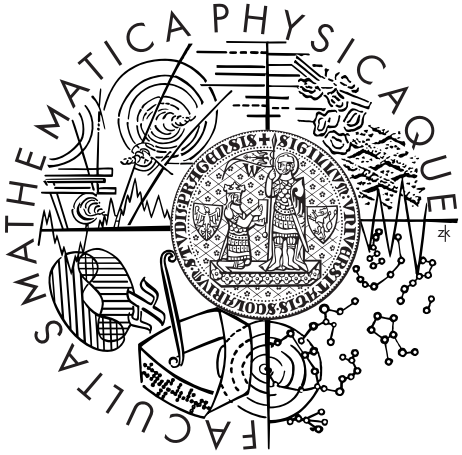
Charles university in Prague
Faculty of Mathematics and Physics

MASTER THESIS



Marek Bernát

Gradientní modely
Gradient models

Institute of Theoretical Physics

Supervisor of the master thesis: prof. RNDr. Kotecký Roman, DrSc.
Study programme: Physics
Specialization: Theoretical physics

Prague 2012

Acknowledgements.

I would like to thank my advisor Roman Kotecký without whose undiminishing support this thesis would not exist. I express my deep gratitude to Iva Hromková who has seen to it that this thesis will be finished.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In ..... date ..... signature

Název práce: Gradientní modely

Autor: Marek Bernát

Katedra: Ústav teoretické fyziky

Vedoucí diplomové práce: prof. RNDr. Kotecký Roman, DrSc.

Abstrakt: Vyšetrovali sme gradientné modely, jednak model s double-well potenciálom a jednak tzv. rozšírený model. V dimenzii 2 sme v rozšírenom modele exaktne spočítali voľné energie roztrúsenných hranových konfigurácií a pre ľubovoľnú dimenziu sme získali odhady týchto voľných energií. Kombináciou týchto odhadov s argumentom o existencii zlých kontúrov a odhadom počtu kontúrov sme boli metódou reflekčnej pozitivity schopní ukázať, že pri nízkych teplotách existuje v rozšírenom modele fázový prechod. Ďalej sme ukázali, že fázový prechod existuje aj v double-well modele za podmienky, že platí hypotéza týkajúca sa odhadov strednej energie. Okrem týchto výsledkov je v práci prezentovaný aj základný aparát štatistickej fyziky a fakty z príbuzných oblastí, a tiež základné výsledky z oblasti gradientných modelov, takže práca zároveň slúži aj ako úvod do tejto problematiky.

Klíčová slova: Modely na mrieži, gradientné modely, nelineárna elasticita, Gibbsovské stavy

Title: Gradient models

Author: Marek Bernát

Department: Institute of Theoretical Physics

Supervisor: prof. RNDr. Kotecký Roman, DrSc.

Abstract: We have investigated gradient models, one of them was a model with double-well potential and the other one a so called extended model. In dimension two we have calculated exact free energies of the disseminated edge configurations for the extended model and for arbitrary dimension we have derived bounds on these free energies. Combining these bounds with an argument on exstince of bad contours together with the estimate of the number of these contours and using the method of reflection positivity we have been able to show that at low temperatures there is a phase transition in the extended model. We have further shown that the phase transition exists also in the double-well model as long as a conjecture on estimates of mean energy holds. Besides these results the thesis also contains basic tools of statistical physics and facts from related fields, as well as basic results on gradient models, so that our work can serve as an introduction into these areas.

Keywords: Lattice models, gradient models, nonlinear elasticity, Gibbs states

Contents

Introduction

In this thesis we study gradient models – models that do not depend on the absolute values of the field but only on the differences at neighboring sites. We will see that these models describe interfaces between multiple thermodynamic phases. After a brief review of known results we introduce a new model with double-well potential. Using the tools of modern statistical physics – in particular the reflection positivity – we will show that this model contains a phase transition of the second order at low temperatures.

In chapter one we will introduce the general theory of Gibbs measures which is necessary for the formal treatment of phase transitions.

In chapter two we will derive the basic properties of the torus and the group of its symmetries generated by reflections through the hyperplanes. Using these results we will then introduce the notion of reflection positivity and its all-important consequence – the chessboard estimates.

In chapter three we review gradient models. For gaussian fields and, more generally, for fields with convex potential there can be no phase transition. We will therefore focus our attention on non-convex models and discuss basic results in this area. Then we will introduce the model that is the subject of this work – the double-well model – and another extended model that will help simplify the analysis of the double-well model. We will also investigate groundstates of these models and verify that the extended model is reflection positive.

In chapter four we will perform the calculation of the free energy of the double-well model which will be necessary for the use of chessboard estimates. It will turn out that it is possible to completely solve the problem using Fourier analysis. Nevertheless, the closed-form solution gets unwieldy in higher dimensions and therefore we split the discussion. In two dimensions we provide an exact solution while in arbitrary dimension we only obtain bounds on the free energy.

In chapter five we first outline the strategy for the proof of coexistence. We then establish basic results on geometric objects called contours, especially bounds on their number in arbitrary dimension. Using these results together with chessboard estimates and the bounds on free energy obtained in chapter four are able to show the presence of two phases in the extended model on the torus and, conjecturally, also for the double-well model.

In chapter six we will introduce the models on infinite lattices and show how the results of chapter five imply the existence of two distinct infinite volume phases.

1. General theory

The fundamental notion of the classical statistical mechanics is that of a Gibbs measure. We know that in equilibrium the probability of occurrence of individual microstate with energy E is proportional to $\exp(-\beta E)$ where β is the inverse temperature of the heat bath. This definition, while simple and appealing, is useless in the infinite volume setting where the energy of the system can become infinite. Nevertheless, it will be necessary for us to deal with infinite volumes since it is only there that the features of phase transitions are most apparently pronounced in the form non-analiticies of macroscopic observables.

This means some other approach is called for in the infinite volume setting. Still, it would be dubious to forget what we already know about finite systems where the Boltzmann ensembles work so well. We will resolve this problem in a natural manner by obtaining the infinite volume measure as an extension of the corresponding finite volume measures. The prescription of this finite volume behaviour will be formalized in the notion of *specification*.

To introduce all the relevant concepts we will first need to introduce some notation. Reader who is not familiar with basic notions of measure and probability theory is advised to consult Appendix C.

1.1 Basic notions of statistical mechanics on a lattice

All of the events of interest to us will occur in some discrete set S (usually this set will come equiped with an extra structure; for concretness one can think of $\mathbb{Z}^2$). At every point x of this set there will sit a *s*pin taking on the values from the set E_x (the terminology comes from sources such as Ising model where $E_x := \{-1, +1\}$ for all $x \in S$). In this work E_x will always be either $\mathbb{R}$ or some finite set which means that we can actually treat these sets as measure spaces $(E_x, \mathscr{E}_x, \lambda_x)$ in a natural way – when $E_x = \mathbb{R}$ we consider it equipped with its Borel σ-algebra (induced by the standard Euclidean topology) and Lebesgue measure while for E_x finite we take a counting measure on the power set of E_x.

Having introduced these basic building blocks, we can now create objects that will be of more interest to us. We begin with the product σ-algebra $\mathscr{F}_\Lambda := \prod_{x\in\Lambda} \mathscr{E}_x$ on the set $\prod_{x\in\Lambda} E_x$ for any $\Lambda \in S$. We will denote $\mathscr{F}_S$ simply by $\mathscr{F}$ and $\mathscr{F}_{S\setminus\Lambda}$ by $\mathscr{T}_\Lambda$. Let us also denote the joint set of all spins $\prod_{x\in S} E_x$ by Ω and the set of all finite subset of S by $\mathscr{S}$.

Definition 1.1. Specification $(\gamma_\Lambda)_{\Lambda\in\mathscr{S}}$ *is a family of proper probability kernels* γ_Λ *from* $\mathscr{T}_\Lambda$ *to* $\mathscr{F}$ *such that for any* $\Lambda \subset \Delta \subset \mathscr{S}$ *we have* $\gamma_\Delta\gamma_\Lambda = \gamma_\Delta$.

This definition formalizes our intuition about probability measures in finite volume. For any finite Λ and boundary condition ω we obtain a probability measure $\gamma_\Lambda(\cdot|\omega)$ on $\mathscr{F}$ and the collection of this measures is moreover compatible. We will denote by $\mathcal{G}(\gamma)$ the set of probability measures on $\mathscr{F}$ that satisfy $\mu\gamma_\Lambda = \mu$ for every $\Lambda \in \mathscr{S}$ and we will say that μ is admitted by the specification γ. This set of equations expresses the fact that μ conditioned on any finite subset Λ gives a measure that is prescribed by γ_Λ. In other words, μ is an infinite volume extension of the collection of compatible finite volume probability measures, which is precisely what we were aiming for.

In order to investigate the properties of the set $\mathcal{G}(\gamma)$ we will first need to construct some concrete specifications. For any $\Lambda \in \mathscr{S}$ we define a measure kernel from $\mathscr{T}_\Lambda$ to $\mathscr{F}$

$$\lambda_\Lambda(\cdot|\omega) := \left(\prod_{x\in\Lambda} \lambda_x\right) \times \delta_{\omega_{S\setminus\Lambda}} \tag{1.1}$$

It can be checked that the collection $(\lambda_\Lambda)_{\Lambda\in\mathscr{S}}$ forms a specification that we will call a λ-specification. While this notion is crucial to what will follow, it is also completely trivial from the physical point of view since there can be no phase transitions in the absence of interactions.

To remedy this problem we shall introduce the physical notions of potential and energy of the system.

Definition 1.2. Potential Φ *is a collection* $(\Phi_A)_{A\in\mathscr{S}}$ *such that for every* $A \in \mathscr{S}$ *the* $\Phi_A : \Omega \to \mathbb{R}$ *be a* $\mathscr{F}_A$*-measurable function.*

Using this notion we can define the *energy* (also known as the *Hamiltonian*) for a finite volume Λ under boundary condition $\omega_{S\setminus\Lambda}$ interacting by means of the potential Φ

$$H^\Phi_\Lambda(\omega) := \sum_{A\in\mathscr{S}, A\cap\Lambda\neq\emptyset} \Phi_A(\omega) \tag{1.2}$$

Finally, we arrive at the Boltzmann probabilities $h^\Phi_\Lambda(\cdot) := \exp(-H^\Phi_\Lambda(\cdot))$. We say that the potential Φ is λ-admissible if $\lambda_\Lambda h^\Phi_\Lambda < \infty$ for every $\Lambda \in \mathscr{S}$. If this is the case we introduce the *partition sum* $Z^\Phi_\Lambda := \lambda_\Lambda h^\Phi_\Lambda$. We can now define the *Gibbsian specification* with potential Φ as

$$(\gamma^\Phi_\Lambda)_{\Lambda\in\mathscr{S}} := (h^\Phi_\Lambda/Z^\Phi_\Lambda)\lambda_\Lambda \tag{1.3}$$

The set of measures $\mathcal{G}(\gamma^\Phi)$ admitted by this specification will be denoted by $\mathcal{G}(\Phi)$ and the admitted measures will be called *Gibbs measures*.

1.2 Construction of Gibbs measures

The most natural way of constructing Gibbs measures is as a limit in the space of probability measures $\mathscr{P}(\Omega, \mathscr{F})$ on the measurable space $(\Omega, \mathscr{F})$.

We will therefore need to deal with both existence of the limit and also with deciding whether this limit is actually a Gibbs measure for the specification wanted. These problems put conflicting demands on the topology on $\mathscr{P}(\Omega, \mathscr{F})$ – there are more cluster points when the topology is coarse but at the same time the topology needs to be fine enough so that the cluster point is really a Gibbs measure. It turns out that there is a convenient choice of topology – so called $\mathscr{L}$-topology – that works reasonably well to solve both of these problems. We will now state the theorem (Corollary 4.13 in [1]) that addresses the issue of existence. Note that Georgii works in a slightly different setting of using nets instead of sequences and of specializing $(E_x, \mathscr{E}_x, \lambda_x) \equiv (E, \mathscr{E}, \lambda)$ for all $x \in S$. These minor differences do not change anything about the validity of these results.

Theorem 1.3. *Suppose $(E_x, \mathscr{E}_x)$ is standard Borel for all $x \in S$. Let $(\Phi^n)_{n\in\mathbb{N}}$ be a sequence of λ-admissible potentials, $(\Lambda_n)_{n\in\mathbb{N}}$ a sequence in $\mathscr{S}$ with $\Lambda_n \to S$, $(\nu_n)_{n\in\mathbb{N}}$ a sequence in $\mathscr{P}(\Omega, \mathscr{F})$ and $\mu_n = \nu_n \gamma_{\Lambda_n}^{\Phi^n}$ for all $n \in \mathbb{N}$. Suppose that for all $x \in S$ there is a sequence $(K_x^l)_{l\in\mathbb{N}} \in \mathscr{E}_x$ such that the following conditions are satisfied:*

- *$0 < \lambda_x(K_x^l) < \infty$ for all $x \in S, l \in \mathbb{N}$*
- *$\lim_{l\to\infty} \limsup_{n\in\mathbb{N}} \mu_n(\sigma_x \notin K_x^l) = 0$ for all $x \in S$; and*
- *each $\Lambda \in \mathscr{S}$ is contained in some $\Delta \in \mathscr{S}$ such that for all $l \in \mathbb{N}$*

$$\limsup_{n\in\mathbb{N}} \sup_{\omega \in \prod_{x\in\Delta} K_x^l \times \prod_{x\in S\setminus\Delta} E_x} |H_\Lambda^{\Phi^n}(\omega)| < \infty.$$

Then the sequence $(\mu_n)_{n\in\mathbb{N}}$ has a cluster point.

2. Torus and reflection positivity

2.1 Torus and cubical complex

We define a *continuous torus* T_L^d as a quotient space of $\mathbb{R}^d$ by a lattice $(L\mathbb{Z})^d$. More precisely, define a relation $\sim$ on $\mathbb{R}^d$ with $x \sim y$ if there exists $z \in \mathbb{Z}^d$ such that $x = y + L \cdot z$. This is an equivalence relation since $\mathbb{Z}^d$ is a group. We then define the continuous torus (or simply *torus* when there is no possibility of confusion with $\mathbb{T}_L^d$) as $\mathbb{R}^d/\sim$.

Next we proceed to the definition of the cubical complex. For that we will first need a notion of *index function* $I : \{1, \ldots, d\} \to \{0, 1\}$ and its sum $|I| = \sum_k I(k)$. Then we define an *m-cube $u(X, I)$ with lowest corner $X \in \mathbb{Z}^d$ and index function* I, $|I| = m$ by

$$u(X, I) \equiv \{x \in \mathbb{R}^n \,|\, X_j \leq x_j \leq X_j + I(j)\}\,.$$

Let U be any set of subspaces of T_L^d. We will write $[U] := \bigcup_{u \in U} u$ when interpretting U as a subset of T_L^d. A set U of m-cubes will be called an *m-skeleton*. Arbitrary set of cubes (without regard to their dimension) will be called a *cubical complex*.

