Macroscopic behaviour of Lipschitz random surfaces

Pieter Gezinus Lammers

Queens' College, University of Cambridge

June 2020

This thesis is submitted for the degree of Doctor of Philosophy.

Abstract

A *random field* is a random function ϕ from the square lattice $\mathbb{Z}^d$ to some fixed standard Borel space $(E, \mathcal{E})$. A *random surface* is a random field with the extra condition that $E \in \{\mathbb{Z}, \mathbb{R}\}$ where $\mathcal{E}$ is the standard σ-algebra. For random surfaces, one often studies the gradient $\nabla\phi$ of the random function of interest. Random fields and random surfaces serve as toy models for analysing several phenomena in statistical physics: examples include percolation models, the Ising model, dimer models, the discrete Gaussian free field, and uniformly random Lipschitz functions.

We analyse the specific free energy functional for a class of random fields and for a class of random surfaces. In either case, we are interested in the nature of the minimisers of the specific free energy, and we give a new characterisation of these minimisers even when the model fails to be quasilocal. This immediately leads to a notion of free energy in the spirit of Burton and Keane. In the case of random fields, we derive a concise theory which includes several existing results, and use this theory to prove new results for the Loop $O(n)$ model and the Griffiths singularity model.

The study of the minimisers of the specific free energy is part of a larger programme, where the ultimate goal is to derive strict convexity of the surface tension for random surface models which are monotone in boundary conditions. We prove this conjecture in the case that the model is also Lipschitz, although we also impose some very mild conditions on the representation of the model in terms of an interaction potential to guarantee well-definedness of the statistical mechanical quantities. This in contrast to the work of Sheffield, where the case for strict convexity depends strongly on special properties of the potential, namely that it is a convex nearest-neighbour potential. The results in this thesis include a large deviations principle for (simultaneously) the macroscopic shape and the microscopic statistics of the surface under consideration. Applications include models induced by *submodular potentials*, that is, potentials which satisfy the Fortuin-Kasteleyn-Ginibre lattice condition. This answers an open question of Sheffield in the Lipschitz case: we derive that the surface tension is strictly convex. We furthermore prove a conjecture of Menz and Tassy: we derive strict convexity of the surface tension for uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree, for any $d, k \geq 2$. This is remarkable as the target space is not $\mathbb{Z}$ or $\mathbb{R}$.

Finally, we prove new results for a generalisation of the hexagonal dimer model to higher dimensions. We give a much more direct version of Sheffield's proof for strict convexity of the surface tension, tailored to the special structure of the model. The same structure implies an identity for the covariance structure of the model in terms of its random geometry. We also derive a generalised Kasteleyn theory: the partition function of the model equals the Cayley hyperdeterminant of the hypergraph which is the natural dual to the graph supporting the generalised model.

To my parents and my brothers

Acknowledgements

The submission of this thesis marks the end of a personal journey. I now look back and am so happy to have embarked on this journey, because of the many people that I met and because of the countless things that I learned from those people.

I am most grateful to James Norris, for supervising me during the past four years. Without his confidence in me by taking me as a PhD student and his encouragement to explore my mathematical interests, I would not be in the same place. I looked forward to every meeting we have had, and left with new energy. I feel so fortunate to have had a supervisor who always had my interest at heart, and who knows how to motivate me.

My special thanks go also to Nathanaël Berestycki. I am grateful for the many different ways in which Nathanaël has helped me: by supervising me while James was away from Cambridge, by introducing me to Martin Tassy and encouraging us to work together, and by commenting on my work. Each of these things have been very valuable to me, and they are much appreciated.

Thanks to Martin Tassy, for sharing his ideas with me from the very first time we met, which led to an intense collaboration. Martin invited me into his house many times, and it feels like we are friends more than colleagues.

I would like to thank Geoffrey Grimmett and Scott Sheffield in advance for taking part in my examination. I realise that it is a substantial commitment, and that I will benefit greatly from having examiners who are so close to the topics in this thesis.

Thanks to Ioan Manolescu for inviting me to his workshop and later for a visit, which has helped me tremendously in becoming part of the community.

Thanks to Erik van den Ban for supervising me as an undergraduate student, which I consider my starting point in becoming a researcher.

Thanks to my father, who sensed my interest in mathematics and stimulated and helped develop this interest in many different ways when I was young.

Many others have helped and inspired me during the writing of my dissertation, by discussing ideas or working together: in particular, I would like to thank Nilanjana Datta, Aernout van Enter, Alexander Glazman, Vadim Gorin, Ulrik Hansen, Matan Harel, Peter van Hintum, Richard Kenyon, Benoît Laslier, Marcin Lis, Sébastien Ott, Ron Peled, Fabio Toninelli, and Peter Winkler.

I am grateful to Tessa, as well as to all my friends on the PhD program, for creating such a fantastic atmosphere around the department.

I would like to thank some people for helping and inspiring me in ways not related to mathematics. Thanks in particular to my mother, grandmother, and brothers, and to Andrej, Eduardo, Floris, Gary, Jasper, Koen, Marilou, Milo, and Titia.

Finally, I would like to thank my best friend and girlfriend Rebecca, for being such a friendly and positive spirit, and for helping me through the most stressful parts of the program.

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. Chapter 2 and Chapter 3 derive from joint work with Martin Tassy. The two authors contributed equally to each chapter.

Contents

List of figures

Chapter 1

Introduction

1.1 Preface

The behaviour of many physical systems is understood to be random, or at least chaotic, on the atomic scale. The purpose of statistical mechanics is to understand the macroscopic behaviour of such systems, through a probabilistic, mathematical analysis on the microscopic scale. It turns out that the macroscopic time evolution of these systems is often predictable and essentially deterministic. This principle—that macroscopic order emerges from microscopic disorder—is at the core of statistical mechanics.

The development of statistical mechanics is relatively new. Initial efforts within the sciences focused instead on a deterministic understanding of the natural phenomena that we, as humans, can observe directly. Mathematics turned out to be an indispensable tool for building this understanding from the very start. The calculus of variations, for example, plays a vital role in the rich tradition in classical physics to relate the equilibria of physical systems to the local minima $f : D \to \mathbb{R}$ of the fundamental integral

$$\int_D \sigma(x, f(x), \nabla f(x))dx$$

over all functions f that are differentiable almost everywhere. Here D is a bounded open subset of Euclidean space $\mathbb{R}^d$, and the function σ, which is called the *free energy function*, encodes the physical properties of the system. This approach applies to a broad spectrum of phenomena, and was already present implicitly in the work of Zenodorus (c. 200 BC), Pappus (c. AD 300) and Galileo, albeit from a geometrical rather than an analytical perspective [23]. The analytical formulation of the calculus of variations first appeared in the work of Fermat, who postulated the *principle of least time* to describe the refraction of light when passing from one medium to another. Another key development is the introduction of Lagrangian mechanics: a variational reformulation of classical mechanics due to Lagrange. Contributors to the calculus of variations include a number of great scientists: Newton, Leibniz, the Bernoullis, Euler, Jacobi, Weierstrass, and Hilbert.

