem 6.10.** *Let  $d = 2$ . For each translation-invariant, ergodic gradient Gibbs measure  $\mu$  with zero tilt, there exists a positive-definite  $d \times d$  matrix  $q = q(\mu)$  such that for any smooth  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  with compact support and  $\int f(x) dx = 0$ ,*

$$\int dx \phi_{\lfloor x/\epsilon \rfloor} f(x) \xrightarrow[\epsilon \downarrow 0]{\mathcal{D}} \mathcal{N}(0, (f, Q^{-1}f)) \quad (6.49)$$

where  $\mathcal{N}(0, C)$  denotes a normal random variable with mean zero and covariance  $C$  and  $Q$  is the elliptic operator

$$Qf(x) := \sum_{i,j=1}^d q_{ij} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \quad (6.50)$$

The basis of this result — derived in all  $d \geq 1$  by Spohn and the present author — is the fact that, conditional on the  $\kappa$ 's, the  $\phi$ 's are Gaussian with mean zero and covariance given by the inverse of the generator of a reversible random walk in random environment. The Gaussian limit is a consequence of an (annealed) invariance principle for such random walks and some basic arguments in homogenization theory. The restriction to zero tilt appears crucially in the proof.

#### 6.4 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 with such property is as follows:

*Orbital compass model:* Here the spins on  $\mathbb{Z}^d$  take values in a unit sphere in  $\mathbb{R}^d$ , i.e.,  $S_x \in \mathbb{S}^{d-1}$  with  $x \in \mathbb{Z}^d$ . The Hamiltonian is

$$H(S) := \sum_x \sum_{\alpha=1}^d (S_x^{(\alpha)} - S_{x+\hat{e}_\alpha}^{(\alpha)})^2 \quad (6.51)$$

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.

Despite a formal similarity with the Heisenberg model, note that only one component of the spin is coupled in each lattice direction. Notwithstanding, *every* constant configuration is still a minimum-energy state of (6.51). Further 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 operations are possible that preserve the minimum-energy property.

The key question is now what happens with this huge ground-state degeneracy at positive temperatures. Here is a theorem one can prove about the two-dimensional system:

**Theorem 6.11.** *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*

$$E_{\mu_j}(|S_x \cdot \hat{e}_j|) \geq 1 - \epsilon, \quad j = 1, 2 \quad (6.52)$$

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

$$E_\mu(S_x) = 0 \quad (6.53)$$

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

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

$$\beta\Delta^2 > \frac{1}{\delta} \quad \text{and} \quad \beta\Delta^3 < \delta \quad (6.54)$$

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

$$E_{\otimes \mu_0} \left( e^{-\beta H_L(S)} \prod_{x \in \mathbb{T}_L} 1_{\{|S_x - \hat{v}_\theta| < \Delta\}} \right) = \left( \frac{2\pi}{\beta} \right)^{L^2/2} e^{-L^2[F(\theta) + o(\epsilon)]} \quad (6.55)$$

where

$$F(\theta) := \frac{1}{2} \int \frac{dk}{(2\pi)^2} \log \{ \sin^2(\theta) |1 - e^{ik_1}|^2 + \cos^2(\theta) |1 - e^{ik_2}|^2 \} \quad (6.56)$$

The quantity  $F$  has the interpretation of the *spin-wave free energy* where the term “spin wave” refers to slowly varying deformations of a constant ground state. A convexity argument — based on the identity  $\sin^2(\theta) + \cos^2(\theta) = 1$  — now shows that  $F$  is minimized by  $\theta = 0, \pi/2, \pi, 3\pi/2$ , i.e., exactly in one of the coordinate directions. This corroborates the intuition that only the configurations with most of the spins aligned in one of these directions will be relevant at low temperatures. However, to extract a proof of phase coexistence, we will have to again invoke a Peierls’ argument.