Since for an m-skeleton U the space $[U]$ is naturally a topological manifold (with boundary) we can introduce the operation of taking a boundary ∂. In terms of cubes, this takes the m-skeleton U to an $(m-1)$-skeleton V such that V is a union of all $(m-1)$-cubes that are contained in one and only one m-cube of U. We will write $V = \partial U$ so that this boundary operation on sets agrees with the boundary operation on spaces – $[\partial U] = \partial [U]$. It is straightforward to check that $\partial^2 = 0$. In other words $\partial[U]$ is a manifold without boundary.

Let V be the collection of 0-cubes of T_L^d and E be the collection of oriented 1-cubes of T_L^d. We will call the graph $G = (V, E)$ a *discrete torus* and denote it by $\mathbb{T}_L^d$.

Given a d-skeleton U, let $E(U)$ be the set of oriented 1-cubes contained in it. We will call any function $\xi : E(U) \to \{-1, 1\}$ such that $\xi(-e) = -\xi(e)$ an *edge configuration on U*. The space of all such configurations on U will be denoted $K(U)$.

2.2 Reflectional symmetries

For the rest of the discussion suppose L is even. For any $i, 1 \leq i \leq m$ and any $0 \leq k < L/2$ we define the hyperplane

$$P_{(i,k)} := \{x \in T_L^d | x_i = k \text{ or } x_i = k + L/2\}.$$

Any such hyperplane P defines a reflection on T_L^d

$$\theta_P(x) \equiv x + 2(k - x_i)\mathbf{e}_i \pmod{L}.$$

Also denote by τ_y the translation

$$\tau_y(x) \equiv (x + y) \pmod{L}.$$

In the following $G = (V, E)$ will stand for a graph with vertex set consisting of d-cubes and edges between those d-cubes that share a common $(d-1)$-cube. For any such $e = (c, d) \in E$ we will denote by P_e the hyperplane containing $c \cap d$ and we will write θ_e in place of θ_{P_e}.

Lemma 2.1. *For all $e, f \in E$ such that P_e is not parallel with P_f it holds $\theta_e\theta_f = \theta_f\theta_e$. If P_e is parallel with P_f we have that $\theta_e\theta_f = \tau_{m\mathbf{e}_i}$ for some $1 \leq i \leq d$ and $0 \leq m < L$ and if moreover $e_e = f_b$ then $\theta_e\theta_f = \tau_{2\mathbf{e}_i}$.*

Proof. This follows directly from the definition of reflection since there are some $1 \leq i \leq d$, $1 \leq j \leq d$ and $0 \leq k, l < L/2$ such that for any $(x_1, \cdots, x_d) \in \mathbb{T}_L$ we have

$$\theta_e(x) \equiv x + 2(k - x_i)\mathbf{e}_i \pmod{L}$$

and

$$\theta_f(x) \equiv x + 2(l - x_j)\mathbf{e}_j \pmod{L}$$

so that we see that the order of these operations does not matter for $i \neq j$ while for $i = j$ we get

$$\theta_e\theta_f x \equiv x + 2(l - x_i)\mathbf{e}_i + 2(k - (2l - x_i)))\mathbf{e}_i \equiv x + 2(k - l)\mathbf{e}_i \pmod{L}. \quad \square$$

From this lemma it follows that the group $\mathcal{R}$ generated by the said reflections is isomorphic to the d-fold product of dihedral groups $(D_L)^d$ [1].

Now let z be a d-cube with lower right corner at origin. An edge configuration defined on $\{z\}$ will be called an *elementary configuration.*

Definition 2.2. *Let a elementary configuration ξ be given. We define a* disseminated configuration $D(\xi) \in K(\mathbb{T}_L)$ *by*

$$(D\xi)(e) = \xi(\rho(P)^{-1}e)$$

where

$$\rho(P) := \prod_{i=1}^{n} \theta_{(v_i, v_{i+1})}$$

and $P = (v_i)_{i=1}^n$ is any path in G from z to a d-cube x containing e.

[1] Recall that the dihedral group D_k is the group of symmetries of a regular k-gon.

We will need to show that this definition does not depend on the path P. To this end we will first need to establish basic facts about the paths in G and the corresponding transformations generated by them.

Lemma 2.3. *Let L be a loop in G. Then $\rho(L) = \mathrm{Id}$*

Proof. Suppose $d = 1$. First assume the loop is of the form

$$L = \underbrace{L' + \cdots + L'}_{n \text{ times}}$$

for some $n \in \mathbb{N}$ and L' the loop that wraps around the torus . By lemma 2.1 for every pair of consecutive edges we get a translation and so

$$\rho(L) = \rho(L')^n = \left(\tau_{2\mathbf{e}_1}^{L/2}\right)^n = \mathrm{Id}^n = \mathrm{Id}$$

Arbitrary loop is composed of parts of the above type and parts of the form $P - P$ where P is a path of consecutive edges which contribute a factor of

$$\rho(P - P) = \rho(P)^{-1} \circ \rho(P) = \mathrm{Id}.$$

From the above the claim in arbitrary dimension d follows since the projection L^i of L to the subgroup of $\mathbb{T}_L$ generated by $\mathbf{e}_i$ for some $1 \leq i \leq d$ is again a loop and by lemma (2.1) we can reorder the factors in the product in the statement as

$$\rho(L) = \prod_{i=1}^{d} \rho(L^i) = \prod_{i=1}^{d} \mathrm{Id} = \mathrm{Id}$$

where the second equality follows from the $d = 1$ part of the proof. □

Corollary 2.4. *Suppose an edge e is given and P_1, P_2 are two paths such that both $\rho(P_1)z$ and $\rho(P_2)z$ contain e. Then $\rho(P_1)e = \rho(P_2)e$.*

Proof. The d-cubes $\rho(P_1)z$, $\rho(P_2)z$ share an edge. Using at most $d-1$ reflections through hyperplanes containing e we can map one cube into the other and therefore we can assume both paths terminate in the same d-cube. But then $P_1 - P_2$ forms a loop and it follows that $\mathrm{Id} = \rho(P_1 - P_2) = \rho(P_1) \circ \rho(P_2)^{-1}$. □

Because of this result we see that reflecting from cube z to cube x does not depend on the path between the cubes and we will write $\theta_x := \rho(P)$, P being any path from z to x.

Corollary 2.5. *Let ξ be an elementary configuration. Then $D(\xi)$ is reflection invariant – for any hyperplane P and any edge of T it holds that $(D(\xi))(\theta_P e) = (D(\xi))(e)$.*

Proof. Let x be a d-cube containing e and choose paths P_1 from z to x and P_2 from z to $\theta_P x$. Also pick any path P_3 from x to $\theta_P x$ such that $\rho(P_3) = \theta_P$. Then $P_1 + P_3 - P_2$ forms a loop which means $\rho(P_1)^{-1} = \rho(P_3 - P_2)$ and therefore

$$(D(\xi))(\theta_P e) = \xi\left((\rho(P_2)^{-1} \circ \rho(P_3))(e)\right) = \xi(\rho(P_1)^{-1} e) = (D(\xi))(e).$$

□

2.3 Reflection positivity

Pick a hyperplane P in the torus and let θ denote the associated reflection through this hyperplane. This reflection partitions the torus into three subsets $\mathbb{T}_L^+$, $\mathbb{T}_L^-$ and $\mathbb{T}_L^0$ such that $\theta(\mathbb{T}_L^\pm) = \mathbb{T}_L^\mp$ and θ acts trivially on $\mathbb{T}_{L^0}$.

This operation has a natural extension to the algebra of random variables on $\mathscr{F}_{\mathbb{T}_L}$. Here the algebra is defined over both vertices and edges. For the fields living on vertices the extension is given as $(\theta f)(\omega_x) := f(\omega_{\theta(x)})$ and for the fields on the edges as $(\theta f)(\kappa_e) := f(\kappa_{\theta(e)})$. In particular, denoting by $U^\alpha := \mathscr{F}_{\mathbb{T}_L^\alpha}$ where α is one of $+$, $-$ or 0 we get $\theta(U^\pm) = \theta(U^\mp)$ and $\theta(U^0) = U^0$.

Definition 2.6. *A measure μ on $\mathscr{F}_{\mathbb{T}_L}$ is said to be* reflection positive *with respect to θ if for every $f, g \in U^+$ it satisfies*

- $\mathbb{E}_\mu(f\theta f) \geq 0$
- $\mathbb{E}_\mu(f\theta g) = \mathbb{E}_\mu(g\theta f)$

If the measure is reflection positive with respect to reflections through any hyperplane we will say it is *reflection positive* (without further qualifications).

Lemma 2.7. *Let P be a hyperplane in the torus and θ the corresponding reflection. Suppose that the torus Hamiltonian can be written as*

$$-H_L = A + \theta A + \sum_\alpha C_\alpha \theta C_\alpha \tag{2.1}$$

with $A, C_\alpha \in U^+$. Then the torus Gibbs measure μ_L is reflection positive with respect to θ.

We will say that event A is *elementary* if it is $\mathscr{F}_x$-measurable for some d-cube x (this means it only depends on the fields living on x). In the following we write $D(A) := \bigcap_{t\in\mathbb{T}_L} \theta_t(A)$ to denote disseminated event obtained from an elementary event A. In case $A = \{\xi\}$ this coincides with the dissemination operation considered in the previous section.

Lemma 2.8. *Let μ be a reflection positive measure and let $\mathcal{A}$ be a collection of elementary events. Then the following sub-additivity formula for disseminated events holds*

$$\mu\left(D\left(\bigcup_{A\in\mathcal{A}} A\right)\right) \leq \sum_{A\in\mathcal{A}} \mu(D(A)). \tag{2.2}$$

Theorem 2.9. Chessboard estimates. *Let μ be a reflection positive measure and let A be an elementary event. Let R be a subset of $\mathbb{T}_L$ Then*

$$\mu\left(\bigcap_{x\in R} \theta_x(A)\right) = \prod_{x\in R} \mu(D(A))^{1/|\mathbb{T}_L|}. \tag{2.3}$$

Proofs of the above statements (as well as further notes on reflection positivity) can be found in [2].

3. Gradient models

3.1 Preliminaries

3.1.1 Motivation

The main motivation for the investigation of gradient models comes from the study of interfaces between multiple phases.

For concreteness consider the Ising model having spins $+, -$ at every vertex of $\mathbb{Z}^2$ and interactions occuring only between adjacent vertices. Any configuration in this model can be equivalently represented by contours that separate the regions of the like spins. The energy of such a configuration will then be proportional to the sum of the sizes of the contours. For small temperatures in the ferromagnetic regime the spins will prefer to have the same sign or equivalently, most contours will be large. If we make a simplification here and suppose there are precisely two huge regions of alike spins and moreover that the $+$ spins are located mostly in the bottom part of $\mathbb{Z}^2$ while the $-$ spins are located mostly at the top we can regard this scenario as a model for the interface being represented by a height field in $\mathbb{Z}$ (formally one assumes that these regions are induced by boundary conditions with $+$ spins in the bottom part and $-$ spins in the top part). Moreover the length of the interface (that determines the energy of the configuration) only depends on the differences of the height field at neighboring sites. In other words the energy will be a function of the gradient of the field.

In the following we will consider a scalar field $\phi : \mathbb{Z}^d \to \mathbb{R}$, again regarded as a height field and energy given as a function of $\nabla\phi$. Such a model is translation invariant with respect to translations $\phi \mapsto \phi + C$ and so there is no preferred $\phi_0 \in \mathbb{R}$ that the field would concentrate around. For this reason we call ϕ massless.

3.1.2 Convex models

Let a scalar field $\phi : \mathbb{Z}^d \to \mathbb{R}$ be given. We want to study massless models, i.e. models where the interactions only depend on the gradient $\nabla\phi$ of the field. The most basic such model is one where the interactions occur only between nearest neighbors by means of a quadratic potential. Such a field is called Gaussian and the model can be solved completely explicitly. In this case there can be no phase transition. On an intuitive level this can be seen from the fact that the Boltzmann weights come from the factors such as $\exp(-\beta(\nabla x)^2)$ which is invariant with respect to rescaling $x \mapsto \alpha x, \beta \mapsto \frac{1}{\alpha^2}\beta$. But this means that there is no real notion of temperature in the model. More generally Funaki and Spohn have shown in [3] that there can be no phase coexistence whenever the potential is convex.

3.1.3 Non-convex models

In order to study phase transitions in massless models, some form of non-convexity has to be introduced.

In [4] Biskup and Kotecký studied a model where the potential is of the form

$$\exp(-V(y)) = p\exp(-\kappa_O y^2) + (1-p)\exp(-\kappa_D y^2) \tag{3.1}$$

with $\kappa_O \gg \kappa_D$ and were able to show that in this case a phase transition occurs. This potential represents a thin deep well inside a shallow thick well. At small temperatures ($p \sim 1$) gradient of the field will behave as a Gaussian field with coupling κ_O. This means that the gradient of the field will be localized in the thin well. At the transition point $0 < p^* < 1$ the gradient of the field jumps out of the thin well and as $p \to 0$ the field starts to behave as a Gaussian field with coupling κ_D. The phase transition therefore corresponds to a macroscopic change of the characteristic scale of fluctuations.

In our work we will introduce non-convexity by means of a double well potential

$$\exp(-V(y)) = \exp(-K(y-a)^2) + \exp(-K(y+a)^2).$$

To investigate the model, it is natural to first look at its ground states because these, under certain conditions, govern the low-temperature behavior. The minimization of the energy occurs for any ϕ such that $(\nabla\phi)(e) = \pm b$ where e is an edge in $\mathbb{Z}^d$ and $b > 0$ is the minimum of $V - V(y) \geq V(b)$ for all $y \in \mathbb{R}$. Therefore the set of ground states is huge and it would be hard to study the behavior even at the zero temperature. In fact, in two dimensions at zero temperature the model is isomorphic to an ice model, a variant of the six-vertex model.

The situation would simplify if we could reduce the number of ground states and this can indeed be done in a natural way by introducing a further interaction between next-nearest neighbors. Such an interaction will favor the configurations that are nearly constant on both even respectively odd sublattices. Together with the potential V this forces the ground states to be $\phi \equiv C$ on the even sublattice and $\phi \equiv C \pm a$ on the even sublattice.

Because of the translational symmetry $\phi \mapsto \phi + C$ in any gradient model there is an additional subtlety, since we obtain a continuum of phases indexed by the average height of the field. To take care of this nuisance we just fix the height of the field at origin $\phi(0) := 0$.

3.2 Models on the torus

The core of this thesis rests crucially on the fact that we can work on a torus $\mathbb{T}_L$ and since the torus contains only finite number of vertices we can sidestep the infinite volume issues discussed in chapter one and work directly with fields

and Hamiltonians defined on the torus. Once all the necessary work on the torus is done we will define the full model on $\mathbb{Z}^d$ in chapter 6 and show how all its properties follow from the present analysis on the torus.