We shall focus in this thesis on the application of the calculus of variations to surfaces. Variational calculus is used in this context to understand the shape of materials subject to external forces and boundary conditions. The following is a tangible example in the context of surfaces: it is empirically understood that a soap film tends to minimise its surface area when subjected to boundary constraints. A soap bubble is spherical, because a sphere minimises the area of the surface, subject to a volume constraint.

Essential to the theory of variational calculus is the assumption that the model of interest is continuous in spirit. This contradicts our understanding of the nature of matter at the atomic scale. The first application of statistical mechanics in this case is to derive the variational characterisation of the macroscopic equilibria of a system, from the microscopic description of that system. The free energy function σ, which is also called *surface tension* in the context of surfaces, plays a pivotal role in this connection between small and large. The surface tension is a macroscopic quantity, but relates directly to the *specific free energy* of the microscopic equilibria of the model. Suppose for a moment that the surface tension is *strictly convex*. This would imply that the fundamental integral has a unique minimiser, which in turn implies that the system has a unique macroscopic equilibrium. Broadly speaking, this assumption leads to the premise of statistical mechanics that macroscopic stability emerges from microscopic disorder. This thesis fits into a larger programme, where the ultimate objective is to justify this assumption from a rigorous, mathematical perspective. The most significant result in this thesis, is that we prove strict convexity of the surface tension for nearly all random surface models which are Lipschitz and which are monotone in boundary conditions.

The surface tension is the thread that connects the three subsequent chapters. Each chapter corresponds to a separate research article, and can, as such, be considered an independent piece of work. Chapter 2 and Chapter 3 derive from work written in collaboration with Martin Tassy, while Chapter 4 is the sole work of the author of this thesis.

The remainder of this chapter consists of three sections. Section 1.2 provides a brief overview of the thesis. Section 1.3 contains the definitions for the formal discussion of random surfaces. Section 1.4 presents the main results of each chapter.

1.2 Overview

A *random field* is a random function ϕ from the square lattice $\mathbb{Z}^d$ to some fixed standard Borel space $(E, \mathcal{E})$. A *random surface* is a random field with the extra condition that $E \in \{\mathbb{Z}, \mathbb{R}\}$ where $\mathcal{E}$ is the standard σ-algebra. For random surfaces, one often studies the gradient $\nabla\phi$ of the random function of interest. Random fields and random surfaces serve as toy models for analysing several phenomena in statistical physics: examples include percolation models, the Ising model, dimer models, the discrete Gaussian free field, and uniformly random Lipschitz functions.

In Chapter 2, we analyse the specific free energy functional for a class of random fields. We are interested specifically in the nature of the minimisers of this functional, and we give a characterisation of these minimisers even when the model fails to be quasilocal. This immediately leads to a notion of free energy in the spirit of Burton and Keane. We derive a concise theory which includes several existing results, and use this theory to prove new results for the Loop $O(n)$ model and the Griffiths singularity model.

The study of the minimisers of the specific free energy is part of a larger programme, where the ultimate goal is to derive strict convexity of the surface tension for random surface models which are monotone in boundary conditions. This conjecture was first established by Sheffield for convex nearest-neighbour potentials in his seminal work *Random Surfaces*. In Chapter 3, we prove the conjecture in the case that the model under consideration is both monotone and Lipschitz, although we also impose some very mild conditions on the representation of the model in terms of an interaction potential to guarantee well-definedness of the statistical mechanical quantities. This in contrast to the work of Sheffield, where the case for strict convexity depends strongly on special properties of the potential, namely that it is a convex nearest-neighbour potential. Chapter 3 relies crucially on the ideas from Chapter 2 on minimisers of the specific free energy functional, suitably adapted to the setting of random surfaces. The results in Chapter 3 include a large deviations principle for (simultaneously) the macroscopic shape and the microscopic statistics of the surface under consideration. Applications include models induced by *submodular potentials*, that is, potentials which satisfy the Fortuin-Kasteleyn-Ginibre lattice condition. This answers an open question of Sheffield in the Lipschitz case: we derive that the surface tension is strictly convex. We furthermore prove a conjecture of Menz and Tassy: we derive strict convexity of the surface tension for uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree, for any $d, k \geq 2$. This is remarkable as the target space is not $\mathbb{Z}$ or $\mathbb{R}$.

In Chapter 4, we discuss a generalisation of the hexagonal dimer model to higher dimensions. We prove additional results for this model, which also falls into the class of Chapter 3. We give a much more direct version of Sheffield's proof for strict convexity of the surface tension, tailored to the special structure of the model. The same structure implies an identity for the covariance structure of the model in terms of its random geometry. We also derive a generalised Kasteleyn theory: the partition function of the model equals the Cayley hyperdeterminant of the hypergraph which is the natural dual to the graph supporting the generalised model.

1.3 The formalism of random surfaces

We are concerned with the study of *random surfaces*, which are special cases of *random fields*. We shall follow the notation of Georgii [20] or otherwise Sheffield [54] wherever possible. The reader should be warned that the notation for the context of random surfaces sometimes differs from the notation for general random fields—we shall make notice of this when encountering these differences. Chapter 2 uses the notation of random fields, while Chapter 3 uses the notation of random surfaces. Chapter 4 concerns a specific model of random surfaces, and is much lighter in notation relative to the other two.

We introduce the definitions and notations in four stages. In the first stage, we introduce random fields and random surfaces, which are merely distributions. In the second stage, we introduce *specifications* and *potentials*, which encode the interactions of the model of interest. Once the definition of a specification is established, we can rightfully say that a random field is a *Dobrushin-Lanford-Ruelle (DLR) measure*, that is, an equilibrium of the model. In the third stage, we introduce the notions of *entropy*, *specific free energy*, and *surface tension*. One improvement that is made in this thesis, is that we include models with infinite-range interactions in the analysis. To this end, we introduce some original constructions in the fourth stage. We recall the definition of *quasilocality*, and introduce objects to describe the behaviour of random fields and random surfaces when quasilocality fails.