Fix  $\kappa > 0$  and let  $\Delta := \beta^{-\frac{5}{12}}$  and  $B := \log \beta$  and let  $\mathcal{B}_E$  and  $\mathcal{B}_{SW}$  denote the following events:

- (1)  $\mathcal{B}_E := \{ \text{a pair of neighboring spins in } \Lambda_B \text{ differ by an angle } \geq \Delta \}$
- (2)  $\mathcal{B}_{SW}$  is the set of configurations in the complement of  $\mathcal{B}_E$  in which 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 directions

The event  $\mathcal{B}_E$  captures the situations when two neighboring spins are not quite close to each other leading to excess energy order  $\Delta^2$ . As a result of that,

$$\mathfrak{z}(\mathcal{B}_E) \leq 3B^3 e^{-c_3 \beta \Delta^2} \quad (6.57)$$

The event  $\mathcal{B}_{SW}$  collects the configurations where the energy is good but the fluctuations are not sufficiently powerful. The calculation in Lemma 6.12 and a simple use of the subadditivity lemma show

$$\mathfrak{z}(\mathcal{B}_{SW}) \leq \frac{c_1}{\Delta} e^{-c_2 B^3 \kappa^2} \quad (6.58)$$

for some constants  $c_1, c_2 > 0$ . Thus, for our choices of  $\Delta$  and  $B$ , once  $\beta \gg 1$  the density of blocks where  $\mathcal{B}_E \cup \mathcal{B}_{SW}$  occurs in any typical configuration from the torus measure will be rather small. 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.

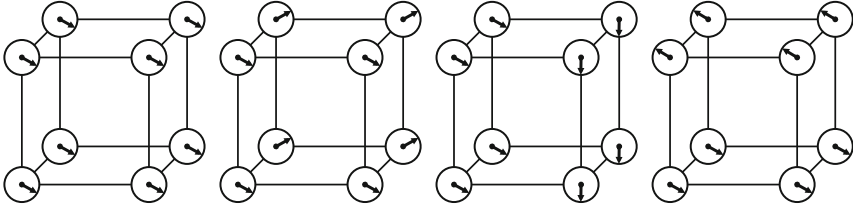
*120-degree model:* A somewhat more complicated version of the interaction, but with the spins  $S_x$  taking values in the unit circle  $\mathbb{S}^1$ , can be contrived in  $d = 3$ . The Hamiltonian will actually look just as for the orbital compass model except that  $S_x^{(\alpha)}$  are not Cartesian components but projections on the three third-roots of unity  $\hat{b}_1, \hat{b}_2, \hat{b}_3$  in  $\mathbb{S}^1$ . Explicitly,

$$H(S) := \sum_x \sum_{\alpha=1,2,3} (S_x \cdot \hat{b}_\alpha - S_{x+\hat{e}_\alpha} \cdot \hat{b}_\alpha)^2 \quad (6.59)$$

Again, all constant configurations are ground states and further ground states may again be obtained by judicious reflections. Fortunately, 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. Namely, 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 we say rigorously about this model:

**Theorem 6.13.** *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*

$$E_{\mu_j}(S_x \cdot \hat{w}_j) \geq 1 - \epsilon, \quad j = 1, \dots, 6 \tag{6.60}$$

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

The ideas underlying this theorem are quite similar to the orbital compass model. First we find out that the spin-wave free energy for fluctuations about the ground state pointing in direction  $\theta$  is given by

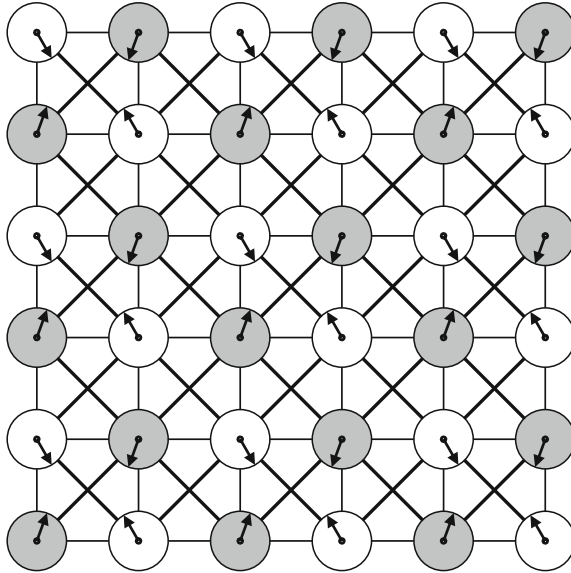
$$F(\theta) := \frac{1}{2} \int \frac{dk}{(2\pi)^3} \left[ \log \sum_{\alpha=1,2,3} q_\alpha(\theta) |1 - e^{ik_\alpha}|^2 \right] \tag{6.61}$$