3.2.1 Double-well model

Let $\omega : \mathbb{T}_L \to \mathbb{R}$ with $\omega(0) = 0$ be a scalar field living on the vertices of the torus. We define an energy of such a field by means of the Hamiltonian

$$H_L(\omega) = \frac{1}{2} \sum_{e \in E(\mathbb{T}_L)} V^{\mathrm{dw}}(\omega_{e_e} - \omega_{e_b}) + \frac{1}{4} \sum_{e,f \in E(\mathbb{T}_L), e_e = f_b, e \perp f} V^{\mathrm{sw}}(\omega_{f_e} - \omega_{e_b}) \tag{3.2}$$

with the double-well potential

$$\exp(-V^{\mathrm{dw}}(y)) := \exp(-K_1(y-a)^2/2) + \exp(-K_1(y+a)^2/2) \tag{3.3}$$

and the single-well potential

$$V^{\mathrm{sw}}(y) := K_2 y^2/2. \tag{3.4}$$

Note that both potentials are symmetric and the factors $\frac{1}{2}$ and $\frac{1}{4}$ in the Hamiltonian appear due to the overcounting coming from working with oriented edges. Also suppose that $K_1 \geq 0$, $K_2 \geq 0$.

Let us now look at the ground states of this model. First introduce the configurations

$$\omega^z(x) := \begin{cases} 0 & \sum_j x_j \equiv 0 \pmod 2 \\ z & \sum_j x_j \equiv 1 \pmod 2 \end{cases} \tag{3.5}$$

that vanish on the even sublattice and are equal to z on the odd sublattice.

We claim that the only two ground states of H_L are $\omega^{\pm b}$ where $b > 0$ is the positive minimum of V^{dw}.

Proposition 3.1. *The Hamiltonian H_L is minimized at $\omega^{\pm b}$ and for $\omega \neq \omega^{\pm b}$ we have that $H_L(\omega) > H_L(\omega^{\pm b})$.*

Proof. The V^{sw} part of the Hamiltonian vanishes for $\omega^{\pm b}$ precisely because these configurations are constant on both even and odd sublattices. As for the V^{dw} part, we know that it is minimized at $y = \pm b$. Since we have $\omega^{\pm b}(e_e) - \omega^{\pm b}(e_b) = \pm b$ for all $e \in E(\mathbb{T}_L)$ the Hamiltonian is necessarily minimized at $\omega^{\pm b}$.

It is also clear from the discussion that all other configurations ω do not minimize the Hamiltonian because they are either not constant on the even/odd sublattices, thereby introducing a non-zero V^{sw} contribution, or else, if they are constant on both sublattices, they are not equal to $\pm b$ on the odd sublattice and introduce a bigger V^{dw} contribution than that coming from $\omega^{\pm b}$. □

We define the Gibbs measure in the usual manner

$$\mu_L(\mathrm{d}\omega) := \frac{1}{Z_L}\exp(-H_L(\omega))\mathrm{d}\omega \tag{3.6}$$

with the partition function

$$Z_L := \int e^{-H_L(\omega)}\mathrm{d}\omega. \tag{3.7}$$

Here $\mathrm{d}\omega := \prod_{x\in\mathbb{T}_L\setminus\{0\}}\mathrm{d}\omega_x$. The zeroth term is omitted because of the translational symmetry of the model $H_L(\omega + C) = H_L(\omega)$.

3.2.2 Extended model

We can simplify the study of the double-well model by looking at a related model obtained by splitting the double-well into two single wells. Similar idea was used in [4] to split the potential (3.1). Suppose $\kappa : E(\mathbb{T}_L) \to \{-1, 1\}$ is such that $\kappa(e) = -\kappa(-e)$. We will denote the set of such configurations by $K_L := K(\mathbb{T}_L)$ and call any of its elements an *edge configuration*. Using it we can introduce an extended Hamiltonian

$$H^e_L(\omega,\kappa) := \frac{K_1}{4}\sum_{e\in E(\mathbb{T}_L)}(\omega_{e_e}-\omega_{e_b}-\kappa_e a)^2 + \frac{K_2}{8}\sum_{e,f\in E(\mathbb{T}_L),e_e=f_b,e\perp f}(\omega_{f_e}-\omega_{e_b})^2 \tag{3.8}$$

and the corresponding extended Gibbs measure

$$\mu^e_L(\mathrm{d}\eta\times\mathrm{d}\kappa) := \frac{1}{Z_L}\exp(-H^e_L(\eta,\kappa))\mathrm{d}\omega\times\mathrm{d}\kappa \tag{3.9}$$

where $\mathrm{d}\kappa$ is the counting measure on the edge configurations K_L.

The extended model is related to the main model by the following observation

$$\begin{aligned}
\mu_L(A) =& \frac{1}{Z_L}\int_A \exp(-H_L(\omega))\mathrm{d}\omega \\
=& \frac{1}{Z_L}\int_A \prod_{e\in E(\mathbb{T}_L)}\left(\sum_{k\in\{-1,1\}}\exp(-K_1(\omega_{e_e}-\omega_{e_b}-ka)^2/2)\right) \\
&\times \prod_{e,f\in E(\mathbb{T}_L),e_e=f_b,e\perp f}\exp(-K_2(\omega_{f_e}-\omega_{e_b})^2/2) \\
=& \sum_{\kappa\in K_L}\frac{1}{Z_L}\int_A \exp(-H^e_L(\omega,\kappa))\mathrm{d}\omega = \sum_{\kappa\in K_L}\mu^e_L(A\times\kappa).
\end{aligned} \tag{3.10}$$

This means any quantity of interest in the main model can be obtained from the extended model. The simplification rests on the fact that the extended model is almost Gaussian in the sense that for each $\kappa \in K_L$ the Hamiltonian $H^e_L(\omega,\kappa)$ is a quadratic form in ω. For this very reason it will be convenient to define

$$Z_{L,(\kappa)} := \int \exp(-H^e_L(\omega,\kappa))\mathrm{d}\omega. \tag{3.11}$$

and from (3.10) we have

$$Z_L = \sum_{\kappa \in K_L} Z_{L,(\kappa)}. \tag{3.12}$$

To conclude this section we also state the counterpart to proposition (3.1) concerning the ground states of H_L^e.

Denote by z the d-cube at origin the functions $b^{\pm} : V(z) \to \{0, \pm 1\}$,

$$b^{\pm}(x) = \begin{cases} 0 & \sum_j x_j \equiv 0 \pmod 2 \\ \pm 1 & \sum_j x_j \equiv 1 \pmod 2. \end{cases} \tag{3.13}$$

In other words $b^{\pm}$ vanishes on the even vertices and is equal to ± 1 on the odd vertices. Using these functions we can define certain elementary configurations as $\xi^{\pm}(e) := b^{\pm}(e_e) - b^{\pm}(e_b)$ (this is indeed an edge configuration because 1-cubes only connect even vertices with odd vertices). Denote by $\kappa_0^{\pm} := D(\xi^{pm})$ the disseminated configurations obtained from $\xi^{\pm}$. We claim the following.

Proposition 3.2. *The Hamiltonian H_L^e is always non-negative. It vanishes at $\omega^{\pm a} \times \kappa_0^{\pm}$. For any other state $\omega \times \kappa \neq \omega^{\pm a} \times \kappa_0^{\pm}$ the Hamiltonian is positive.*

Proof. The discussion of the K_2 part of H_L^e is the same as in the proof of the proposition (3.1). As for the K_1 part we have $(\nabla \omega^{\pm a})_e = a\kappa_0^{\pm}(e)$ for all $e \in E(\mathbb{T}_L)$ and therefore this part vanishes as well.

Now, let any other state $\omega \times \kappa$ be given. If it is not constant on both even and odd sublattices we get a non-zero K_2 contribution and the statement holds. On the other hand if ω is constant on both sublattices then since $\kappa(e) = \pm 1$ for all $e \in E(\mathbb{T}_L)$ and $\omega \neq \omega^{\pm a}$ we again get a non-zero contribution. Therefore assume $\omega = \omega^{\pm a}$. If $(\nabla \omega^{\pm a})_e = a\kappa(e)$ for every edge then κ actually coincides with $\kappa_0^{\pm}$. This cannot happen by hypothesis and so for at least one edge e we again get a non-zero K_1 contribution. We have thus exhausted all the cases and the proof is complete. □

3.3 Reflection positivity of the extended model

The Hamiltonian H_L^e given by (3.8) can be seen to be reflection positive with respect to θ by writing it as

$$
\begin{aligned}
H_L^e(\omega,\kappa) = {} & \frac{K_1}{4}\sum_{e\in E(\mathbb{T}_L)^+}((\nabla\omega)_e-\kappa_e a)^2+\frac{K_1}{4}\sum_{e\in E(\mathbb{T}_L)^-}((\nabla\omega)_e-\kappa_e a)^2\\
&+\frac{K_2}{8}\sum_{e,f\in E(\mathbb{T}_L)^+,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2+\frac{K_2}{8}\sum_{e,f\in E(\mathbb{T}_L)^-,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2\\
&+\frac{K_2}{4}\sum_{e\in E(\mathbb{T}_L)^+,f\in P,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2+\frac{K_2}{4}\sum_{e\in E(\mathbb{T}_L)^-,f\in P,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2\\
&+\frac{K_2}{4}\sum_{f\in E(\mathbb{T}_L)^+,e\in P,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2+\frac{K_2}{4}\sum_{f\in E(\mathbb{T}_L)^-,e\in P,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2\\
&+\frac{K_1}{4}\sum_{e\in P}((\nabla\omega)_e-\kappa_e a)^2+\frac{K_2}{8}\sum_{e,f\in P,e_e=f_b,e\perp f}((\nabla\omega)_e+(\nabla\omega)_f)^2.
\end{aligned}
$$

The formula looks complicated but the underlying idea is simple, the pairs on the first four lines transform upon reflection into each other while the last two terms are left invariant by the reflection since they only involve terms living on the hyperplane. This is thus the decomposition of the Hamiltonian required for the lemma (2.7) to apply.

4. Free energy calculations

4.1 Momentum representation

In this section we will compute the free energies of disseminated configurations which will later be used to estimate probabilities of the associated elementary configurations. We begin by exploiting the discrete translational symmetry of the Hamiltonian (3.8). Using Fourier transform (section D.2) we can write

$$H_L^e(\widehat{\omega},\kappa) = \frac{K_1}{2}\sum_{k\in\mathbb{T}_L^*}\left(|\widehat{\omega}_k|^2\widehat{D}_k^{(\mathrm{nn})} - 2a\sum_{r\in\{\mathbf{e}_1,\dots,\mathbf{e}_d\}}\widehat{\omega}_k^*\widehat{\partial}_k^{r*}\widehat{\kappa}_k^r + a^2\right) + \frac{K_2}{2}\sum_{k\in\mathbb{T}_L^*}|\widehat{\omega}_k|^2\widehat{D}_k^{(\mathrm{nnn})} \tag{4.1}$$

where we introduced $\widehat{\kappa}_k^r$ as a Fourier transform of $\kappa_x^r := \kappa_{(x,x+r)}$ and $\widehat{D}_k^{(\mathrm{nn})}$ and $\widehat{D}_k^{(\mathrm{nnn})}$ are laplacians for the gradient operators defined by (D.7) and (D.8).

It will turn out to be useful to introduce

$$\widehat{A}_k := \frac{K_1}{2}\widehat{D}_k^{(\mathrm{nn})} + \frac{K_2}{2}\widehat{D}^{(\mathrm{nnn})} \qquad \widehat{B}_k^\kappa := \frac{aK_1}{2}\sum_{r\in\{\mathbf{e}_1,\dots,\mathbf{e}_d\}}\widehat{\partial}_k^{r*}\widehat{\kappa}_k^r. \tag{4.2}$$

If the configuration κ is understood, we will simply write $\widehat{B}_k$.

Using this notation we can rewrite the Hamiltonian as

$$H_L^e(\widehat{\omega},\kappa) = \sum_{k\in\mathbb{T}_L^*}\left(\widehat{A}_k|\widehat{\omega}_k|^2 - 2\widehat{\omega}_k^*\widehat{B}_k\right) + K_1a^2L^dd/2. \tag{4.3}$$

The reality of the Hamiltonian is no longer apparent in this form. To fix this, let us recall the fact for any real function $f \in \mathbb{R}^{\mathbb{T}_L} \subset \mathbb{C}[\mathbb{T}_L]$ we have $\widehat{f}_k^* = \widehat{f}_{-k}$ and so we can group together terms with k and $-k$. Let us also write $\omega_k = r_k + is_k$. Then we have

$$H_L^e(\widehat{\omega},\kappa) = \sum_{k\in\mathbb{T}_L^*}\left(\widehat{A}_k(r_k^2+s_k^2) - 2\mathbf{Re}\widehat{B}_kr_k - 2\mathbf{Im}\widehat{B}_ks_k\right) + K_1a^2L^dd/2 \tag{4.4}$$

and completing the squares

$$\begin{aligned} H_L^e(\widehat{\omega},\kappa) = \sum_{k\in\mathbb{T}_L^*\setminus\{0\}}&\left(\widehat{A}_k\left(r_k - \frac{\mathbf{Re}\widehat{B}_k}{\widehat{A}_k}\right)^2 + \widehat{A}_k\left(s_k - \frac{\mathbf{Im}\widehat{B}_k}{\widehat{A}_k}\right)^2 - \frac{\widehat{B}_k^2}{\widehat{A}_k}\right) \\ &+ K_1a^2L^dd/2. \end{aligned} \tag{4.5}$$

Note that the omission of the $k=0$ term is justified because $\widehat{A}_0$ and $\widehat{B}_0$ vanish (a consequence of the absence of mass in the model) while $\widehat{A}_k \neq 0$ for $k \neq 0$.

As alluded to before, we will need to obtain the free energies of the disseminated κ configurations. Using the correspondence (3.11) they can be expressed as

$$Z_{L,(\kappa)} = \int \exp(-H_L^e(\omega,\kappa)) \prod_{x\in\mathbb{T}_L\setminus\{0\}} \mathrm{d}\omega_x \tag{4.6}$$

The integral is taken over (L^d-1)-dimensional real subspace of $\mathbb{C}[\mathbb{T}_L]$ spanned by the basis $\{\delta_x \mid x \in \mathbb{T}_L \setminus \{0\}\}$. By Fourier transform we can also parametrize it by $\widehat{\omega}_k$ with respect to a basis of irreducible modules $\{\rho_k \mid k \in \mathbb{T}_L^* \setminus \{0\}\}$. The real parametrization is given by $\{\widetilde{r}_k | k \in K^r\} \cup \{\widetilde{s}_k | k \in K^s\}$ where K^r is an arbitrary fundamental domain of the action $\Gamma^r : k \mapsto -k$ of $\mathbb{Z}/2\mathbb{Z}$ on $\mathbb{T}_L^* \setminus \{0\}$ and $K^s := K^r \setminus \mathbb{T}_2^*$. The need for the introduction of K^s stems from the fact that for $k \in \mathbb{T}_2^*$ we have $k = -k$ and so $\widehat{\omega}_k$ modes are already real with no $\widetilde{s}_k$ part.