1.3.1 Random fields and random surfaces

If $(X, \mathcal{X})$ is any measurable space, then write $\mathcal{P}(X, \mathcal{X})$ for the set of probability measures on $(X, \mathcal{X})$, and $\mathcal{M}(X, \mathcal{X})$ for the set of σ-finite measures μ with $\mu(X) > 0$. It is always tacitly understood that all measurable spaces are standard Borel spaces: this is useful as it implies immediately that regular conditional probability distributions are well-defined. The pair $(E, \mathcal{E})$ denotes a fixed standard Borel space.

Fix a dimension $d \in \mathbb{N}$. The set $S := \mathbb{Z}^d$ is called the *parameter set*, and $(E, \mathcal{E})$ is called the *state space*. Elements of S are called *sites*. A *configuration* is a function ω that assigns to each site $x \in S$ a state $\omega_x \in E$. Write $\Omega := E^S$ for the set of configurations, and $\mathcal{F}$ for the product σ-algebra $\mathcal{E}^S$ on Ω. A *random field* is a probability measure on configurations: the set of random fields is $\mathcal{P}(\Omega, \mathcal{F})$.

A *random surface* is a random field where the standard Borel space $(E, \mathcal{E})$ is either $\mathbb{Z}$ or $\mathbb{R}$ endowed with the standard σ-algebra. In this case, we view configurations as functions from $\mathbb{Z}^d$ to E, and they are also called *height functions*. The symbols ϕ and ψ are preferred over the symbol ω when referring to height functions. In the setting of random surfaces, it makes sense to speak of the *gradient* $\nabla\phi$ of some height function $\phi : \mathbb{Z}^d \to E$. This gives rise in particular to the *gradient σ-algebra* $\mathcal{F}^\nabla$, which is generated by the functions $\phi(y) - \phi(x)$ with x and y ranging over $\mathbb{Z}^d$. Of particular interest to us is the set $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$, the collection of *gradient measures*.

Consider again the general setting of random fields. Write d_1 for the graph metric on the square lattice $\mathbb{Z}^d$. Write $\Lambda \subset\subset \mathbb{Z}^d$ whenever Λ is a finite subset of $\mathbb{Z}^d$. If ω is a configuration, then write $\omega_\Lambda \in E^\Lambda$ for the restriction of ω to Λ, for any subset $\Lambda \subset \mathbb{Z}^d$. If ζ is another configuration and $\Delta \subset \mathbb{Z}^d$ another subset disjoint from Λ, then write $\omega_\Lambda\zeta_\Delta$ for the unique element in $E^{\Lambda\cup\Delta}$ which restricts to ω on Λ and to ζ on Δ. Define the σ-algebras $\mathcal{F}_\Lambda := \sigma(\omega_x : x \in \Lambda)$ and $\mathcal{T}_\Lambda := \sigma(\omega_x : x \in \mathbb{Z}^d \smallsetminus \Lambda)$ for any $\Lambda \subset \mathbb{Z}^d$, and write $\mathcal{T} := \cap_{\Lambda\subset\subset\mathbb{Z}^d}\mathcal{T}_\Lambda$. An event (or function) is called a *cylinder* event (or function) if it is $\mathcal{F}_\Lambda$-measurable for some $\Lambda \subset\subset \mathbb{Z}^d$, and it is called *tail-measurable* if it is $\mathcal{T}$-measurable. In the setting of random surfaces, we furthermore define

$$\mathcal{F}_\Lambda^\nabla := \mathcal{F}_\Lambda \cap \mathcal{F}^\nabla, \qquad \mathcal{T}_\Lambda^\nabla := \mathcal{T}_\Lambda \cap \mathcal{F}^\nabla, \qquad \mathcal{T}^\nabla := \mathcal{T} \cap \mathcal{F}^\nabla.$$

To see convergence of a model on the macroscopic scale, it is essential that the model exhibits shift-invariance on the microscopic scale. Morally, this means that the model is not able to distinguish between different points of Euclidean space and alter its behaviour accordingly. For any $x \in \mathbb{Z}^d$, write $\theta_x : \mathbb{Z}^d \to \mathbb{Z}^d$ for the map $y \mapsto y + x$. Such maps are called *shifts*. Write $\mathcal{L}$ for a fixed full-rank sublattice of $\mathbb{Z}^d$, and write $\Theta = \Theta(\mathcal{L})$ for the group of shifts $\{\theta_x : x \in \mathcal{L}\}$. We shall always choose $\mathcal{L} := \mathbb{Z}^d$ when considering general random fields, and allow other lattices only in the context of random surfaces. If $\omega \in \Omega$ and $\theta \in \Theta$, then $\theta\omega$ denotes the unique configuration satisfying $(\theta\omega)(x) = \omega(\theta x)$ for all x. Similarly, define

$$\theta A := \{\theta\omega : \omega \in A\}, \quad \theta\mathcal{A} := \{\theta A : A \in \mathcal{A}\}, \quad \theta\mu : \theta\mathcal{A} \to [0, \infty], \, \theta\mu(\theta A) \mapsto \mu(A)$$

for $A \subset \Omega$, for $\mathcal{A}$ a sub-σ-algebra of $\mathcal{F}$, and for μ a measure on $\mathcal{A}$. Any of these three objects is called *$\mathcal{L}$-invariant* or *Θ-invariant* if they are invariant under θ for any $\theta \in \Theta$. If $\mathcal{A}$ is an $\mathcal{L}$-invariant σ-algebra on Ω, then write $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{A})$ or $\mathcal{P}_\Theta(\Omega, \mathcal{A})$ for the collection of $\mathcal{L}$-invariant probability measures on $(\Omega, \mathcal{A})$. Note that $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{A})$ is the set of probability measures on $(\Omega, \mathcal{A})$ such that ω and $\theta\omega$ have the same distribution for any $\theta \in \Theta$.

Let $\mathcal{A}$ denote a sub-σ-algebra of $\mathcal{F}$. The *topology of local convergence* or *$\mathcal{L}$-topology* on $\mathcal{P}(\Omega, \mathcal{A})$ is the coarsest topology that makes the map $\mu \mapsto \mu(A)$ continuous for any cylinder event A in $\mathcal{A}$. In the setting of random surfaces, we call a measurable function $f : \Omega \to \mathbb{R}$ a *continuous cylinder function* if f is $\mathcal{F}_\Lambda$-measurable for some $\Lambda \subset\subset \mathbb{Z}^d$, and if f is continuous as a function on E^Λ with respect to the natural topology on E^Λ (recall that $E \in \{\mathbb{Z}, \mathbb{R}\}$). In this case, the *topology of weak local convergence* is the coarsest topology on $\mathcal{P}(\Omega, \mathcal{A})$ that makes the map $\mu \mapsto \mu(f)$ continuous for each bounded continuous cylinder function f. The topology of local convergence and the topology of weak local convergence coincide whenever $E = \mathbb{Z}$.