where  $q_1 := \sin^2(\theta)$ ,  $q_2 := \sin^2(\theta - 120^\circ)$  and  $q_3 := \sin^2(\theta + 120^\circ)$ . A surprisingly sophisticated argument is then required to show that  $F$  is minimal only for  $\theta$  of the form  $\frac{\pi}{3}j$ ,  $j = 1, \dots, 6$ . Once we have this information, the rest of the argument follows a route very similar to that for the orbital compass model (including the introduction of the scales  $\kappa$  and  $\Delta$  and the corresponding events  $\mathcal{B}_E$  and  $\mathcal{B}_{SW}$ ). Details appeared in a paper by Chayes, Nussinov and the present author.

*n.n. and n.n.n. antiferromagnet:* Finally, we will 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:

$$H(S) := \gamma \sum_x [S_x \cdot S_{x+\hat{e}_1} + S_x \cdot S_{x+\hat{e}_2}] + \sum_x [S_x \cdot S_{x+\hat{e}_1+\hat{e}_2} + S_x \cdot S_{x+\hat{e}_1-\hat{e}_2}] \quad (6.62)$$

Assuming  $|\gamma| < 2$ , the minimum energy state is obtained 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! Here is a configuration of this form:



For this system we can nevertheless prove the following theorem:

**Theorem 6.14.** *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*

$$-E_{\mu_j}(S_x \cdot S_{x+\hat{e}_1 \pm \hat{e}_2}) \geq 1 - \epsilon \quad (6.63)$$

and

$$E_{\mu_j}(S_x \cdot S_{x+\hat{e}_j}) \geq 1 - \epsilon, \quad j = 1, 2 \quad (6.64)$$

*There are no shift-ergodic  $\mu \in \mathfrak{G}_\beta$ ,  $\beta \geq \beta_0$ , for which either (6.63) or at least one of (6.64) does not hold.*

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

$$F(\theta) := \frac{1}{2} \int_{[-\pi, \pi]^2} \frac{d\mathbf{k}}{(2\pi)^2} \log D_{\mathbf{k}}(\theta) \quad (6.65)$$

where

$$D_{\mathbf{k}}(\theta) := |1 - e^{i(k_1+k_2)}|^2 + |1 - e^{i(k_1-k_2)}|^2 + \gamma \cos(\theta) (|1 - e^{ik_1}|^2 - |1 - e^{ik_1}|^2) \quad (6.66)$$

As  $D(\theta) = \alpha D(0) + (1 - \alpha)D(\pi)$ , with  $\alpha := \frac{1}{2}(1 + \cos(\theta))$ , Jensen's inequality for the logarithm directly shows that  $F$  is minimized by  $\theta = 0$  or  $\theta = \pi$ . In 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. Details of all calculations appeared in a paper by Chayes, Kivelson and the present author.

Notice that, despite the fact that the lattices maintain a specific *relative* orientation at low temperatures, a Mermin-Wagner argument ensures that every Gibbs measure is invariant under a rigid rotation of all spins.

## 6.5 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 [35]. Representations of the kind (6.4) have been used already before, e.g., by Külske [74, 75] and Zahradník [107]. The method of proof presented here draws on the work of Dobrushin, Kotecký and Shlosman [33, 71, 69] which was used to control order-disorder transitions in a number of systems; most notably, the  $q$ -state Potts model with  $q \gg 1$  [69]. These methods may be combined with graphical representations of Edwards-Sokal [39] (or Fortuin-Kasteleyn [43]) to establish rather complicated phase diagrams, e.g., [26, 12]. Recently, the method has been used to resolve a controversy about a transition can occur in 2D non-linear vector models [41, 42].