Because the Fourier transform is an isometry we have

$$Z_{L,(\kappa)} = \int \exp(-H_L^e(\widehat{\omega}, \kappa)) 2^{|K^s|} \prod_{k \in K^r} \mathrm{d}\widetilde{r}_k \prod_{k \in K^s} \mathrm{d}\widetilde{s}_k \tag{4.7}$$

where the factors of two come from the Jacobian of the parametrization $\widehat{\omega}_k = \widetilde{r}_k \pm i\widetilde{s}_k$.

Plugging in (4.5) and shifting the integration variables we obtain

$$\begin{aligned} Z_{L,(\kappa)} = \int \exp\left(- \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \left(\widehat{A}_k \left(r_k^2 + s_k^2 \right) - \frac{\widehat{B}_k^2}{\widehat{A}_k} \right) \right) 2^{|K^s|} \prod_{k \in K^r} \mathrm{d}\widetilde{r}_k \prod_{k \in K^s} \mathrm{d}\widetilde{s}_k \\ \times \exp(-K_1 a^2 L^d d/2) \end{aligned} \tag{4.8}$$

and noting that for every $k \in K^s$ there are two terms proportional to both $\widetilde{s}_k$ and $\widetilde{r}_k$ in the exponential (since $\widehat{A}_k = \widehat{A}_{-k}$) that precisely cancel with the powers of two we have

$$Z_{L,(\kappa)} = \prod_{k \in \mathbb{T}_L^* \setminus \{0\}} \sqrt{\frac{\pi}{\widehat{A}_k}} \exp\Big(\sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \frac{\widehat{B}_k^2}{\widehat{A}_k} - K_1 a^2 L^d d/2 \Big) \tag{4.9}$$

Finally, we arrive at the free energy

$$F_{L,(\kappa)} := -\log Z_{L,(\kappa)} = \sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \Big(\frac{1}{2} \log \Big(\widehat{A}_k - \pi \Big) - \frac{\widehat{B}_k^2}{\widehat{A}_k} \Big) + K_1 a^2 L^d d/2. \tag{4.10}$$

4.2 Free energies in two dimensions

Since our proof is based on chessboard estimates we will need to obtain free energies of the disseminated configurations obtained from elementary configurations by reflections. In two dimensions this problem can be solved directly. We will refer to 2-cubes as *plaquettes*.

Now, regarding the plaquette configurations, we will not need to compute free energies associated to all of them because many configurations will produce disseminated configurations that only differ by a translation and consequently will have the same free energy. Utilizing this knowledge the $2^4 = 16$ allowed plaquette configurations can be enumerated in the following way:

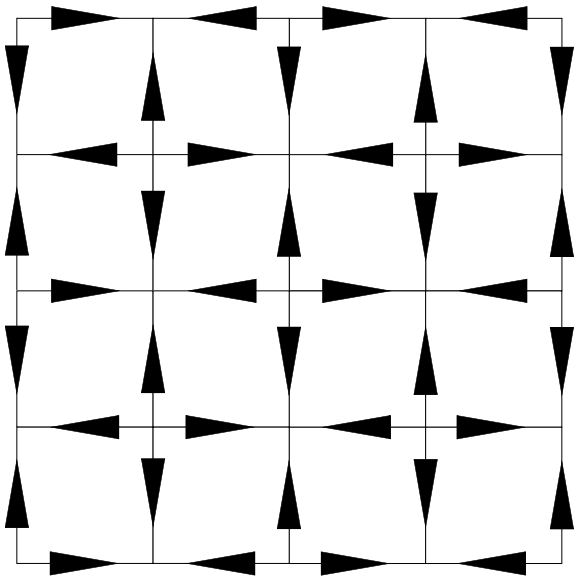


Figure 4.1: Disseminated state G

- Two elementary configurations that correspond to ground states. These generate completely antisymmetric disseminated configurations $\kappa_0^{\pm}$ displayed in figure (4.1). In this and following figures we show the direction of edges where $\kappa(e) = +1$. We will refer to it as state G.

- Eight elementary configurations that generate completely symmetric disseminated configurations. These are displayed in figure (4.2) and we will refer to it as state S.

- Four states that generate disseminated configurations symmetric in one direction and antisymmetric in the other direction. These are displayed in figure (4.3) and we will call them PA and PA'. All of these configurations will have the same energy because they can be related by a flip along diagonal which is also a symmetry of the Hamiltonian.

- Two states that generate completely antisymmetric configurations but different from $\kappa_0^{\pm}$. We will call them state A and they can be seen in figure (4.4).

Since disseminated configurations are 2-periodic we only need to specify them on the plaquette lying at origin by 2×2 matrices κ_i^r where $r \in \{\mathbf{e}_1, \mathbf{e}_2\}$ and $i \in \{G, S, PA, A\}$. In the following table we list these elementary configurations in the first column and the corresponding Fourier modes of $\widehat{\kappa}_k^r$, $k \in \{(0,0), (0,\pi), (\pi,0)(\pi,\pi)\}$ (note that all other modes vanish trivially because of the 2-periodicity).

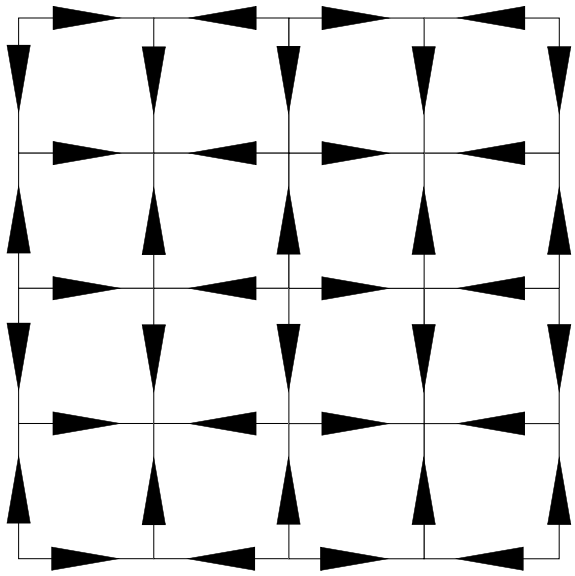


Figure 4.2: Disseminated state S

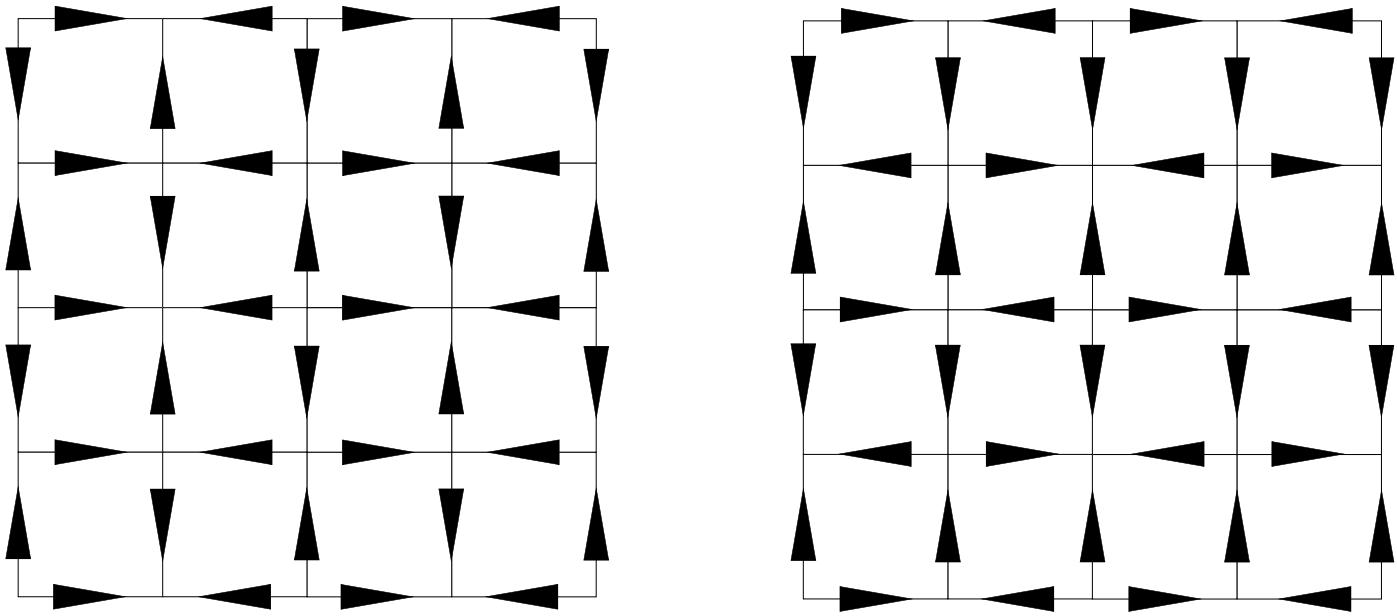


Figure 4.3: Disseminated states PA and PA'

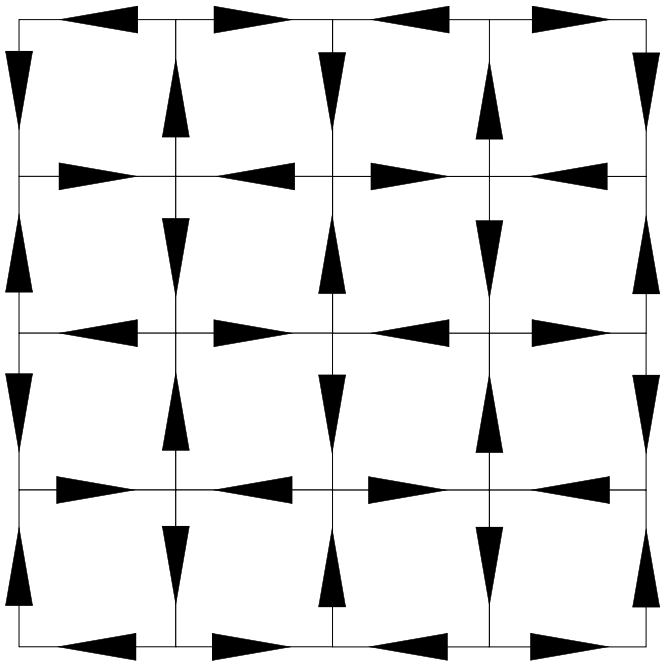


Figure 4.4: Disseminated state A

	κ on plaquette at origin	corresponding Fourier modes $\widehat{\kappa}$
$\kappa_G^{e_1}$	$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & L \end{pmatrix}$
$\kappa_G^{e_2}$	$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & L \end{pmatrix}$
$\kappa_S^{e_1}$	$\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & L \\ 0 & 0 \end{pmatrix}$
$\kappa_S^{e_2}$	$\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ L & 0 \end{pmatrix}$
$\kappa_{PA}^{e_1}$	$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & -L \end{pmatrix}$
$\kappa_{PA'}^{e_2}$	$\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ L & 0 \end{pmatrix}$
$\kappa_A^{e_1}$	$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & L \end{pmatrix}$
$\kappa_A^{e_2}$	$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & -L \end{pmatrix}$

The resulting free energies $F_L^j := F_{L,(\kappa_j)}$ – after discarding terms not containing $\widehat{B}_k$ which are are not important for the comparisons here – are then given as

$$\begin{aligned} F_L^G &= -K_1 a^2 L^2, \\ F_L^S &= -K_1 a^2 L^2 \frac{1}{1 + 2K_2/K_1} \\ F_L^{PA} &= -K_1 a^2 L^2 \left(\frac{1}{2 + 4K_2/K_1} + 1/4 \right), \\ F_L^A &= 0. \end{aligned} \tag{4.11}$$

For $K_2 > 0$ we have $F_L^G < F_L^S$, $F_L^G < F_L^{PA}$ and $F_L^G < F_L^A$. For $K_2 = 0$ we have $F^G = F^S$ which signals reappearance of the ground state degeneracy. On the other hand, if $K_2 \gg K_1 \gg 1$ we see that $F_L^G \to -\infty$ whereas the rest of the free energies $F_L^i \to 0$. This reflects the fact that the configurations $\omega^{\pm a} \times \kappa_0^{\pm}$ are the ground states of H_L^e.

4.3 Free energies in arbitrary dimension

When working in dimensions $d > 2$, it becomes infeasible to enumerate all of the disseminated states because any d-cube contains $d \cdot 2^{d-1}$ edges and so there will be $2^{d \cdot 2^{d-1}}$ elementary configurations (for $d = 3$ we already have $2^{12} = 4096$ of them[1]). This means that direct computations of free energies are no longer possible and we will need to resort to estimates.

Let us recall that the free energy of the disseminated configuration κ is given by (4.10) as

$$F_{(\kappa)} = -\sum_{k \in \mathbb{T}_2^* \setminus \{0,(\pi,\dots,\pi)\}} \frac{\widehat{B}_k^{(\kappa)^2}}{\widehat{A}_k} - \frac{\widehat{B}_{(\pi,\dots,\pi)}^{(\kappa)^2}}{\widehat{A}_{(\pi,\dots,\pi)}}. \tag{4.12}$$

where we discarded terms not containing $\widehat{B}_k$. We have also used the fact that for disseminated configurations – which are 2-periodic – $\widehat{\kappa}_k^r$ vanishes whenever $k \cdot e_i$ is different from 0 or π for any $1 \le i \le d$. We have also singled out the $k = (\pi, \dots, \pi)$ mode because it is naturally connected to groundstates and we will see that in certain regime this mode is in fact a major contribution to the free energy.

We begin the discussion with the study of the behavior of $\widehat{D}^{(\mathrm{nn})}$ and $\widehat{D}^{(\mathrm{nnn})}$. First note that for $r \in \{e_1, \dots, e_d\}$ we have $|1 - e^{-ik\cdot r}|^2 = 4\delta_{k\cdot r = \pi}$ and so we easily obtain $\widehat{D}_k^{(\mathrm{nn})} \ge 4$ for $k \neq 0$. We get a similar although less trivial bound also for $\widehat{D}^{(\mathrm{nnn})}$.