1.3.2 Specifications and potentials

A *specification* is a family $\gamma = (\gamma_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^d}$ of probability kernels, such that

1. γ_Λ is a probability kernel from $(\Omega, \mathcal{T}_\Lambda)$ to $(\Omega, \mathcal{F})$ for each $\Lambda \subset\subset \mathbb{Z}^d$,
2. $\mu\gamma_\Lambda(A) = \mu(A)$ for any $\Lambda \subset\subset \mathbb{Z}^d$, $A \in \mathcal{T}_\Lambda$, and $\mu \in \mathcal{P}(\Omega, \mathcal{F})$,
3. $\gamma_\Lambda\gamma_\Delta = \gamma_\Lambda$ for any $\Delta \subset \Lambda \subset\subset \mathbb{Z}^d$.

The specification defines the local behaviour of the model, and we think of $\gamma_\Lambda(\cdot, \omega)$ as the *local Gibbs measure* in $\Lambda \subset\subset \mathbb{Z}^d$ with boundary conditions $\omega \in \Omega$. A specification γ is called *$\mathcal{L}$-invariant* if $\gamma_\Lambda(\cdot, \theta\omega) = \theta\gamma_{\theta\Lambda}(\cdot, \omega)$ for any $\Lambda \subset\subset \mathbb{Z}^d$, $\omega \in \Omega$, and $\theta \in \Theta$. Some random field $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ is called a *Dobrushin-Lanford-Ruelle (DLR) measure* for the specification γ if $\mu\gamma_\Lambda = \mu$ for any $\Lambda \subset\subset \mathbb{Z}^d$.

In the setting of random surfaces, we call a specification γ *monotone* if $\phi \leq \psi$ implies $\gamma_\Lambda(\cdot, \phi) \preceq \gamma_\Lambda(\cdot, \psi)$ for any $\Lambda \subset\subset \mathbb{Z}^d$. Call the specification *Lipschitz* if there exists some constant $K \in (0, \infty)$ such that $\gamma_\Lambda(\cdot, \phi)$ is supported on K-Lipschitz functions for any $\Lambda \subset\subset \mathbb{Z}^d$ and for any K-Lipschitz function ϕ (a more subtle notion is introduced at a later stage). We finally use the group structure of $(E, +)$ to define gradient specifications. Call γ a *gradient specification* if the distribution of $\psi + a$ in $\gamma_\Lambda(\cdot, \phi)$ equals that of ψ in $\gamma_\Lambda(\cdot, \phi + a)$ for any $\Lambda \subset\subset \mathbb{Z}^d$, $\phi \in \Omega$, and $a \in E$, where ψ denotes the random height function in each local Gibbs measure. Note that each kernel γ_Λ restricts to a kernel from $(\Omega, \mathcal{T}_\Lambda^\nabla)$ to $(\Omega, \mathcal{F}^\nabla)$ whenever γ is a gradient specification. If μ is a gradient measure and γ a gradient specification, then call μ a *gradient DLR measure* if $\mu\gamma_\Lambda = \mu$ for any $\Lambda \subset\subset \mathbb{Z}^d$.

Let us now return to the general setting of random fields. Although specifications exist in their own right, they are often induced by a potential. An *interaction potential* $\Phi = (\Phi_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^d}$ is a family of *potential functions* $\Phi_\Lambda : \Omega \to \mathbb{R} \cup \{\infty\}$ where each function Φ_Λ is required to be measurable with respect to $\mathcal{F}_\Lambda$. A potential Φ is called *$\mathcal{L}$-invariant* or *periodic* if $\Phi_{\theta\Lambda}(\omega) = \Phi_\Lambda(\theta\omega)$ for all $\theta \in \Theta$ and for any $\omega \in \Omega$. In the sequel, Φ shall always denote a fixed periodic potential. In the setting of random surfaces, it is furthermore assumed that Φ is a *gradient potential*, meaning that each function Φ_Λ is $\mathcal{F}_\Lambda^\nabla$-measurable.

Next, introduce the Hamiltonian. For $\Lambda \subset\subset \mathbb{Z}^d$ and $\Delta \subset \mathbb{Z}^d$ containing Λ, let $H_{\Lambda,\Delta}$ denote the $\mathcal{F}_\Delta$-measurable function from Ω to $\mathbb{R} \cup \{\infty\}$ defined by

$$H_{\Lambda,\Delta} := \sum\nolimits_{\Gamma \subset\subset \mathbb{Z}^d \text{ with } \Gamma \subset \Delta \text{ and with } \Gamma \text{ intersecting } \Lambda} \Phi_\Gamma.$$

In particular, we write $H_\Lambda := H_{\Lambda,\mathbb{Z}^d}$ and $H_\Lambda^0 := H_{\Lambda,\Lambda}$. It is not obvious from this definition that the sum is well-defined, but we shall not concern ourselves with that problem in this section. The function H_Λ is called the *Hamiltonian* of Λ and H_Λ^0 is called the *interior Hamiltonian* of Λ.

If the specification of interest is generated by a potential, then the standard Borel space $(E,\mathcal{E})$ is endowed with a *reference measure* $\lambda \in \mathcal{M}(E,\mathcal{E})$. In the case of random surfaces, this reference measure is always the counting measure (whenever $E=\mathbb{Z}$) or the Lebesgue measure (whenever $E=\mathbb{R}$).