Theorem 6.7 has been proved by Funaki and Spohn [53]. 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 [52], Velenik [104] or Sheffield [92].) Theorem 6.8 was proved by Biskup and Kotecký [17]; Theorem 6.10 was derived by Biskup and Spohn [18].

The interest in models in Sect. 6.4 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 [13] (120-degree model) Biskup, Chayes, Nussinov and Starr [15, 14] (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 [83]. The conclusions hold also the 2D quantum orbital-compass model with large quantum spins [15]. 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 [11]. This model actually goes back to the papers by Shender [93] and Henley [63] which first spelled out the original order-by-disorder physics arguments.

All three “phase coexistence” theorems in Sect. 6.4 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 [13, 11] 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 [16]; the non-existence clauses in Theorems 6.11, 6.13 and 6.14 are direct consequences of the main result of [16] 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.

## 7 Topics Not Covered

There are naturally many interesting topics dealing with reflection positivity that have not been covered by these notes. Here we will attempt to at least provide a few relevant comments and give pointers to the literature where an interested reader may explore the subject to the desired level of detail.

The first (and large) area which was neglected is that of *quantum* models. Here one faces the principal difficulty that the spin variables

are replaced by operators which, generally, do not commute with one another. Nevertheless, reflection positivity can be proved for reflections through planes between sites under the condition that the Hamiltonian is of the form (5.10). (For reflections through planes of sites the non-commutativity of involved objects makes the above technology largely unavailable.) Thus, chessboard estimates and, by a passage via the Duhamel two-point function, also infrared bound can again be established. This and the resulting applications to proofs of phase transitions in, e.g., the quantum Heisenberg *antiferromagnet* and XY-model constitute the papers of Dyson, Lieb and Simon [38] and Fröhlich and Lieb [48]. A pedagogical account of these can be found in the notes by Tóth [103].

Unlike for the classical models, in the quantum setting reflection positivity appears to be a somewhat peculiar condition. Generally, it requires that the involved operators can be represented by either real or purely imaginary matrices. This is where the technique fails in the case of the quantum Heisenberg ferromagnet (Speer [99], but see also Kennedy [67] and Conlon and Solovej [27]). Notwithstanding, the technique continued to be applied in the quantum world to derive useful conclusions; e.g., to study long range order in two-dimensional antiferromagnets (Kennedy, Lieb and Shastry [68]), to resolve the so called flux phase problem in the Hubbard model (Lieb [78], see also Macris and Nachtergaele [81]) or to prove uniqueness of the ground state in the half-filled band therein (Lieb [77]). The latter work invokes *spin-reflection positivity*; a new idea later further exploited by, e.g., Tian [102] and Tasaki [101]. Other applications of reflection positivity in itinerant-electron models appear in, e.g., Macris [79] and Macris and Lebowitz [80].

As already mentioned, one can use RP to develop a rigorous link between the phase transitions in quantum and classical systems (Biskup, Chayes and Starr [15]). Here the main idea is the conversion of the quantum chessboard estimate to the classical one by means of an extension of Berezin-Lieb inequalities to matrix elements in the basis of coherent states.

Another topic not sufficiently represented in these notes is that of dimer or other combinatorial models. Here we wish to mention, e.g., the conclusions concerning the six-vertex model and hard-core lattice gasses (Fröhlich, Israel, Lieb and Simon [47]) or the liquid-crystal models based on interacting dimers (Heilmann and Lieb [62] and Abraham and Heilmann [1]). There is also a novel application to characterization of graph homomorphisms (Freedman, Lovász and Schrijver [44]).