Proposition 4.1. *$\widehat{D}_k^{(\mathrm{nnn})} \ge 4$ for $k \in \mathbb{T}_L^* \setminus \{0, (\pi, \dots, \pi)\}$ and vanishes otherwise.*

Proof. If $k = 0$ then

$$\widehat{D}_k^{(\mathrm{nnn})} = \sum_{r \in \{\mathbf{e}_1 \pm \mathbf{e}_2, \dots, \mathbf{e}_{d-1} \pm \mathbf{e}_d\}} \left|1 - e^{ik\cdot r}\right|^2 = \sum_{r \in \{\mathbf{e}_1 \pm \mathbf{e}_2, \dots, \mathbf{e}_{d-1} \pm \mathbf{e}_d\}} |1 - 1|^2 = 0.$$

Now, since $r = \pm e_i \pm e_j$ for some $1 \le i < j \le d$ we see that $r \cdot (\pi, \dots, \pi) = \pm\pi \pm \pi = 2m\pi$ for some $m \in \mathbb{Z}$ and so every summand in $\widehat{D}_{(\pi,\dots,\pi)}^{(\mathrm{nnn})}$ is zero. On the other hand, if $k \neq 0$ and $k \neq (\pi, \dots, \pi)$ then there must be some $1 \le i \le d$

[1] Of course, symmetry considerations reduce the number of states one needs to look at – we have seen in the previous chapter that we only needed to deal with 4 states instead of 16 in $d = 2$ – but even so the number of states is huge already in $d = 3$.

and $1 \leq j \leq d$, $i \neq j$ such that $k \cdot e_i = 0$ and $k \cdot e_j = \pi$ and let $r' = e_i + e_j$. Then $k \cdot r' = \pi$ and consequently

$$\widehat{D}_k^{(\mathrm{nnn})} = \left|1 - e^{ik\cdot r'}\right|^2 + \sum_{r \in \{\mathbf{e}_1 \pm \mathbf{e}_2, \ldots, \mathbf{e}_{d-1} \pm \mathbf{e}_d\} \setminus \{r'\}} \left|1 - e^{ik\cdot r}\right|^2 \geq 4. \qquad \square$$

Next we will investigate the quantities $\widehat{B}_k^\kappa$. It turns out we can get a uniform bound across all k and κ.

Proposition 4.2. *For all $k \in \mathbb{T}_L^*$ and for all disseminated configurations κ we have $\left|\widehat{B}_k^\kappa\right| \leq aK_1 dL^{d/2}$.*

Proof.

$$\left|\widehat{B}_k^\kappa\right| = aK_1 \left|\sum_r \delta_{k\cdot r = \pi} \widehat{\kappa}_k^r\right| \leq aK_1 d \frac{1}{L^{d/2}} \left|\sum_{x \in \mathbb{T}_L} \kappa_x^r e^{ik\cdot x}\right| \leq$$

$$\leq aK_1 d \frac{1}{L^{d/2}} \sum_{x \in \mathbb{T}_L} \left|\kappa_x^r e^{ik\cdot x}\right| = aK_1 dL^{d/2}$$

where the last equality follows from the fact that $\kappa_x^r = \pm 1$ for all $x \in \mathbb{T}_L$. $\square$

As we have mentioned $\widehat{B}_{(\pi,\ldots,\pi)}^\kappa$ is a special mode. This is because of the following observation.

Proposition 4.3. *Denote by κ_0 any ground state configuration (the configuration κ_0^+ differs from κ_0^- just by translation and so we need not distinguish them here). Then $\left|\widehat{B}_{(\pi,\ldots,\pi)}^{\kappa_0}\right| = aK_1 L^{d/2} d$. Moreover for any disseminated configuration $\kappa \neq \kappa_0^\pm$ we have $\left|\widehat{B}_{(\pi,\ldots,\pi)}^{\kappa}\right| \leq aK_1 L^{d/2}(d-1)$.*

Proof. The expression of B^κ for the $(\pi, \ldots, \pi)$ mode simplifies to

$$\widehat{B}_k^\kappa = aK_1 \sum_{r \in \{\mathbf{e}_1, \ldots, \mathbf{e}_d\}} \widehat{\kappa}_{(\pi,\ldots,\pi)}^r.$$

Now $\widehat{\kappa}_{(\pi,\ldots,\pi)}^{e_i}$ is equal to $\pm L^{d/2}$ if the κ^{e_i} is antisymmetric in every direction and vanishes otherwise. It follows that maximum possible value for $\left|B_{(\pi,\ldots,\pi)}^\kappa\right|$ is $aK_1 L^{d/2} d$. We split the rest of the discussion into two cases.

- States which have κ^{e_i} periodic in direction e_j for any $1 \leq i \leq d$, $1 \leq j \leq d$. If this occurs the $r = e^i$ summand vanishes and the whole expression is bounded by $aK_1 L^{d/2}(d-1)$.
- Completely antisymmetric states. To maximize $\left|\widehat{B}_{(\pi,\ldots,\pi)}^\kappa\right|$ we need the signs of all the $\pm aK_1 L^{d/2}$ contributions to agree. This happens precisely for the ground states $\kappa_0^\pm$.

$\square$

From the propositions (4.1) and (4.2) we can get a lower bound on the free energies of disseminated configurations

$$F_{L,(\kappa)} = -\frac{\widehat{B}^2_{(\pi,\dots,\pi)}}{A_{(\pi,\dots,\pi)}} - \sum_{k\in\mathbb{T}_L\setminus\{0,(\pi,\dots,\pi)\}} \frac{\widehat{B}^2_k}{\widehat{A}_k} \geq -\frac{\widehat{B}^2_{(\pi,\dots,\pi)}}{A_{(\pi,\dots,\pi)}} - \frac{(aK_1d)^2(2L)^d}{4K_2}. \quad (4.13)$$

Finally, we can estimate differences between free energies of the ground state and any other disseminated state.

Lemma 4.4. *Let κ_0 be a ground state configuration, κ any disseminated configuration and let a coupling K_1 be given. Then there exists K_2' such that for all $K_2 > K_2'$ it holds $F_{L,(\kappa)} - F_{L,(\kappa_0)} \geq a^2K_1L^d/2$.*

Proof. Using proposition (4.3) and the bound (4.13) we get

$$F_{L,(\kappa)} - F_{L,(\kappa_0)} \geq \frac{-\left(\widehat{B}^{\kappa}_{(\pi,\dots,\pi)}\right)^2 + \left(\widehat{B}^{\kappa_0}_{(\pi,\dots,\pi)}\right)^2}{A_{(\pi,\dots,\pi)}} - \frac{(aK_1d)^2(2L)^d}{4K_2} \geq$$

$$\geq a^2K_1L^d\frac{2d-1}{2d} - \frac{(aK_1d)^2(2L)^d}{4K_2} = a^2K_1L^d\left(1 - \frac{1}{2d} - \frac{K_1d^22^d}{4K_2}\right).$$

Now simply choose $K_2' = 2K_1d^22^d$. □

5. Coexistence bounds

5.1 Contour counting

Let U be a set of m-cubes in T_L^d. To any such set we associate the graph $GU := (U, E)$ where $\{x, y\} \in E \subset \binom{U}{2}$ if and only if $x \cap y$ is an $(m-1)$-cube.

Denote by W the set of all d-cubes of T_L^d and suppose W is partitioned into three non-empty subsets $W = W_+ \cup W_- \cup W_B$. We say that a cube x is of type $q \in \{+, -, B\}$ if $x \in W_q$.

Let now $q \in \{+, -\}$ and let Q be a set of $(d-1)$-cubes such that GQ is connected and every $z \in Q$ can be written as $z = x \cap y$ where $x \in W_q$ and $y \in W_B$. We will call such a Q a *contour* (of type q). We say that contour Q is *closed* if its diameter is less than L and Q has no boundary. If Q is closed then by theorem (A.3) $T_L^d \setminus [Q]$ has two components, one of them with diameter less than L that we will call *interior of* Q.

Given an m-skeleton U we can introduce the projection to the ith axis $\pi^i[U]$. Since U is just a union of cubes we can also introduce the 1-skeleton $\pi^i U := \{X_i \in \mathbb{Z}/L\mathbb{Z} \,|\, \exists u(X, I) \in U : I(i) = 1\}$ so that $[\pi^i U] = \pi^i[U]$. We then define the *diameter* of U to be $\operatorname{diam}(U) := \max_{1 \le i \le d} |\pi^i U|$. When $\operatorname{diam}(U) < L$ we can define a *bounding box* to be the smallest box that is also a d-skeleton that contains U in its interior.

We claim the following.

Theorem 5.1. *Suppose no cube of type $+$ is adjacent to a cube of type $-$ in GW and let $P_+ \in W_+$ and $P_- \in W_-$. Then there exists a contour Q such that either $|Q| \ge L$ or Q is of type $+$ and contains P_+ in its interior or Q is of type $-$ and contains P_- in its interior.*

Proof. Denote by J_q, $q \in \{+, -\}$, the maximal connected component of GW_q containing P_q. Then J_q is a d-skeleton and has a boundary $(d-1)$-skeleton $K_q := \partial J_q$ which splits into m_q connected components $K_q = C_q^1 \cup \cdots \cup C_q^{m_q}$ when considered as a vertex set of GK_q. Notice that each of C_q^j is a contour of type q.

Now if the diameter of any of the C_q^j is greater than L the proof is finished. So let us assume this is not the case. Then all of the contours are closed (they have no boundary since they are themselves components of boundaries).

In the set of contours $\{C_q^1, \ldots, C_q^{m_q}\}$ there is precisely one that contains at least one of J_+ or J_- in its interior (note that it can contain both of them). Denote it by D_q. Now if D_+ contains $[J_+]$ in its interior we are finished. If it does not contain $[J_+]$ then from the way we have chosen D_+ we see that it must contain $[J_-]$ in its interior. But then it must also contain D_- in its interior (since D_- is a part of the boundary of $[J_-]$) which in turn also contains $[J_-]$ in its interior. So we see that D_- is the contour we are looking for. □

Proposition 5.2. *Let a $(d-1)$-cube x be given. Then there are at most $(6d-6)^{2n}$ contours of length n that contain x.*

Proof. Let Q be a contour of length n that contains x. Since GQ is connected by lemma (B.4) there exists a cycle C of length $2n-1$ that visits every vertex of GQ. Therefore the number of contours of length n containing x is bounded from above by the number of cycles of length $2n-1$ containing x. The number of cycles is in turn bounded by q^{2n-2} where q is the number of $(d-1)$-cubes adjacent to a given $(d-1)$-cube. Since every $(d-1)$-cube contains $2(d-1)$ $(d-2)$-cubes and there are three possibilities how a given $(d-2)$-cube can be adjacent to another $(d-1)$-cube we have that $q = 6(d-1)$. □

Proposition 5.3. *Let a closed contour Q of length n and diameter less than L be given and let x be a d-cube in the interior of Q. Then there is a $(d-1)$-cube $y \in Q$ such that Manhattan distance of x and y is less than $\frac{n}{4(d-1)}$.*

Proof. Consider the projection $\pi^1 Q$ of Q to the first axis. Since $[Q]$ is connected the projection is connected as well. Now consider the set O of $(d-1)$-cubes orthogonal to the first axis that lie above x. Because Q is a boundary of an m-skeleton containing x, at least two elements of O must belong to Q. But now the projection of these $(d-1)$-cubes to the first axis must lie in $[\pi^1 Q]$. From these two $(d-1)$-cubes pick the one closer to x and denote it by y. Because $n \geq 2(d-1)|\pi^1 Q|$ and $\mathrm{dist}(x,y) \leq |\pi^1 Q|/2$ we are finished. □

5.2 Peierls estimates

In this section we will write $\mathbb{P}$ for the probability measure μ_L^e. For ground states it holds that $\kappa_0^+ = -\kappa_0^-$. In other words, it is never true that $\kappa(e) = \kappa_0^+(e)$ and $\kappa(f) = \kappa_0^-(f)$ for two distinct edges $e, f \in E(\mathbb{T}_L)$ when $\kappa = \kappa_0^{\pm}$. We would like to show that a similar fact applies generally for any configuration at all, not just for the ground states.

Theorem 5.4. *Let $\epsilon > 0$ and let $e, f \in E(\mathbb{T}_L)$. Then we can find the couplings $K_1 > 0, K_2 > 0$ such that*

$$\mathbb{P}\left(\kappa(e) = \kappa_0^+(e) \,\&\, \kappa(f) = \kappa_0^-(f)\right) < \epsilon. \tag{5.1}$$

To prove this statement we will use the following.

Theorem 5.5. *Let $\epsilon > 0$ and let x, y be two d-cubes in T_L^d. Let A denote the event that $\kappa(e) = \kappa_0^+(e)$ for all edges e of x and $\kappa(f) = \kappa_0^-(e)$ for all edges f of y. Then we can find the couplings $K_1 > 0, K_2 > 0$, independent of L, such that*

$$\mathbb{P}(A) < \epsilon. \tag{5.2}$$

If (5.5) holds then a fortiori so must (5.4) – just pick any two d-cubes containing the given edges and apply the bound (5.2). Therefore it suffices to prove the second theorem.

Proof. Let x, y be d-cubes and let κ be any configuration such that $\kappa(e) = \kappa_0^+(e)$ for all edges e of x and $\kappa(f) = \kappa_0^-(f)$ for all edges f of y. Using κ we can partition the set of all d-cubes into three types. Let z be any d-cube. If $\kappa(e) = \kappa_0^\pm(e)$ for all edges e of z then we say z is of type $\pm$. Otherwise z will be of type B. It holds that a d-cube of type $+$ cannot be adjacent to a d-cube of type $-$ because for any edge e lying in their common $(d-1)$-cube we would have $\kappa_0^+(e) = \kappa(e) = \kappa_0^-(e)$ – a contradiction. Therefore we see that the assumptions of theorem (5.1) are satisfied and there must exist a contour of certain properties. In other words we get a bound

$$\mathbb{P}(A) \leq \sum_n \mathbb{P}(C_n) \cdot c_n \tag{5.3}$$

where C_n is the event that a contour of length n exists and c_n is the number of contours of length n. Recall that every contour Q (of type q) separates d-cubes of type q from d-cubes of type B. This means that every contour determines a set of d-cubes of type B. Now since every d-cube contains $2d$ $(d-1)$-cubes it follows that there must be at least $|Q|/2d$ d-cubes of type B. Therefore, using chessboard estimates (2.9) and the sub-additivity lemma (2.8), we can bound the probability $\mathbb{P}(C_n)$ as

$$\mathbb{P}(C_n) \leq \left(\sum_\kappa \mathbb{P}(\kappa) \right)^{[n/2d]} \tag{5.4}$$

where the sum runs over all disseminated configurations distinct from $\kappa_0^\pm$. Taking $K_2 > 2K_1 d^2 2^d$ it follows from the lemma (4.4) on free energy bounds that

$$(\mathbb{P}(\kappa))^{|\mathbb{T}_L|} = \frac{Z_{L,(\kappa)}}{Z_L} \leq \frac{Z_{L,(\kappa)}}{Z_{L,(\kappa_0)}} = \exp(-(F_{L,(\kappa)} - F_{L,(\kappa_0)})) = \epsilon_0^{|\mathbb{T}_L|} \tag{5.5}$$

where

$$\epsilon_0 = \exp(-a^2 K_1/2)$$

which can be made arbitrarily small by choosing K_1 suitable large. Using all of this we get

$$\mathbb{P}(C_n) \leq \left(2^{d \cdot 2^{d-1}} \epsilon_0 \right)^{[n/2d]} \tag{5.6}$$

Let us now turn to the problem of estimating the number of contours of the given length. Suppose first that the contour has length $n \geq L$. Then we have L^d options of where to put one of its $(d-1)$-cubes so using proposition (5.2) we get that the number of such contours is bounded by $n^d \cdot (6d-6)^{2n}$. Similarly, suppose the contour has length n less than L. Without loss of generality suppose it contans the d-cube x in its interior. Then proposition (5.3) implies we have

$\frac{n}{4(d-1)}$ possibilities where to put one of its $(d-1)$-cubes and using proposition (5.2) we obtain the bound $\frac{n}{4(d-1)} \cdot (6d-6)^{2n} \leq n^d \cdot (6d-6)^{2n}$.