The potential Φ generates a specification $\gamma^\Phi = (\gamma_\Lambda^\Phi)_{\Lambda\subset\subset\mathbb{Z}^d}$ defined by

$$\gamma_\Lambda^\Phi(A,\omega) := \frac{1}{Z_\Lambda^\Phi(\omega)}\int_{E^\Lambda} 1_A(\zeta\omega_{\mathbb{Z}^d\smallsetminus\Lambda})e^{-H_\Lambda^\Phi(\zeta\omega_{\mathbb{Z}^d\smallsetminus\Lambda})}d\lambda^\Lambda(\zeta),$$

for any $\Lambda\subset\subset\mathbb{Z}^d$, $\omega\in\Omega$, and $A\in\mathcal{F}$, where $Z_\Lambda^\Phi(\omega)$ is the normalising constant

$$Z_\Lambda^\Phi(\omega) := \int_{E^\Lambda} e^{-H_\Lambda^\Phi(\zeta\omega_{\mathbb{Z}^d\smallsetminus\Lambda})}d\lambda^\Lambda(\zeta).$$

We drop the superscript Φ in this notation unless the choice of potential is ambiguous. Of course, $\gamma_\Lambda(\cdot,\omega)$ is a well-defined probability measure on $(\Omega,\mathcal{F})$ only if $Z_\Lambda(\omega)\in(0,\infty)$. Say that ω has *finite energy* if $\Phi_\Lambda(\omega)<\infty$ for any $\Lambda\subset\subset\mathbb{Z}^d$, and say that ω is *admissible* if it has finite energy and $Z_\Lambda(\omega)\in(0,\infty)$ for any $\Lambda\subset\subset\mathbb{Z}^d$. To draw a sample ζ from $\gamma_\Lambda(\cdot,\omega)$, set first ζ equal to ω on the complement of Λ, then sample ζ_Λ proportional to $e^{-H_\Lambda}\lambda^\Lambda$. Similarly, if μ is a probability measure on $(\Omega,\mathcal{T}_\Lambda)$ supported on admissible configurations, then $\mu\gamma_\Lambda$ is a probability measure on $(\Omega,\mathcal{F})$; to sample from $\mu\gamma_\Lambda$ one first obtains an auxiliary sample ω from μ; then one draws the final sample ζ from $\gamma_\Lambda(\cdot,\omega)$.

It is important to observe that γ is always a gradient specification in the context of random surfaces. This is due to the fact that Φ is a gradient potential which makes H_Λ measurable with respect to $\mathcal{F}^\nabla$, and because the reference measures λ and λ^Λ are invariant under translations.

1.3.3 The entropy functional and its derivatives

Consider two σ-finite measures $\mu,\nu\in\mathcal{M}(X,\mathcal{X})$ on a standard Borel space $(X,\mathcal{X})$. The *entropy* of μ relative to ν is defined by

$$\mathcal{H}(\mu|\nu) := \begin{cases}\mu(\log f)=\nu(f\log f) & \text{if } \mu\ll\nu \text{ where } f:=d\mu/d\nu,\\ \infty & \text{otherwise.}\end{cases}$$

The *max-entropy* of μ relative to ν is defined by

$$\mathcal{H}^\infty(\mu|\nu) := \log\inf\{\lambda\geq 0:\mu\leq\lambda\nu\} = \begin{cases}\operatorname{ess\,sup}\log f & \text{if } \mu\ll\nu \text{ where } f:=d\mu/d\nu,\\ \infty & \text{otherwise.}\end{cases}$$

The definition of the max-entropy is due to Datta [8]. Note that both entropies are nonnegative when μ and ν are probability measures—if they are indeed probability measures, then each entropy equals zero if and only if $\mu=\nu$. If $\mathcal{Y}$ is a sub-σ-algebra of $\mathcal{X}$, then define $\mathcal{H}_\mathcal{Y}(\mu|\nu) := \mathcal{H}(\mu|_\mathcal{Y}|\nu|_\mathcal{Y})$. Finally, define the *max-diameter* of a nonempty set $\mathcal{A}\subset\mathcal{M}(X,\mathcal{X})$ by

$$\operatorname{Diam}^\infty\mathcal{A} := \sup_{\mu,\nu\in\mathcal{A}}\mathcal{H}^\infty(\mu|\nu)\geq 0,$$

where we observe equality if and only if $\mathcal{A}$ contains exactly one measure. Note that $\operatorname{Diam}^\infty\mathcal{A}<\infty$ if and only if all measures in $\mathcal{A}$ are absolutely continuous with respect to one another, with a uniform bound on the logarithm of the Radon-Nikodym derivatives.

Next comes the introduction of the free energy and the specific free energy. For random fields, these are defined relative to another random field or relative to the specification. For random surfaces, these are defined relative to the potential generating the specification. The definition for general random fields is more flexible; the reason for this is extensively discussed in Chapter 2, as well as in Section 3.2.

Let us first discuss general random fields. Fix some reference random field $\nu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$. If $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ is some other random field, then the *free energy* of μ in some set $\Lambda \subset\subset \mathbb{Z}^d$ with respect to ν is defined by

$$\mathcal{H}_\Lambda(\mu|\nu) := \mathcal{H}_{\mathcal{F}_\Lambda}(\mu|\nu).$$

If μ is $\mathcal{L}$-invariant, then the *specific free energy* of μ with respect to ν is defined by

$$\mathcal{H}(\mu|\nu) := \lim_{n\to\infty} |\Delta_n|^{-1} \mathcal{H}_{\Delta_n}(\mu|\nu),$$

where Δ_n represents the hypercube $\{-n, \ldots, n\}^d \subset\subset \mathbb{Z}^d$. The limit is not convergent in all cases, but we shall impose the appropriate restrictions on ν in Chapter 2 which guarantee consistency of the definition.

If a specification is chosen as reference distribution, then we define the specific free energy instead by

$$\mathcal{H}(\mu|\gamma) := \lim_{n\to\infty} |\Delta_n|^{-1} \mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}),$$

where $\nu \in \mathcal{P}(\Omega, \mathcal{F})$ is an arbitrary random field serving as the mixed boundary condition for the specification. We shall again impose the appropriate restrictions on γ so that this is well-defined and independent of the choice of ν.

For random surfaces, the free energy and the specific free energy are defined with respect to the reference measure λ and the potential Φ. If $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ is a gradient measure, then the *free energy* of μ in Λ with respect to Φ is defined by

$$\mathcal{H}_\Lambda(\mu|\Phi) := \mathcal{H}_{\mathcal{F}_\Lambda^\nabla}(\mu|e^{-H_\Lambda^{0,\Phi}}\lambda^{\Lambda-1}) = \mathcal{H}_{\mathcal{F}_\Lambda^\nabla}(\mu|\lambda^{\Lambda-1}) + \mu(H_\Lambda^{0,\Phi}).$$

Here $\lambda^{\Lambda-1}$ denotes the natural gradient measure obtained by choosing one vertex of Λ as a reference point, and taking the product over the λ for all other sites in Λ. If μ is furthermore $\mathcal{L}$-invariant, then define the *specific free energy* of μ with respect to Φ by the limit

$$\mathcal{H}(\mu|\Phi) := \lim_{n\to\infty} |\Pi_n|^{-1} \mathcal{H}_{\Pi_n}(\mu|\Phi),$$

where $\Pi_n := \{0, \ldots, n-1\}^d \subset\subset \mathbb{Z}^d$. Observe the difference between the sets Δ_n and Π_n; this difference is cosmetic in nature, and both definitions for the specific free energy are in fact invariant under interchanging Δ_n and Π_n.