The origin of reflection positivity lies within the field theory as part of the Osterwalder-Schrader axioms. A reader interested in this direction should employ the relevant search outlets to explore the literature on the subject. For statistical mechanics, interesting applications come in the proofs of phase transitions in Euclidean field theories, e.g., that of quark confinement (Borgs and Seiler [20]) or chiral symmetry breaking (Salmhofer and Seiler [91]) in gauge theories.

Finally, there is the recent clever application of chessboard estimates to control the rigidity of Dobrushin interfaces in the Ising (and some other) three dimensional models (Shlosman and Vignaud [96]). This direction will likely be further exploited to study interface states in continuum-spin systems.

## 8 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 3 and 4 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 8.1.** 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 [90], who managed to apply the lace expansion to a modified random current representation of the Ising model. However, 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

Among approaches in this direction is the *spherical approximation* for the  $O(n)$  model, in which the constraint  $|S_x| = 1$  at every spin is replaced by a constraint on  $\sum_x |S_x|^2$ .

The IRB is often viewed as a rigorous version of *spin-wave theory*. This theory, initiated in the work of Dyson [37] and others, describes continuous deformations of the lowest energy states by means of an appropriate Gaussian field theory. In Chapter 6 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 8.2.** Prove symmetry breaking at low temperatures in systems with continuous internal symmetry — e.g., the  $O(2)$ -model — without the use of the IRB. Chessboard estimates are allowed.

An interesting resource for thinking about this problem may be the paper of Bricmont and Fontaine [21].

Further motivation to look at this problem comes from quantum theory: The quantum Heisenberg ferromagnet is not RP (see Speer [99]) 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. However, if we had a more robust proof of symmetry breaking in the classical model, e.g., using chessboard estimates, one might hope to extend the techniques of Biskup, Chayes and Starr [15] 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 antiferromagnets, e.g., hard core lattice gas, which is a model with variables  $n_x \in \{0, 1\}$  and the “Gibbs” weight proportional to

$$\lambda^{\sum_x n_x} \prod_{\langle x, y \rangle} (1 - n_x n_y), \quad (8.1)$$

or the  $q$ -state Potts antiferromagnet, which is the model in (2.7) 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 8.3.** 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 [54, 55].

## References

1. D.B Abraham and O.J. Heilmann, *Interacting dimers on the simple cubic lattice as a model for liquid crystals*, J. Phys. A: Math. Gen. **13** (1980) 1051–1062.
2. M. Aizenman, *Geometric analysis of  $\phi^4$  fields and Ising models. I, II.*, Commun. Math. Phys. **86** (1982) 1–48.
3. M. Aizenman, J.T. Chayes, L. Chayes and C.M. Newman, *Discontinuity of the magnetization in one-dimensional  $1/|x - y|^2$  Ising and Potts models*, J. Statist. Phys. **50** (1988), no. 1-2, 1–40.
4. M. Aizenman and R. Fernández, *On the critical behavior of the magnetization in high-dimensional Ising models*, J. Statist. Phys. **44** (1986) 393–454.
5. M. Aizenman and R. Fernández, *Critical exponents for long-range interactions* Lett. Math. Phys. **16** (1988), no. 1, 39–49.
6. 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.
7. M. Biskup, *Reflection positivity of the random-cluster measure invalidated for non-integer  $q$* , J. Statist. Phys. **92** (1998) 369–375.
8. M. Biskup, C. Borgs, J.T. Chayes and R. Kotecký, *Gibbs states of graphical representations of the Potts model with external fields*, J. Math. Phys. **41** (2000), no. 3, 1170–1210.
9. 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.
10. 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.
11. 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.
12. 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.
13. M. Biskup, L. Chayes, and Z. Nussinov, *Orbital ordering in transition-metal compounds: I. The 120-degree model*, Commun. Math. Phys. **255** (2005), no. 2, 253–292.
14. M. Biskup, L. Chayes and Z. Nussinov, *Orbital ordering in transition-metal compounds: II. The orbital compass model*, in preparation.
15. M. Biskup, L. Chayes and S. Starr, *Quantum spin systems at positive temperature*, Commun. Math. Phys. **269** (2007), no. 3, 611–657
16. 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.