Putting it all together we get

$$\mathbb{P}(A) \leq \sum_n \left(2^{d \cdot 2^{d-1}} \epsilon_0\right)^{[n/2d]} \cdot n^d \cdot (6d-6)^{2n} \tag{5.7}$$

This series can be bounded by a geometric series $C \cdot \sum_{n=n_0} s^n$ where $C > 0$ and $n_0 > 0$ and with arbitrarily small coefficient s (by choosing ϵ_0 sufficiently small). The sum of such a series and therefore the sought for probability is bounded from above by

$$\mathbb{P}(A) \leq C \cdot s^{n_0}/(1-s) < 2C \cdot s^{n_0} =: \epsilon. \tag{5.8}$$

□

5.3 Localization bounds

We would like to extend the result of the previous chapter – that only considers the behavior of the κ field – to the ω field. It is intuitively clear that, fixing κ, the ω fields likes to occur near $a\kappa$ simply because this is where minimum of the Hamiltonian $H(\cdot, \kappa)$ is. But to be able to show an actual localization we will also need to bound the fluctuations of ω. Before approaching this problem we will need to obtain another estimate. Denote by $\mathbb{P}^{\pm}(\cdot) := \mu_e^L(\cdot|\kappa(f) = \kappa_0^{\pm}(f))$ the conditional probabilities of the extended Gibbs measure where $f \in E(\mathbb{T}_L)$ and write $\mathbb{E}^{\pm}$ for their corresponding expectations.

Conjecture 5.6. *It is possible to choose the couplings $K_2 \gg K_1 \gg 1$ in a way such that*

$$\mathbb{E}^{\pm}\left(\sum_{e \in E(\mathbb{T}_L)} ((\nabla\omega)(e) - a\kappa(e))^2\right) \leq Ca^2 dL^d$$

where $C \leq 1 - \delta$ for some positive constant δ independent of L.

We demonstrate here the reasons we believe this conjecture is true as well as one possible approach to proving it. The sum under expectation is, up to multiple of four, nothing else than the K_1 part of $H_L^e(\omega, \kappa)$ and therefore can be computed as

$$-\frac{\partial \log Z}{\partial K_1} = \frac{1}{Z} \sum_{\kappa \in K(\mathbb{T}_L)} \frac{\partial F_{(\kappa)}}{\partial K_1} \tag{5.9}$$

where the free energy is given by equation (4.10). Because both $\widehat{A}_k$ and $\widehat{B}_k$ depend only linearly on K_1 – with derivatives evaluating to $\frac{\partial \widehat{A}_k}{\partial K_1} = \widehat{D}^{(\mathrm{nn})}/2$ and $\frac{\partial \widehat{B}_k}{\partial K_1} = \widehat{B}_k/K_1$ we can evaluate the free energy derivative as

$$\frac{1}{Z} \sum_{\kappa \in K(\mathbb{T}_L)} \left(\sum_{k \in \mathbb{T}_L^* \setminus \{0\}} \left(\frac{\widehat{D}^{(\mathrm{nn})}}{4\widehat{A}_k} - \frac{\widehat{B}_k^2}{\widehat{A}_k} \left(\frac{K_1 \widehat{D}^{(\mathrm{nn})} + 2K_2 \widehat{D}^{(\mathrm{nnn})}}{2K_1 \widehat{A}_k} \right) \right) + a^2 L^d d/2 \right). \tag{5.10}$$

We will not consider the $\frac{\widehat{D}^{(\mathrm{nn})}}{4\widehat{A}_k}$ term since it is less than $1/2K_1$ so by choosing K_1 big enough it can be made very small in comparison with other terms. Therefore we see that the whole expression is bounded by $a^2L^d d/2$. This means that the conjecture holds if weakened to $C = 2 + \delta$ for any $\delta > 0$.

Observe now that the term proportional to $\frac{\widehat{B}_k^2}{\widehat{A}_k}$ is only slightly smaller than that in the expression (4.10) for the free energy itself. Using the results on free energies of disseminated states obtained in section 4.3 we can see that for ground states $\kappa_0^{\pm}$ we get zero contribution to the average energy (again, disregarding the $\frac{\widehat{D}^{(\mathrm{nn})}}{4\widehat{A}_k}$ term). Naturally, the states κ that are close to the groundstate will also have very small contribution to the average energy and because they are the most likely ones when the couplings are high enough (with the probability going as $\sim \exp(-F)$) it is natural to conclude that the average energy can be made small by a suitable choice of the couplings.

Let us now return to the problem of bounding the fluctuations of ω. Given that the conjecture holds we have the following result.

Theorem 5.7. *Let an $\epsilon > 0$ and edge $e \in E(\mathbb{T}_L)$ be given. Then there exists couplings K_1, K_2 and a positive constant $C \leq 1/2 - \delta$ such that*

$$\mathbb{P}^{\pm}\left(\left((\nabla\omega)(e) - a\kappa^{\pm}(e)\right)^2 \geq \lambda^2\right) \leq \epsilon + \frac{Ca^2}{\lambda^2}. \tag{5.11}$$

Here $\delta > 0$ is a positive constant indepedent of L.

Proof. Pick any K_1 and K_2 obtained from the theorem (5.4) for the given ϵ and a given edges e, f. Then

$$\begin{aligned}
&\mathbb{P}^{\pm}\left(\left((\nabla\omega)(e) - a\kappa^{\pm}(e)\right)^2 \geq \lambda^2\right) = \\
&= \mathbb{P}^{\pm}\left(\left((\nabla\omega)(e) - a\kappa^{\pm}(e)\right)^2 \geq \lambda^2 \,\&\, \kappa(e) = \kappa^{\mp}(e)\right) \\
&+ \mathbb{P}^{\pm}\left(\left((\nabla\omega)(e) - a\kappa^{\pm}(e)\right)^2 \geq \lambda^2 \,\&\, \kappa(e) = \kappa^{\pm}(e)\right) \leq \\
&\leq \epsilon + \mathbb{P}^{\pm}\left(((\nabla\omega)(e) - a\kappa(e))^2 \geq \lambda^2\right) \leq \\
&\leq \epsilon + \frac{\mathbb{E}^{\pm}\left(((\nabla\omega)(e) - a\kappa(e))^2\right)}{\lambda^2}
\end{aligned}$$

where the last line follows from the Chebyshev's inequality. We have thus reduced the proof to showing that the expectation of $(\nabla\omega)(e) - a\kappa(e))^2$ is less than C. Observe that because of the translational and rotational symmetry of Hamiltonian H_L^e this expectation equals to the expectation of the average taken over all the edges

$$\mathbb{E}^{\pm}\left(\frac{1}{|E(\mathbb{T}_L)|}\sum_{e\in E(\mathbb{T}_L)}((\nabla\omega)(e) - a\kappa(e))^2\right).$$

The proof now follows from the conjecture above. ∎

6. Infinite volume models

We now proceed to the definition of the double-well model on $\mathbb{Z}^d$. After that we will show how this model relates to the torus models studied in the prior chapters. Using those results we will be able to construct two infinite volume Gibbs measures thereby showing an occurence of phase transition.

6.1 Potentials

6.1.1 Double-well model

We take the discrete set the model will live as $S = S_1 := \mathbb{Z}^d$ and for every $x \in \mathbb{Z}$ the spin state will be $E_x := \mathbb{R}$. The potential representing this model is

$$\Phi_A(\omega) := \begin{cases} V^{\mathrm{dw}}(\omega_x - \omega_y) & A = \{x, y\};\ x, y \text{ nearest neighbors} \\ V^{\mathrm{sw}}(\omega_x - \omega_y) & A = \{x, y\};\ x, y \text{ next-nearest neighbors} \\ 0 & \text{otherwise.} \end{cases} \tag{6.1}$$

where the symmetric pair potentials V^{dw}, V^{sw} are given by equations (3.3) and (3.4).

6.1.2 Extended model

The extended model will live not only on the vertices of $\mathbb{Z}^d$ but also on its edges. We therefore take $S = S_2 := \mathbb{Z}^d \times E(\mathbb{Z}^d)$. The spin state $E_x := \mathbb{R}$ for $x \in \mathbb{Z}^d$ and $E_x := \{-1, 1\}$ for $x \in E(\mathbb{Z}^d)$. The potential for this model is

$$\Phi^e_A(\omega) := \begin{cases} \frac{K_1}{4}(\omega_x - \omega_y - a\omega_{(x,y)})^2 & A = \{x, y, (x, y)\};\ x, y \text{ nearest neighbors} \\ \frac{K_2}{2}(\omega_x - \omega_y)^2 & A = \{x, y\};\ x, y \text{ next-nearest neighbors} \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

note that here we are using ω in the sense of chapter 1 to mean the total field composed of both scalar and discrete fields of the previous chapters.

6.2 Gibbsian constructions

Using the potentials defined above we can use the constructions of chapter one and define the Hamiltonians $H^{\Phi}_{\Lambda}(\omega)$ and $H^{\Phi^e}_{\Lambda}(\omega)$ using (1.2) and define the corresponding Gibbsian specifications $(\gamma^{\Phi}_{\Lambda})_{\Lambda \in \mathscr{S}_1}$ and $(\gamma^{\Phi^e}_{\Lambda})_{\Lambda \in \mathscr{S}_2}$ using (1.3) where $\mathscr{S}_1$ is taken to be the collection of finite subsets of $\mathbb{Z}^d$ and similarly $\mathscr{S}_2$ the collection of finite subsets of $\mathbb{Z}^d \times E(\mathbb{Z}^d)$. All the potentials are of course admissible since they decay sufficiently fast.

Following the example (4.20) in [1], given a potential Φ, one can introduce the Δ_L-periodic potential $\tilde{\Phi}^{\Delta_L}$ where Δ_L is a box of d-cubes with all sides of length L. Consider then the associated Gibbsian specification $\gamma^{\tilde{\Phi}_{\Delta_L}}$. The restriction of $\gamma_{\Delta_L}^{\tilde{\Phi}_{\Delta_L}}(\cdot|\omega)$ to $\mathscr{F}_{\Delta_L}$ is independent of ω and we call it the *Gibbs ditribution in* Δ_L *for* Φ *with periodic boundary conditions*. Moreover these Gibbs distributions are precisely equal to μ_L respectively μ_L^e from chapter 3 when we specialize to potentials Φ respectively Φ^e from the preceding section. Now let $f \in E(\Delta_L)$. We define

$$\mu^{e,\pm}(A)_L := \mu_L(A \,|\, \omega(f) = \pm\kappa_0^{\pm}(f)). \tag{6.3}$$

Any cluster point of the sequence of measures $\lim_{L\to\infty} \gamma_{\Delta_L}^{\tilde{\Phi}_{\Delta_L}}(\cdot|\omega)$ is Gibbsian for the specification γ^{Φ}. This is again shown in (4.20) in [1].

To show that a cluster point exists we use the theorem (1.3). We only need to show that

$$\lim_{l\to\infty} \limsup_{n\in\mathbb{N}} \mu_n(\omega(x) \notin K_x^l) = 0 \tag{6.4}$$

for x a nearest neighbor of 0 because other hypotheses are trivially satisfied and for general $x \in \mathbb{T}_L$ the claim follows from translational invariance of the model. The limit is zero because for large values of $\nabla\omega$ (which is implicit since $\omega(0) = 0$) the field does not feel the presence of the double-well and thus the model behaves like a Gaussian model. But for Gaussian model the bound follows trivially.

This construction works also works for the extended because edge configurations do not be taken into account (since they have finite state space). We thus see that infinite volume Gibbs states given by the double-well and extended specifications exist.

There is one small caveat here – in the above estimate one needs to assume $d > 2$ because the gradient Gaussian field in two dimensions has infinite fluctuations. Nevertheless, it is possible to construct a model for the gradient of the Gaussian field and in that model the fluctuations turn out to be finite. It is in this sense that the results of [3] and [4] hold also for $d = 2$. We note that one can also extend the results of the present work to $d = 2$ using the same methods.

6.3 Phase coexistence in the infinite volume

Using the theorem (5.4), it is clear that for any $e \in E(\Delta_L)$ we can find couplings K_1, K_2 such that $\mu_L^{e,\pm}(\kappa_e = \mp 1) \leq \epsilon$ and moreover this is independent of L. Therefore, taking the $L \to \infty$ limit (which we now know exists) it must also be true that $\mu^{e,\pm}(\kappa_e = \mp 1) \leq \epsilon$. But this means that the measures $\mu^{e,\pm}$ are distinct.

Similarly, applying the theorem (5.7) for $\epsilon < \delta/2$ and $\lambda = a$ to the bounds $\mu_L^{\pm}\left(((\nabla\omega)(e) - a\kappa^{\pm}(e))^2 \geq a^2\right)$ we see that for μ_L^+ and $\kappa^+(e) = +1$ the gradient of the field $(\nabla\omega)(e)$ is with probability more than $1/2 + \delta/2$ inside the interval $[0, 2a]$. Similarly for μ_L^- and interval $[-2a, 0]$. In other words the states $\mu_L^{\pm}$ prefer

to be localized around the values prescribed by $\kappa_0^{\pm}$. Because these bounds are independent of L, after taking the limit $L \to \infty$ we get the same estimates also for the $\mu^{\pm}$ measures which are therefore distinct.

A. Algebraic topology

Algebraic topology is a subject that assigns to topological spaces certain algebraic invariants which often reduces a hard geometrical problem into a simple algebraic one. It is a rich subject that also studies the general interplay between topology and algebra and quickly leads to some of the deepest parts of mathematics. Here we will content ourselves with the elementary parts of the theory which, nevertheless, provide a unified and simplied view of many standard constructions that are often encountered.

A.1 CW complex

We begin by specifying a class of spaces we will be working with. In stark contrast with the point-set topology that concerns itself with the study of general (and seemingly pathological) topological spaces, algebraic topology requires from the start that the spaces be nicely behaved. We will forgo a precise definition of 'nicely behaved' – as well as other topological and set-theoretical issues – since it obscures rather than illuminates the concepts we will be dealing with (interested reader is encouraged to study [5] though).

The CW complex X is a space defined as $X = \bigcup_n X^n$ where the n-skeletons X^n are defined inductively. For $n = 0$ X^0 is just a set of points. Given X^{n-1} one constructs X^n using gluing operations to attach n-dimensional balls along their boundary sphere S^{n-1} by means of continuous maps $\phi : S^{n-1} \to X^{n-1}$. These maps can then be extended to the full maps $\phi : D^n \to X^n$ and the images of the balls $\phi(D^n)$ are called n-cells of X.