Finally, in the setting of random surfaces, we can introduce the notion of a *slope*, which gives rise to the *surface tension*. Consider a shift-invariant gradient measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$. If $\phi(y) - \phi(x)$ is μ-integrable for any $x, y \in \mathbb{Z}^d$, then μ is said to have *finite slope*. If μ has finite slope, then shift-invariance of μ implies that the function

$$\mathcal{L} \to \mathbb{R},\, x \mapsto \mu(\phi(x) - \phi(0))$$

is additive. In particular, this means that there is a unique linear functional $u \in (\mathbb{R}^d)^*$ such that

$$u(x) = \mu(\phi(x) - \phi(0))$$

for any $x \in \mathcal{L} \subset \mathbb{R}^d$. This linear functional u is called the *slope* of μ, and we write $S(\mu)$ for it. The *surface tension* at some slope u is now defined to be the infimum of the specific free energy functional $\mathcal{H}(\cdot|\Phi)$ over all shift-invariant gradient measures of that slope; formally, we write

$$\sigma(u) := \inf_{\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla) \text{ with } S(\mu) = u} \mathcal{H}(\mu|\Phi).$$

1.3.4 Constructions for the infinite-range setting

This final subsection introduces some constructions for dealing with the infinite-range setting. The letter γ shall denote the specification of the model of interest throughout. It was already mentioned that our analysis includes models which are not necessarily quasilocal. The constructions in this subsection are original, with the exception of the definition of quasilocality. We introduce these definitions from the perspective of random surfaces, although the definitions are the same for general random fields.

Write π_Λ for the natural probability kernel from $(\Omega, \mathcal{F})$ to $(E^\Lambda, \mathcal{E}^\Lambda)$ which restricts measures to Λ. Consider two finite sets $\Lambda \subset \Delta \subset\subset \mathbb{Z}^d$. Denote by $\mathcal{A}_{\Lambda,\Delta,\phi}$ the set of probability measures on $(E^\Lambda, \mathcal{E}^\Lambda)$ of the form $\mu\gamma_\Lambda\pi_\Lambda$, where μ is any measure in $\mathcal{P}(\Omega, \mathcal{F})$ subject only to $\mu\pi_\Delta = \delta_{\phi_\Delta}$. In other words, $\mathcal{A}_{\Lambda,\Delta,\phi}$ is the set of local Gibbs measures in Λ (and restricted to Λ) given (mixed) boundary conditions which match ϕ on Δ. Write $\mathcal{C}(\mathcal{A})$ for the closure of any set $\mathcal{A} \subset \mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda)$ in the strong topology, and define

$$\mathcal{A}_{\Lambda,\phi} := \cap_{\Delta \subset\subset \mathbb{Z}^d} \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi}).$$

A height function $\phi \in \Omega$ is called a *point of quasilocality* if $\mathcal{A}_{\Lambda,\phi} = \{\delta_\phi\gamma_\Lambda\pi_\Lambda\} = \{\gamma_\Lambda(\cdot, \phi)\pi_\Lambda\}$ for any $\Lambda \subset\subset \mathbb{Z}^d$. Write Ω_γ for the set of points of quasilocality. Call a measure $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ an *almost Gibbs measure* whenever $\mu(\Omega_\gamma) = 1$ and $\mu = \mu\gamma_\Lambda$ for any $\Lambda \subset\subset \mathbb{Z}^d$. The definition of an almost Gibbs measure is the same for gradient measures $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$—noting that Ω_γ is $\mathcal{F}^\nabla$-measurable because γ is a gradient specification. Almost Gibbs measures are called *Gibbs measures* whenever $\Omega_\gamma = \Omega$.

1.4 Main results

1.4.1 General random fields

Chapter 2 investigates the nature of the minimisers of the specific free energy relative to specifications which are *weakly dependent*, a term coined by Lewis, Pfister, and Sullivan. For the definition of weak dependence, we require more notation. Let γ denote the shift-invariant specification of interest, and recall the definition of the restriction kernel π in the previous section. Write $\mathcal{A}_\Lambda(\gamma)$ for the set

$$\mathcal{A}_\Lambda(\gamma) := \{\mu\gamma_\Lambda\pi_\Lambda : \mu \in \mathcal{P}(\Omega, \mathcal{F})\} \subset \mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda).$$

In other words, $\mathcal{A}_\Lambda(\gamma)$ is the set of local Gibbs measures in Λ with arbitrary mixed boundary conditions, and restricted to Λ. The specification γ is called *weakly dependent* if $\operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma) = o(n^d)$ as $n \to \infty$. This means that the interaction between the states of sites in Δ_n with those outside Δ_n is of order $o(|\Delta_n|)$. The weakly dependent framework includes independent percolation models, the random-cluster model, the Ising model, the Loop $O(n)$ model, and the Griffiths singularity model.

General results

Theorem 1.4.1. *Consider a weakly dependent specification γ. Then the map $\mathcal{H}(\cdot|\gamma) : \mathcal{P}_\Theta(\Omega, \mathcal{F}) \to [0, \infty]$ is well-defined with compact lower level sets in the topology of local convergence. Moreover, the set of* minimisers $h_0(\gamma) := \{\mathcal{H}(\cdot|\gamma) = 0\}$ *is nonempty, and the following are equivalent for random fields $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$:*

1. *$\mu \in h_0(\gamma)$,*
2. *$\nu^n \gamma_{\Delta_n} \to \mu$ in the $\mathcal{L}$-topology for some sequence $(\nu^n)_{n\in\mathbb{N}} \subset \mathcal{P}(\Omega, \mathcal{F})$,*
3. *$\mu\pi_{\Delta_n} \in \mathcal{C}(\mathcal{A}_{\Delta_n}(\gamma))$ for each $n \in \mathbb{N}$.*

Moreover, if $\mu \in h_0(\gamma)$, then

1. *μ is almost Gibbs if $\mu(\Omega_\gamma) = 1$,*
2. *μ has finite energy, in the sense of Burton and Keane,*
3. *$\mu^\omega_\Lambda \in \mathcal{A}_{\Lambda,\omega}$ for μ-almost every ω, for each fixed $\Lambda \subset\subset \mathbb{Z}^d$.*

For the last statement, we write μ^ω_Λ for the regular conditional probability distribution of μ on $(E^\Lambda, \mathcal{E}^\Lambda)$ corresponding to the projection map $\Omega \to E^{\mathbb{Z}^d \smallsetminus \Lambda}$.