17. M. Biskup and R. Kotecký, *Phase coexistence of gradient Gibbs states*, Probab. Theory Rel. Fields **139** (2007), no. 1-2, 1–39.
18. M. Biskup and H. Spohn, *Scaling limit for a class of gradient fields with non-convex potentials*, Ann. Probab. (to appear)
19. T. Bodineau, *Translation invariant Gibbs states for the Ising model*, Probab. Theory Related Fields **135** (2006), no. 2, 153–168.
20. C. Borgs and E. Seiler, *Quark deconfinement at high temperature: a rigorous proof*, Nucl. Phys. B **215** (1983), no. 1, 125–135.
21. J. Bricmont and J.-R. Fontaine, *Infrared bounds and the Peierls argument in two dimensions*, Commun. Math. Phys. **87** (1982/83), no. 3, 417–427.
22. 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.
23. R.M. Burton and M. Keane, *Density and uniqueness in percolation*, Commun. Math. Phys. **121** (1989), no. 3, 501–505.
24. L. Chayes, *Mean field analysis of low dimensional systems*, Commun. Math. Phys. (to appear)
25. L. Chayes, R. Kotecký, and S. B. Shlosman. *Staggered phases in diluted systems with continuous spins*, Commun. Math. Phys. **189** (1997) 631–640.
26. L. Chayes and J. Machta, *Graphical representations and cluster algorithms. Part I: Discrete spin systems*, Physica A **239** (1997) 542–601.
27. J.G. Conlon and J.P. Solovej, *On asymptotic limits for the quantum Heisenberg model*, J. Phys. A **23** (1990), no. 14, 3199–3213.
28. 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.
29. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (Springer Verlag, Inc., New York, 1998).
30. J.-D. Deuschel and D.W. Stroock, *Large deviations*, Pure and Applied Mathematics, 137. Academic Press, Inc., Boston, MA, 1989.
31. J. Dimock and T.R. Hurd, *Sine-Gordon revisited*, Ann. Henri Poincaré **1** (2000), no. 3, 499–541.
32. 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.
33. R.L. Dobrushin and S.B. Shlosman, *Phases corresponding to minima of the local energy*, Selecta Math. Soviet. **1** (1981), no. 4, 317–338.
34. 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.
35. 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.

36. R. Durrett, *Probability: Theory and Examples*, Second edition. Duxbury Press, Belmont, CA, 1996.
37. F.J. Dyson, *General theory of spin-wave interactions*, Phys. Rev. **102** (1956), no. 5, 1217–1230.
38. 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.
39. 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.
40. R.S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Grundlehren der Mathematischen Wissenschaften, vol. 271 (Springer-Verlag, New York, 1985).
41. 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.
42. 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.
43. 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.
44. M. Freedman, L. Lovász and A. Schrijver, *Reflection positivity, rank connectivity, and homomorphism of graphs*, J. Amer. Math. Soc. **20** (2007), no. 1, 37–51.
45. J. Fröhlich, *On the triviality of  $\lambda\varphi_d^4$  theories and the approach to the critical point in  $d \binom{>}{>} 4$  dimensions*, Nucl. Phys. B **200** (1982), no. 2, 281–296.
46. 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.
47. 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.
48. J. Fröhlich and E.H. Lieb, *Phase transitions in anisotropic lattice spin systems*, Commun. Math. Phys. **60** (1978), no. 3, 233–267.
49. 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.
50. J. Fröhlich, B. Simon and T. Spencer, *Infrared bounds, phase transitions and continuous symmetry breaking*, Commun. Math. Phys. **50** (1976) 79–95.
51. J. Fröhlich and T. Spencer, *The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas*, Commun. Math. Phys. **81** (1981), no. 4, 527–602.