A 1-dimensional CW complex is called *graph* (possibly with loops and multiple edges) since S^0 is a set of two points. Therefore the attaching described above is precisely the attaching of intervals D^1 to the points X^0 of the graph.

Let us move into an n-dimensional Euclidean space $\mathbb{R}^d$. The subset

$$\Delta^m := \left\{ (x_1, \ldots, x_m) \in R^d \,\middle|\, 0 \leq \sum_i x_i \leq 1; \quad x_i \geq 0, \ldots, x_m \geq 0 \right\}$$

will be called a *standard m-simplex*. Elements of $Aff(\mathbb{R}^d)\Delta^m$ will be called *m-simplices*.

Consider now an embedding of the CW complex into a Euclidean space. The complex (identified with its image in $\mathbb{R}^d$) will be called *simplicial*. gets mapped to a simplex.

A.2 Homology and cohomology

Let an n-dimensional CW complex X be given and consider the free abelian group $C_m(X)$ generated by the set of all m-cells in X. We will call the elements of $C_m(X)$ *m-chains* – these are just formal linear combinations of m-cells with integral coefficients. One can then define *boundary maps* $\partial_m : C_m(X) \to C_{m-1}(X)$ (we will return to the precise definition later since it can be quite technical). The name for these maps comes from the fact that they satisfy $\partial_{m-1} \circ \partial_m = 0$ which implies $\mathrm{Im}\partial_m \leq \mathrm{Ker}\partial_{m-1}$ or geometrically that every *boundary* (an element of $\mathrm{Im}\partial_m$) is a *cycle* (an element of $\mathrm{Ker}\partial_{m-1}$). Since $C_m(X)$ is an abelian group, we can form the quotient $H_m(X) \equiv \mathrm{Ker}\partial_m/\mathrm{Im}\partial_{m+1}$ – *the mth homology group*. The reason these groups are important is that they are topological invariants – if two spaces X and Y are homotopy equivalent [1] then $H_m(X) \cong H_m(Y)$. Also, any continuous map $f : X \to Y$ induces a homomorphism of groups $f_* : H_m(X) \to H_m(Y)$.

Now consider the group of homomorphisms from $C_m(X)$ to $\mathbb{Z}$. We will denote it by $C^m(X)$ and call its elements *m-cochains*. This group is generated by the homomorphisms

$$\psi_e(f) = \begin{cases} 1 & e = f \\ 0 & \text{otherwise} \end{cases},$$

one for each m-cell e and we have a natural identification $\psi_e \leftrightarrow e$ that extends to an isomorphism $C^m(X) \cong C_m(X)$.

One can then define the *coboundary maps* $\delta_m : C^{m-1}(X) \to C^m(X)$ (the precise definition being again postponed) with $\delta_m \circ \delta_{m-1} = 0$, the group of *coboundaries* $\mathrm{Im}\delta_{m-1}$, the group of *cocycles* $\mathrm{Ker}\delta_m$ and finally the *mth cohomology group* $H^m(x) \equiv \mathrm{Ker}\delta_m/\mathrm{Im}\delta_{m-1}$. These are again topological invariants. In fact, there is an isomorphism $H^m(X) \cong H_m(X)$. This is a very pleasant result since the groups $C^m(X)$ and maps between them are purely algebraic and do not seem to contain, on the first sight, the straightforward geometric content carried by $C_m(X)$.

Let us now come back to the technical details underlying the maps ∂_* and δ_*. There are actually many (co)homology theories that can be defined and for any reasonable space X they give isomorphic (co)homology groups. They only differ in the definition of the groups of m-(co)chains and the (co)boundary maps between them (and, of course, in the amount of work required to obtain useful results).

We provide a brief list of some common homology theories.

- *Simplicial homology* – for space X that is a simplicial complex the m-chains are formed as linear combinations of m-simplices. We can identify

[1] Recall that spaces X and Y are homotopy equivalent if there exist maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f$ can be continuously deformed to an identity Id_X.

any m-simplex S in the chain with the standard simplex Δ^m. Write $S = [v_1, \ldots, v_m]$ after this identification. We can then define the boundary map

$$\partial_m S \equiv \sum_{i=1}^{m} (-1)^i [v_1, \ldots, \hat{v}_i, \ldots, v_m].$$

- *Singular homology* Here the m-chains are defined to be linear combinations of maps from the m-simplex to X. For any such map $\sigma : \Delta^m \to X$ one defines the boundary map

$$\partial_m \sigma \equiv \sum_{i=1}^{m} (-1)^i \sigma\Big|_{[t_1,\ldots,\hat{t}_i,\ldots,t_m]}.$$

- *Cellular homology* This is a theory suited for CW complexes that we adopted in the exposition above – the m-chains are the linear combinations of m-cells themselves. Like for simplices, the boundary of a given m-cell is defined as a linear combination of the $(m-1)$-cells that it was attached to with coefficients ± 1. Unlike for the simplices though, the coefficients do not come from an arbitrary ordering we have put on a simplex but rather can be computed directly as a degree [2] of a certain map

$$\partial_m e_j^m = \sum_k \deg \phi_{jk} e_k^{m-1}$$

where

$$\phi_{jk} : S^{m-1} \to X^{m-1} \to S^{m-1}$$

is a composition of the attaching map of e_j^m and the map that collapses $X^{m-1} - e_k^{m-1}$ to a point.

By dualizing the above definitions we obtain the cohomological versions of the said theories. For example, for simplicial cohomology we have

$$[\delta_m \phi]([v_1, \ldots, v_m]) \equiv \sum_{i=1}^{m} (-1)^i \phi([v_1, \ldots, \hat{v}_i, \ldots, v_m]).$$

and similarly for other theories.

More generally, let a ring R be given and consider the free R-module $C_m(X; R)$ on the m-cells of X. All of the above constructions carry over to this setting without any change, so we arrive at $H_m(X; R)$ resp. $H^m(X; R)$ – *the mth (co)homology R-module*. The group version above is recovered by noting that an abelian group is nothing else than a $\mathbb{Z}$-module.

[2] Recall that the degree of a map $f : S^n \to S^n$ is given as $f_*(1 \in H_n(S^n))$ which can be informally thought of as the number of times the first sphere wraps around the second sphere.

A.3 Certain homological results on manifolds

We will state two important results on duality in homology of manifolds without proof. The proofs can be found in chapter 3.3 of [5].

Theorem A.1. Poincaré duality. *Let X be a connected n-dimensional manifold that is compact, oriented and without boundary. Then $H^k(X) \cong H_{n-k}(X)$ for all $0 \leq k \leq n$.*

Theorem A.2. Alexander duality. *Let X be a compact subspace of $\mathbb{R}^n$. Then $\tilde{H}_k(\mathbb{R}^n \setminus X) \cong \tilde{H}^{n-k-1}(X)$ for all $0 \leq k \leq n-1$.*

Here $\tilde{H}_*(X)$ refers to reduced homology which agrees for $k > 0$ with standard homology $H_*(X)$ and for $k = 0$ we have $\tilde{H}_0(X) := H_0(x)/\mathbb{Z}$. This is convenient because for contracible spaces all reduced homology groups vanish. Also many theorems have more natural statement in terms of reduced homology[3].

From these two theorems we obtain following result.

Theorem A.3. *Let X be a connected n-dimensional submanifold of $\mathbb{R}^{n+1}$ that is compact, oriented and without boundary. Then $Y := \mathbb{R}^{n+1} \setminus X$ has 2 components.*

Proof. Since X is connected we have $H_0(X) = \mathbb{Z}$. Because X is compact, oriented and without boundary we can use Poincaré duality which implies $H_0(X) \cong H^n(X)$. Next, applying Alexander duality we have $\tilde{H}^n(X) \cong \tilde{H}_0(Y)$. Putting it together $\tilde{H}_0(Y) \cong \mathbb{Z}$ and therefore $H_0(Y) \cong \mathbb{Z} \oplus \mathbb{Z}$ which means that Y has two components. □

[3]Reduced cohomology is defined analogously – it agrees with the standard cohomology for $k > 0$ and is reduced at the $k = 0$ term.

B. Graph theory

Recall that we introduced *(undirected) graph* in the previous appendix as a one-dimensional CW complex. In the following we will always assume that a graph contains neither loops nor multiple edges. In this setting we can specify the graph G as a pair (V, E) where V is the set of vertices and $E \subset \binom{V}{2}$ the set of edges. We identify V with the 0-skeleton of the CW-complex and each edge of E with the 1-cell D^1 that it is a boundary of.

One can also consider a *directed graph* which is just a graph where we pick an orientation of each edge. Formally, we talk about the pair (V, E) where V is the set of vertices and $E \subset V \times V$ the set of oriented edges. Unless stated otherwise the word graph will stand for both kinds of graphs and we will also use the same notation for both kinds. The distinction will only be made explicit where necessary.

Definition B.1. *Let a graph $G = (V, E)$ be given. The pair $(V' \subset V, E' \subset E)$ will be called a* subgraph *if it is itself a graph. Given the subset of vertices $W \subset V$ we define a subgraph generated by W as $G(W) = (W, F)$ such that the set of edges F is maximal among all subgraphs of G with vertex set W.*

Graphs can be considered as an abstraction of a transit system where the set of vertices models the stops and the set of edges models the routes between neighboring stops. It is therefore natural to consider the paths that follow the routes.

Definition B.2. *Let a graph $G = (V, E)$ and sequence $P = (v_i)_{i=0}^n$ with $v_i \in V$, $1 \leq i \leq n$ be given. The sequence P will be called a* path *if $\{v_i, v_{i+1}\} \in E$ for $0 \leq i \leq n-1$. If $v_i = v_n$ the path will be called a* loop [1]*. Let two paths $P = (v_i)_{i=0}^n$, $R = (w_j)_{j=0}^m$ such that $v_n = w_0$ be given. We define the* inverse path *$-P \equiv (v_{n-i})_{i=0}^n$ and the* compound path *$P + R \equiv (u_i)_{i=0}^{n+m-1}$ where*

$$u_i = \begin{cases} v_i & 0 \leq i \leq n \\ w_{i-n} & n < i \leq m \end{cases}.$$

Having defined paths one is lead to consider the set of all vertices that can be reached from a given vertex.

Definition B.3. *Let a graph $G = (V, E)$ be given. We will say G is* connected *if there is a path $P = (v_i)_{i=0}^n$ with $v_0 = x$, $v_n = y$ for every $x, y \in V$. A connected subgraph G' of G will be called a* component *of G if $G' = G(U)$ for some $U \subset V$ and there is no $W \subset V$ such that $V \subsetneq W \subsetneq V$ with $G(W)$ connected.*

[1] Note that we do not require paths to respect orientation of edges when G is directed.

Lemma B.4. *Let $G = (V, E)$ be a connected graph. Then there exists a cycle $C = (c_k)_{k=1}^N$, $\{c_i, c_{i+1}\} \in E$, $c_1 = c_N$ that visits every vertex of G and has length $N = 2|V| - 1$.*

Proof. We prove the claim by induction on the number of vertices. When V has only one vertex $V = \{v\}$ the claim holds trivially by taking $C = (v)$. Suppose now that the claim holds for every graph with less than $|V|$ vertices. Let T be any spanning tree of G and let $\{n_1, \dots, n_m\}$ be the set of neighbors of v. T restricted to $V \setminus \{v\}$ then splits into m components, say of sizes $\{s_1, \dots, s_m\}$. By hypothesis there exist cycles $\{C_1, \dots C_m\}$ with lengths $|C_j| = 2s_j - 1$ on the components. We can then form the cycle $C = (v, n_1) + C_1 + (n_1, v, n_2) + C_2 + \cdots + C_m + (n_m, v)$ with length $|C| = 1 + (2s_1 - 1) + 1 + (2s_2 - 1) + \cdots + 1 + (2s_m - 1) + 1 = 1 + 2\sum_{k=1}^m s_k = 2|V| - 1$. □

C. Measure and probability theory

Let Ω be a set.

Definition C.1. *The collection $\mathscr{F} \subset \exp(\Omega)$ is called a* σ-algebra *if it satisfies*

- $\Omega \in \mathscr{F}$
- $\bigcup_{n\in\mathbb{N}} U_n \in \mathscr{F}$ *for any countable collection* $U : \mathbb{N} \to \mathscr{F}$
- $\Omega \setminus U \in \mathscr{F}$ *for any* $U \in \mathscr{F}$

Definition C.2. *A function $\mu : \mathscr{F} \to [0, \infty]$ is called a* measure *if it satisfies*

$$\mu(\bigcup_{n\in\mathbb{N}} U_n) = \sum_{n\in\mathbb{N}} \mu(U_n)$$

for any countable pairwise disjoint collection $U : \mathbb{N} \to \mathscr{F}\,, \quad U_i \neq U_j \quad i \neq j$

The measure μ is called *finite* if $\mu(\Omega) < \infty$ and called a *probability measure* if $\mu(\Omega) = 1$.

The idea behinds these definitions comes from probability theory. The σ-algebra represents a collection of events. We can always observe event Ω denoting that "anything can happen". Such an event has probability 1. If we can observe individual events then we can certainly observe their union (meaning we don't care which of the elementary events happened). If these events are mutually exclusive the total probability must be additive. Also, if we can observe some event then this is equivalent to not observing the complementary event, which should therefore also be observable.

The tuple $(\Omega, \mathscr{F})$ will be called a *measurable space* and a triple $(\Omega, \mathscr{F}, \mu)$ a *measure space*. If the μ is moreover a probability measure we will call it a *probability space*.

A map between measurable spaces $f : (X, \mathscr{X}) \to (Y, \mathscr{Y})$ is said to be $\mathscr{X}/\mathscr{Y}$-measurable if the preimage $f^{-1}(A)$ of every $\mathscr{Y}$-measurable set A is $\mathscr{X}$-measurable. If $Y = \mathbb{R}^d$ and $\mathscr{Y}$ is the Borel σ-algebra generated by the Euclidean topology on $\mathbb{R}^d$ we simply say that the map f is $\mathscr{X}$-measurable.

Definition C.3. *Let $(X, \mathscr{X})$, $(Y, \mathscr{Y})$ be measure spaces. We will call a function $\pi : \mathscr{X} \times Y \to [0, \infty]$ a* measure kernel *from $\mathscr{Y}$ to $\mathscr{X}$ if it satisfies*

- $\pi(\cdot|y)$ *is a measure on $\mathscr{X}$ for every $y \in Y$*
- $\pi(A|\cdot)$ *is $\mathscr{Y}$-measurable for every $A \in \mathscr{X}$.*

The measure kernel is called a *probability kernel* if $\pi(X|y) = 1$ for every $y \in Y$.

The intuition behind this idea comes from the interpretation of Y as the set of boundary conditions and $\mathscr{X}$ as a collection of observable events. For any boundary condition we obtain a measure on the events.