Finally, we deduce a similar theory for the specific free energy functional relative to a weakly dependent random field. A random field μ is called *weakly dependent* if $\mu \in h_0(\gamma)$ for some weakly dependent specification γ.

Theorem 1.4.2. *If γ is a weakly dependent specification and $\mu \in h_0(\gamma)$, then $\mathcal{H}(\cdot|\mu) : \mathcal{P}_\Theta(\Omega, \mathcal{F}) \to [0, \infty]$ is well-defined, and $\mathcal{H}(\cdot|\mu) = \mathcal{H}(\cdot|\gamma)$.*

It is natural for a random field μ to choose a version of its local regular conditional probability distribution as its specification, which makes μ automatically a DLR measure for that specification. This circumvents however the issue of quasilocality, and it is not true that any weakly dependent random field is an almost Gibbs measure for some choice of specification.

The Loop $O(n)$ model

Let us now discuss two applications of this theory, starting with the Loop $O(n)$ model. A *loop configuration* is a subset of the edge set of the hexagonal lattice with the property that it gives each vertex even degree. It is easy to see that each connected component of a given loop configuration is a closed loop or a bi-infinite path through the hexagonal lattice. The Loop $O(n)$ model has two parameters: they are x and n. The relative weight of a configuration ω is given (informally) by $x^{|\omega|} \cdot n^{|C(\omega)|}$, where $|\omega|$ is the number of edges in ω, and $|C(\omega)|$ the number of loops. The model can alternatively be considered an Ising model on the faces of the hexagonal lattice, together with an extra interaction which counts the number of spin clusters. Our theory, together with a brief analysis of the minimisers of the specific free energy functional, lead to the following result.

Theorem 1.4.3. *For any $x, n \in (0, \infty)$, the Loop $O(n)$ model has shift-invariant almost Gibbs measures.*

The Griffiths singularity model

The Griffiths singularity random field consists of an Ising model in a random percolation environment. The state space is $E = \{-1, 0, 1\}$; the state 0 indicates that a site is closed, while the states ± 1 indicate an open site with that spin. Fix $p \in (0, 1)$ and $\beta \geq 0$. The Griffiths singularity random field $K_{p,\beta} \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ is defined as follows: to draw from $K_{p,\beta}$, one first sets each site to 0 with probability $1 - p$ and to 1 with probability p. Then, one samples an Ising model on the open clusters, say with $+1$ boundary conditions to make sense of the Ising model on the infinite cluster whenever it is present.

The specification corresponding to this random field has long-range interactions, which is perhaps counter-intuitive. This model does, however, fit the weakly dependent framework of Chapter 2. There are three phases. If there is no infinite percolation cluster, then the Ising model decomposes as an infinite product of independent Ising models on finite graphs. If there is an infinite percolation cluster, then there are two options, depending on whether or not the Ising model magnetises on this infinite cluster. It was known that the measure $K_{p,\beta}$ is the only minimiser of the specific free energy functional, and that this is the only DLR state of the natural specification corresponding to $K_{p,\beta}$, whenever there is no infinite percolation cluster. We extend this result to the submagnetic phase.

Theorem 1.4.4. *Consider those pairs of parameters p and β for which*

$$K_{p,\beta}(\omega_0 = -1) = K_{p,\beta}(\omega_0 = 1).$$

Then the variational principle holds true, in the sense that the minimisers of the specific free energy coincide exactly with the set of almost Gibbs measures. Moreover, this set contains only a single measure: the measure $K_{p,\beta}$.

Note that the variational principle fails for this model in the magnetic phase for an appropriate choice of dimension $d \in \mathbb{N}$ and parameters p and β, so the result in this thesis is perhaps the sharpest result which relies on generic arguments only.

1.4.2 Lipschitz random surfaces

The purpose of Chapter 3 is to derive strict convexity of the surface tension for random surface models which are Lipschitz and which are monotone in boundary conditions. The proof includes several auxiliary results which are of independent interest. We provide two significant and original applications of our theory: to *submodular potentials*, that is, potentials which satisfy the Fortuin-Kasteleyn-Ginibre (FKG) lattice condition, and to uniformly random graph homomorphisms from the square lattice $\mathbb{Z}^d$ to a k-regular tree, for any $d, k \geq 2$.

Chapter 3 applies to nearly all random surface models which are monotone and Lipschitz. For the thermodynamical formalism, it is required that the model is furthermore generated by a potential which falls into a certain class of potentials. This class is deliberately chosen as large as possible, and in particular, it does not rule out models with infinite-range interactions. As a consequence, its definition is quite involved.

The Lipschitz constraint

It is important for our arguments that the specification does not only exclude height functions which are not Lipschitz: it must also allow all height functions that *are*

Lipschitz with positive probability (if $E = \mathbb{Z}$) or density (if $E = \mathbb{R}$). This extra property will play a role in the derivation of the large deviations principle and in the application of the classical argument of Burton and Keane in the last step of the proof for strict convexity of the surface tension.

To be as general as possible, we must therefore introduce a more subtle definition of the Lipschitz constraint. The Lipschitz constraint is specified in two steps. First, choose an $\mathcal{L}$-invariant edge set $\mathbb{A}$ on $\mathbb{Z}^d$ which makes $(\mathbb{Z}^d, \mathbb{A})$ into a connected graph of bounded degree. Second, specify upper and lower bounds on each edge in $\mathbb{A}$ in a shift-invariant way. A height function ϕ is now called *Lipschitz* if for each edge $\{x, y\} \in \mathbb{A}$, the difference $\phi(y) - \phi(x)$ falls within the specified bounds. The locally defined Lipschitz constraint induces a so called *local Lipschitz constraint* $q : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}$, which behaves more or less like a metric on $\mathbb{Z}^d$. The advantage of this definition is that a height function ϕ is Lipschitz if and only if $\phi(y) - \phi(x) \leq q(x, y)$ for all $x, y \in \mathbb{Z}^d$, in which case we say that ϕ is *q-Lipschitz*. Write U_q for the interior of the set of slopes $u \in (\mathbb{R}^d)^*$ such that $u|_{\mathcal{L}}$ is q-Lipschitz. This set is convex, and it turns out to be the interior of the set of slopes on which the surface tension σ takes finite values. Finally, in the case that $E = \mathbb{R}$, we write q_ε for the local Lipschitz constraint induced by the same graph $\mathbb{A}$, but with the upper bounds decreased by ε and the lower bounds increased by ε. If a height function ϕ is q_ε-Lipschitz for $\varepsilon > 0$, then small perturbations of it are automatically q-Lipschitz.