52. T. Funaki, *Stochastic Interface Models*, Lecture Notes for the International Probability School at Saint-Flour, Lecture Notes in Math., 1869, Springer, Berlin, 2005.
53. T. Funaki and H. Spohn, *Motion by mean curvature from the Ginzburg-Landau  $\nabla\phi$  interface model*, Commun. Math. Phys. **185** (1997) 1–36.
54. 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.
55. 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.
56. P.G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Oxford University Press, New York, 1993.
57. H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter Studies in Mathematics, vol. 9 (Walter de Gruyter & Co., Berlin, 1988).
58. T. Gobron and I. Merola, *First-order phase transition in Potts models with finite-range interactions*, J. Statist. Phys. **126** (2007), no. 3, 507–583.
59. R.B. Griffiths, *Peierls proof of spontaneous magnetization in a two-dimensional Ising ferromagnet*, Phys. Rev. (2) **136** (1964) A437–A439.
60. R.B. Griffiths, *Correlations in Ising ferromagnets*, J. Math. Phys. **8** (1967), no. 3, 478–483.
61. G. Grimmett, *The Random-Cluster Model*, Grundlehren der Mathematischen Wissenschaften, vol. 333, Springer-Verlag, Berlin, 2006
62. O.J. Heilmann and E.H. Lieb, *Lattice models for liquid crystals*, J. Statist. Phys. **20** (1979), no. 6, 679–693.
63. C.L. Henley, *Ordering due to disorder in a frustrated vector antiferromagnet*, Phys. Rev. Lett. **62** (1989) 2056–2059.
64. F. den Hollander, *Large deviations*, Fields Institute Monographs, 14. American Mathematical Society, Providence, RI, 2000.
65. 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.
66. R.B. Israel, *Convexity in the theory of lattice gases*, With an introduction by Arthur S. Wightman. Princeton University Press, Princeton, N.J., 1979.
67. T. Kennedy, *Long range order in the anisotropic quantum ferromagnetic Heisenberg model*, Commun. Math. Phys. **100** (1985), no. 3, 447–462.
68. T. Kennedy, E.H. Lieb and B.S. Shastry, *Existence of Néel order in some spin- $\frac{1}{2}$  Heisenberg antiferromagnets*, J. Statist. Phys. **53** (1988), no. 5-6, 1019–1030,
69. H. Kesten and R. Schonmann, *Behavior in large dimensions of the Potts and Heisenberg models*, Rev. Math. Phys. **1** (1990) 147–182.

70. M. Kosterlitz and D.J. Thouless, *Ordering, metastability and phase transitions in two-dimensional systems*, J. Phys. C: Solid State Phys., **6** (1973) 1181–1203.
71. 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.
72. R. Kotecký and S.B. Shlosman, *First-order phase transitions in large entropy lattice models*, Commun. Math. Phys. **83** (1982), no. 4, 493–515.
73. U. Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, 1985.
74. C. Külske, *The continuous spin random field model: Ferromagnetic ordering in  $d \geq 3$* , Rev. Math. Phys. **11** (1999), no. 10, 1269–1314.
75. C. Külske, *Stability for a continuous SOS-interface model in a randomly perturbed periodic potential*, WIAS Preprint no. 466 (1998); <http://www.math.rug.nl/~kuelske/publications.html>
76. G.L. Lawler, *Intersections of random walks*, Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1991.
77. E.H. Lieb, *Two theorems on the Hubbard model*, Phys. Rev. Lett. **62** (1989), no. 10, 1201–1204. Errata: Phys. Rev. Lett. **62** (1989), no. 16, 1927.
78. E.H. Lieb, *Flux phase of the half-filled band*, Phys. Rev. Lett. **73** (1994) 2158–2161.
79. N. Macris, *Periodic ground states in simple models of itinerant fermions interacting with classical fields*, The nature of crystalline states (Kudowa-Zdrój, 1995). Phys. A **232** (1996), no. 3–4, 648–656.
80. N. Macris and J. Lebowitz, *Ground states and low-temperature phases of itinerant electrons interacting with classical fields: a review of rigorous results*, Quantum problems in condensed matter physics, J. Math. Phys. **38** (1997), no. 4, 2084–210.
81. N. Macris and B. Nachtergaele, *On the flux phase conjecture at half-filling: an improved proof*, J. Statist. Phys. **85** (1996), no. 5–6, 745–761.
82. 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.
83. 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.
84. K. Osterwalder and R. Schrader, *Axioms for Euclidean Green's functions*, Commun. Math. Phys. **31** (1973), 83–112.
85. A. Patrasciou and E. Seiler, *Percolation theory and the existence of a soft phase in 2D spin models*, Nucl. Phys. B (Proc. Suppl.) **30** (1993) 184–191.