The single most important use for a probability kernel is the possibility to use it to pull-back a measure from $\mathscr{Y}$ to $\mathscr{X}$. Let μ be a measure on $\mathscr{Y}$. Then we can define a pulled-back measure $\mu\pi$ on $\mathscr{X}$ by

$$\mu\pi(A) := \int \pi(A|y)\mu(\mathrm{d}y). \tag{C.1}$$

Similarly, let π_2 be a probability kernel from $\mathscr{Z}$ to $\mathscr{Y}$. Then we can form a probability kernel $\pi_2\pi$ from $\mathscr{Z}$ to $\mathscr{X}$ by

$$\pi_2\pi(A|\cdot) := \int \pi(A|y)\pi_2(\mathrm{d}y|\cdot). \tag{C.2}$$

D. Group theory and representation theory

D.1 General theory

Let a finite group G be fiven. Consider a group algebra $(\mathbb{C}[G], *)$ of functions from G to $\mathbb{C}$ which we will denote by $\mathcal{A}$. This algebra carries a natural basis composed of delta functions δ_g with multiplication given by convolution acting on the basis as $\delta_g * \delta_h = \delta_{gh}$ and extended linearly to $\mathcal{A}$ in the following.

Since the group algebra of a finite group is semisimple, it decomposes as

$$\mathcal{A} \cong \bigoplus_{\Gamma \in \widehat{\mathcal{A}}} \mathrm{End}(V^\Gamma)$$

where the direct sum runs over irreducible modules $\Gamma := (V^\Gamma, \rho^\Gamma)$ of $\mathcal{A}$ which coincide with irreducible modules of G (this is proved for example in [6]). The isomorphism is provided by the Fourier transform $\mathcal{F}$

$$\mathcal{F} : \delta_g \mapsto N \sum_{\Gamma \in \widehat{G}} \rho^\Gamma(g)$$

with N an arbitrary normalization constant.

In the case of abelian groups, Schur's lemma implies that all complex irreducible modules are one-dimensional. This means they are of the form $\rho : G \to \mathbb{C}$ and so in fact $\rho \in \mathbb{C}[G]$. The Fourier isomorphism gives us $|G| = |\widehat{G}|$ and so these modules form a basis of $\mathbb{C}[G]$.

It is straightforward to check that point-wise product of two modules gives again a module. Therefore the set of all irreducible modules can be naturally endowed with a multiplication which moreover satisfies the group axioms. The resulting group $\widehat{G}$ is called the Pontryagin dual of G. The collection $(\rho_\cdot(g))_{g \in G}$ of irreducible modules evaluated at the point forms a basis in $\mathbb{C}[\widehat{G}]$.

For cyclic groups of order n we obtain for any module ρ that $(\rho(g))^n = \rho(g^n) = 1$ and so $\rho(g)$ must be an nth root of unity. It is also easy to check that every nth root of unity ζ gives a module $\rho_\zeta(g) := \exp(i\zeta \cdot g)$. We will also write $\rho_k := \rho_\zeta$ for $\zeta = \exp(ik)$. In this notation $\rho_k \rho_l = \rho_{k+l}$ and so $\widehat{G} \cong \mathbb{Z}/n\mathbb{Z}$.

D.2 Fourier analysis on tori

We will work in d dimensions. As usual, we will identify the torus group with its embedding into $\mathbb{Z}^d$. Because the torus $\mathbb{T}_L$ is just a product of cycle groups the general discussion of the above section carries over into this setting. In particular, the irreducible modules are given by $\rho_k(x) := N \exp(ik \cdot x)$ with k being an element

of the dual torus $\mathbb{T}_L^* := (\frac{2\pi}{L})\mathbb{T}_L$ and we have set $N := L^{-d/2}$ as a normalization constant. By the above discussion we know that (ρ_k) form a basis in $\mathbb{C}[G]$ and it is straightforward to verify that it is orthonormal with respect to standard scalar product

$$\langle f, g\rangle := \sum_{k\in\mathbb{T}_L^*} {f_k}^* g_k \tag{D.1}$$

We can therefore express any function $f \in \mathbb{C}[G]$ as

$$f = \sum_{k\in\mathbb{T}_L^*} \widehat{f}_k \rho_k^*(\cdot) \quad , \qquad \widehat{f}_k := (\mathcal{F}f)_k = \sum_{x\in\mathbb{T}_L} f_x \rho_k(x). \tag{D.2}$$

The usefulness of the Fourier transform stems from the fact that for periodic (or almost periodic) functions almost all modes vanish. For the rest of the discussion suppose L is even.

Definition D.1. *Let $f \in \mathbb{C}[G]$ and suppose for all $x \in \mathbb{T}_L$ such that $y \in \mathbb{T}_{L/2}$ we have $f_x = f_{x+y}$. Then we call f* 2-periodic. *We will call the restriction of f to subset $\mathbb{H} := \{0,1\}^L \subset \mathbb{T}_L$ a* generator of f *and denote it $[f] \equiv f\big|_{\mathbb{H}}$.*

Lemma D.2. *Suppose f is a 2-periodic function. Then $\widehat{f}$ vanishes outside of $\mathbb{C}[\mathbb{T}_2^*]$.*

Proof. Let us compute

$$\begin{aligned}
\widehat{f}_k =& L^{-d/2} \sum_{x\in\mathbb{T}_L} f_x \exp(ik\cdot x) \\
=& L^{-d/2} \sum_{z\in\mathbb{H}} \sum_{y\in\mathbb{T}_{L/2}} f_z \exp(ik\cdot z)\exp(ik\cdot y) \\
=& L^{d/2} \sum_{z\in\mathbb{H}} f_z \exp(ik\cdot z) 2^{-d} \delta_{k\in\mathbb{T}_2^*}
\end{aligned}$$

which means that all coefficients with $k \notin \mathbb{T}_2^*$ vanish. □

Various summations over the torus are also comfortably expressed in terms of the Fourier transform.

$$\sum_{x\in\mathbb{T}_L} f_x g_x = L^{-d} \sum_{x\in\mathbb{T}_L} \sum_{k,k'\in\mathbb{T}_L^*} f_k \rho_k(x) g_{k'}^* \rho_{k'}^*(x) = \sum_{k\in\mathbb{T}_L^*} f_k g_k^* \tag{D.3}$$

The other important observation is that the Fourier transform diagonalizes difference operators on the torus. We will often deal with gradients defined by

$$(\nabla^A f)(x) := (f(x) - f(x+r))_{r\in A} \tag{D.4}$$

for arbitrary subset $A \in E(\mathbb{T}_L)$. For sums of such gradients over the torus we obtain

$$\begin{aligned}\sum_{x\in\mathbb{T}_L}((\nabla^A f)(x))^2 &= \sum_{r\in A}\sum_{x\in\mathbb{T}_L}(f(x)-f(x+r))^2 \\ &= \sum_{r\in A}\sum_{k\in\mathbb{T}_L^*}|\widehat{f}_k|^2(1-e^{-ik\cdot r})(1-e^{ik\cdot r}) \\ &= \sum_{k\in\mathbb{T}_L}|\widehat{f}_k|^2\widehat{D}_k^A\end{aligned} \tag{D.5}$$

where we introduced $\widehat{D}_k^A := \sum_{r\in A}|1-e^{-ik\cdot r}|^2$.

We can also easily evaluate the sum involving the product of gradient with a function

$$\begin{aligned}\sum_{x\in\mathbb{T}_L}(\nabla^A f)(x)\cdot g(x) &= \sum_{r\in A}\sum_{x\in\mathbb{T}_L}(f(x)-f(x+r))g^r(x) \\ &= \sum_{r\in A}\sum_{k\in\mathbb{T}_L^*}\widehat{f}_k\widehat{g}_k^{r*}(1-e^{-ik\cdot r})\end{aligned} \tag{D.6}$$

In particular we define the nearest neighbor gradient $\nabla^{(\mathrm{nn})}$ using the following set of edges

$$A^{\mathrm{nn}} = \{\mathbf{e}_1,\ldots,\mathbf{e}_d\} \tag{D.7}$$

and the next-nearest neighbor $\nabla^{(\mathrm{nnn})}$ gradient using the edges

$$A^{\mathrm{nnn}} = \{\mathbf{e}_1 \pm \mathbf{e}_2,\ldots,\mathbf{e}_{d-1}\pm\mathbf{e}_d\}. \tag{D.8}$$

Conclusion

In chapter one we have introduced the basic issues surrounding the construction of infinite volume statistical models. Then we proceeded to the standard technology that addresses these issues – the Gibbs theory. We have introduced the notion of specification – a collection of compatible finite volume measures with boundary conditions. Specifications enable us to talk about infinite volume Gibbs measure as is any measure that satifies the specification. To construct such measures one usually proceed by introducing a sequence of measures and shows it contains a cluster point.

In chapter two we have introduced the necessary geometrical tools – torus, reflections and cubical complex. These objects and results pertaining to them are of course very ancient but they enable us to talk more easily about certain subsets of the torus, in particular m-cubes and their collections. Then, using the discussed structure of reflections, we introduce the very powerful technology of modern statistical physics – the reflection positivity and chessboard estimates – that we crucially rely on in this thesis as well.

In chapter three we have introduced the notion of the gradient field and its physical significance as an interface between distinct phases of a thermodynamical system. We have reviewed the basic results on convex gradient models – these do not admit phase transition – as well as the more complex non-convex models and the few results available for them. After this preliminary discussion and having the statistical and geometric constructions of the first two chapters out of the way we are finally able to give definition of the non-convex model we have studied in this thesis – the double-well model. We have observed some of its properties, most importantly the ground states that give first hint of a possible phase coexistence at low temperatures. Immediately after this we have introduced a yet another model that we dubbed extended – the reason being it simplifies the analysis of the double-well model and actually determines all of its properties. The extended model is observed to be both simpler and harder than the original model – simpler because it is kind of Gaussian and harder because it needs additional degrees of freedom living on edges that need to be dealt with. Finally, we have shown that the extended model is reflection positive – a fundamental ingredient of our proof of phase coexistence.

In chapter four we have performed the crucial calculations that underlie the rest of the work. We have used Fourier transform to pass to momentum representation and were able to give closed formula for the free energy of any edge configuration. Using this result we have explicitly computed free energies of all disseminated configurations in dimension two. Disseminated configurations are those that have been obtained from a single configuration defined on one d-cube

by means of reflections through hyperplanes. Using chessboard estimates one can then bound the probability of elementerary configuration by negative exponential of the free energy of the associated disseminated configuration.

We have given exact expression for free energy in $d = 2$. These results were not used anywhere else in the work though. They were provided only for the sake of carrying out the full calculation – something that is no longer possible in higher dimensions. When working in arbitrary dimension the number of elementary configurations grows too quickly to allow direct computation and more careful analysis of the disseminated configurations was called for. This investigation has been indeed carried out and we have found that the quantities required for the computation of free energy have a nice pattern and we were able to obtain useful bounds on all free energies.

Chapter five is where the results of all the previous chapters arrive at one common intersection. We were able to show that in any torus, independently of its size, the probability of simultaneous occurence of two disticts good events – meaning that one of them is preferred in one ground state and the other one in the second ground state – can be made low by suitable choice of couplings. To this end we have used Peierls estimate. The core of the idea is that any configuration having two distinct good events will contain a contour of bad events – those that are not of the groundstate form. This argument requires two ingredients – first that there is not too many contours of any given length and second that probability of occurence of the contour decays exponentially with its length with large decay constant. The second fact followed from another usage of chessboard estimates because any contour can be obtained by dissemination of the elementary bad event. Since the probability of the elementary bad event can be made small – a fact shown using the knowledge of free energies of the disseminated bad events – so can be the probability of the contour's occurence. The first ingredient of the Peierls argument was a matter of simple combinatorics because contours are connected and so they can be easily enumerated.

To complete the Peierls argument one therefore only needs to show that a configuration with two distinct good events actually does contain a contour surrounding one of those events. This is not as simple as it might seem because we are working on a torus and there is a possibility of contour wrapping non-trivially around the torus. Nevertheless, assuming that the contour is not too long (which is not a problem since then its probability is already very small), we can use the tools of algebraic topology to show that every connected component of the contour splits the torus into inside and outside. Using this fact we were then able to show that there indeed exists contour of the required properties. This concludes the Peierls argument.

The Peierls argument is only strong enough to show the coexistence in the extended model. To be able to show the coexistence also in the original double-

well model, we need further estimate. Using Chebyshev's inequality and averaging arguments we were able to reduce the needed estimate to that of getting a bound on average energy. We were not able to get the required bound but we have given some analysis of the sum involved in the bound and reasons why we think the bound holds.

In chapter six have constructed the infinite volume measures using the Gibbs technology of chapter one. The coexistence arguments of chapter five were then extended to the infinite volume setting thus completing the proof.

We first note that the regime of phase coexistence $K_2 \gg K_1 \gg 1$ is purely qualitative. We could in principle determine the values of couplings for which all of the required ingredients of the proof are satisfied but we would only obtain very crude estimates since both the free energy estimates and the bounds on the number of contours are hopelessly inoptimal – they are good enough only to show that some (probably very high) values of couplings work. As for the double-well model, there the situation is even worse because the phase coexistence depends on the conjecture. Even if the conjecture is true it might require still higher values of the couplings than that required by Peierls' argument to hold. Therefore one can see that there is a further direction of research in getting tight bounds on all required quantities and determining quantitatively the portion of the K_1, K_2 phase diagram where the coexistence occurs.

Another interesting direction would be to study the $K_2 \to 0$ limit of the model. As we have mentioned, in two dimensions the double-well model is equivalent to an ice model. It might be possible to use the relation to the extended model to get some insight into how this ice model arises as $K_2 \to 0$.

One could also try to use the closed formula for free energy of edge configurations to compute standard quantities of interest in the double-well model – statistical moments and so on (the wanted conjecture is contained in this endeavour since it is related to fluctuations of the field).

In conclusion, we have reviewed some standard techniques of modern statistical physics and applied them to show that an extended gradient model contains a phase transition and, conjecturally, that the double-well gradient model also contains a phase transition.

Bibliography

[1] Georgii, H.O. *Gibbs Measures and Phase Transitions*. Walter de Gruyter, 1988. ISBN 978-3-110-10455-4.

[2] Biskup, Marek and Bovier, Anton and den Hollander, Frank and Ioffe, Dima and Martinelli, Fabio and Netocný, Karel and Toninelli, Fabio. *Methods of contemporary mathematical statistical physics*. Berlin: Springer-Verlag, 2009. ISBN 978-3-540-92795-2.

[3] Funaki, T. and Spohn, H. *Motion by Mean Curvature from the Ginzburg-Landau Interface Model* Comm. Math. Phys., January 1997, vol. 185, no. 1, pp. 1–36.

[4] Biskup, Marek and Kotecký, Roman. *Phase coexistence of gradient Gibbs states*. Probability Theory and Related Fields, 2007, vol. 139, no. 1, pp. 1–39.

[5] Hatcher, A. *Algebraic Topology*. Cambridge University Press, 2001. ISBN 978-0-521-79540-1.

[6] Goodman, R. and Wallach, N.R. *Symmetry, Representations, and Invariants*. Springer, 2009. ISBN 978-0-387-79851-6.