86. K. Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, 2. Cambridge University Press, Cambridge, 1989.
87. C.-E. Pfister, *On the symmetry of the Gibbs states in two-dimensional lattice systems*, Commun. Math. Phys. **79** (1981), no. 2, 181–188.
88. D. Ruelle, *Thermodynamic formalism. The mathematical structures of equilibrium statistical mechanics*, Second edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 200
89. D. Ruelle, *Statistical mechanics. Rigorous results*, Reprint of the 1989 edition, World Scientific Publishing Co., Inc., River Edge, NJ; Imperial College Press, London, 1999.
90. A. Sakai, *Lace expansion for the Ising model*, Commun. Math. Phys. **272** (2007), no. 2, 283–344.
91. M. Salmhofer and E. Seiler, *Proof of chiral symmetry breaking in strongly coupled lattice gauge theory*, Commun. Math. Phys. **139** (1991), no. 2, 395–432. Errata: *ibid* **146** (1992), no. 3, 637–638.
92. S. Sheffield, *Random Surfaces*, Asterisque 2005, No. 304, 177pp.
93. E.F. Shender, *Antiferromagnetic garnets with fluctuationally interacting sublattices*, Sov. Phys. JETP **56** (1982) 178–184.
94. S. Shlosman, *Phase transitions for two-dimensional models with isotropic short range interactions and continuous symmetry*, Commun. Math. Phys. **71** (1980) 207–212.
95. 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.
96. S. Shlosman and Y. Vignaud, *Dobrushin interfaces via reflection positivity*, Commun. Math. Phys. **276** (2007), no. 3, 827–86.
97. B. Simon, *The statistical mechanics of lattice gases*, Vol. I., Princeton Series in Physics (Princeton University Press, Princeton, NJ, 1993).
98. A.D. Sokal, *An alternate constructive approach to the  $\varphi_3^4$  quantum field theory, and a possible destructive approach to  $\varphi_4^4$* , Ann. Inst. H. Poincaré Sect. A **37** (1982), no. 4, 317–398.
99. E.R. Speer, *Failure of reflection positivity in the quantum Heisenberg ferromagnet*, Lett. Math. Phys. **10** (1985), no. 1, 41–47.
100. F. Spitzer, *Principles of random walks*, Second edition. Graduate Texts in Mathematics, Vol. 34. Springer-Verlag, New York-Heidelberg, 1976.
101. H. Tasaki, *Ferromagnetism in the Hubbard model: a constructive approach*, Commun. Math. Phys. **242** (2003), no. 3, 445–472.
102. G.-S. Tian, *Lieb’s spin-reflection-positivity method and its applications to strongly correlated electron systems*, J. Statist. Phys. **116** (2004), no. 1-4, 629–680.
103. B. Tóth’s website:  
[www.math.bme.hu/~balint/oktatas/statisztikus\\_fizika/jegyzet/](http://www.math.bme.hu/~balint/oktatas/statisztikus_fizika/jegyzet/)
104. Y. Velenik, *Localization and delocalization of random interfaces*, Probab. Surveys **3** (2006) 112–169.

105. P. Weiss, *L'hypothèse du champ moléculaire et la propriété ferromagnétique*, J. de Physique **6** (1907) 661–689.
106. F.Y. Wu, *The Potts model*, Rev. Modern Phys. **54** (1982) 235–268.
107. 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.