The class of potentials

We allow potentials Φ which belong to the class $\mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$, where $\mathcal{S}_{\mathcal{L}}$ contains finite-range potentials which enforce the Lipschitz constraint from the previous passage, and where $\mathcal{W}_{\mathcal{L}}$ encodes the long-range interactions of the model which are not necessarily quasilocal, but which decay at a sufficiently fast rate for the specific free energy to be well-defined. The definition of the second class is reminiscent of the *weakly dependent* context of Chapter 2. We first focus on the class $\mathcal{S}_{\mathcal{L}}$.

Definition 1.4.5. Let Ψ denote an arbitrary periodic gradient potential. The potential Ψ is called *positive* if $\Psi_\Lambda \geq 0$ for any $\Lambda \subset\subset \mathbb{Z}^d$. The potential Ψ is said to have *finite range* if $\Psi_\Lambda \equiv 0$ whenever the diameter of Λ—in the graph metric d_1 on the square lattice—exceeds some fixed constant. The potential Ψ is called *Lipschitz* if there exists a local Lipschitz constraint $(\mathbb{A}, q)$ such that $\Psi_\Lambda(\phi) = \infty$ if and only if $\Lambda = \{x, y\} \in \mathbb{A}$ and $\phi(y) - \phi(x) > q(x, y)$ for some $x, y \in \mathbb{Z}^d$. If $E = \mathbb{R}$ and Ψ Lipschitz with constraint $(\mathbb{A}, q)$, then Ψ is called *locally bounded* if for any $\varepsilon > 0$ sufficiently small, there exists a fixed constant $C_\varepsilon < \infty$, such that

$$H^\Psi_{\{x\}}(\phi) \leq C_\varepsilon$$

for any $x \in \mathbb{Z}^d$ and for any $\phi \in \Omega$ which is q_ε-Lipschitz at x (in the sense that the Lipschitz condition is not violated on any edge of $\mathbb{A}$ incident to x). A potential Ψ is called a *strong interaction* if Ψ has all of the above properties, that is, if Ψ is a positive Lipschitz periodic gradient potential of finite range, and if it is locally bounded in the case that $E = \mathbb{R}$. We shall write $\mathcal{S}_{\mathcal{L}}$ for the collection of strong interactions.

Let us now focus on the class $\mathcal{W}_{\mathcal{L}}$. By an *amenable function* we mean a function f which assigns a number in $[0, \infty)$ to each finite subset of $\mathbb{Z}^d$, such that:

1. $f(\Lambda) = f(\theta\Lambda)$ for all $\Lambda \subset\subset \mathbb{Z}^d$ and for any $\theta \in \Theta$,

2. $f(\Lambda \cup \Delta) \leq f(\Lambda) + f(\Delta)$ for all $\Lambda, \Delta \subset\subset \mathbb{Z}^d$ disjoint,

3. $f(\Pi_n) = o(n^d)$ as $n \to \infty$, where $\Pi_n := \{0, \ldots, n-1\}^d \subset\subset \mathbb{Z}^d$.

Definition 1.4.6. Let Ξ denote an arbitrary periodic gradient potential. The potential Ξ is called *summable* if it has finite norm

$$\|\Xi\| := \sup_{(x,\phi) \in \mathbb{Z}^d \times \Omega} \sum_{\Lambda \subset\subset \mathbb{Z}^d \text{ with } x \in \Lambda} |\Xi_\Lambda(\phi)|.$$

Define, for any $\Lambda \subset\subset \mathbb{Z}^d$,

$$e^-(\Lambda) := \sup_{\phi \in \Omega} \sum_{\Delta \subset\subset \mathbb{Z}^d \text{ with } \Delta \text{ intersecting both } \Lambda \text{ and } \mathbb{Z}^d \smallsetminus \Lambda} |\Xi_\Delta(\phi)|.$$

The function $e^-(\cdot)$ is called the *lower exterior bound* of Ξ. If the potential Ξ is summable and its lower exterior bound $e^-(\cdot)$ amenable, then Ξ is called a *weak interaction*. Write $\mathcal{W}_\mathcal{L}$ for the collection of weak interactions.

Our results apply to potentials Φ which decompose as the sum of a potential $\Psi \in \mathcal{S}_\mathcal{L}$ and a potential $\Xi \in \mathcal{W}_\mathcal{L}$.

Results on the surface tension

Let us first state that the specific free energy is well-defined, and mention some of its properties.

Theorem 1.4.7. *If $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, then the specific free energy functional*

$$\mathcal{H}(\cdot|\Phi) : \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla) \to \mathbb{R} \cup \{\infty\},\ \mu \mapsto \lim_{n\to\infty} n^{-d} \mathcal{H}_{\Pi_n}(\mu|\Phi)$$

is well-defined, affine, bounded below, lower-semicontinuous, and for each $C \in \mathbb{R}$ its lower level set

$$M_C := \{\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla) : \mathcal{H}(\mu|\Phi) \leq C\}$$

is a compact Polish space, with respect to the topology of (weak) local convergence. In fact, the two topologies coincide on each set M_C.

Recall the definition of the surface tension σ in Section 1.3, and write U_Φ for the interior of the set $\{\sigma < \infty\} \subset (\mathbb{R}^d)^*$. It is trivial to demonstrate that σ is convex, because the slope functional as well as the specific free energy functional are affine.

Theorem 1.4.8. *Let Φ denote a potential in $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ which induces a monotone specification.*

1. *If $E = \mathbb{R}$, then σ is strictly convex on U_Φ,*

2. *If $E = \mathbb{Z}$, then σ is strictly convex on U_Φ if for any affine map $h : (\mathbb{R}^d)^* \to \mathbb{R}$ with $h \leq \sigma$, the set $\{h = \sigma\} \cap \partial U_\Phi$ is convex. In particular, σ is strictly convex on U_Φ if at least one of the following conditions is satisfied:*

   (a) σ is affine on ∂U_Φ, but not on $\bar{U}_\Phi$,

   (b) σ is not affine on the line segment $[u_1, u_2]$ for any distinct $u_1, u_2 \in \partial U_\Phi$ such that $[u_1, u_2] \not\subset \partial U_\Phi$.

Moreover, for any potential $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$,

1. *We have $U_\Phi = U_q$,*
2. *If $E = \mathbb{R}$, then $\sigma(u)$ tends to ∞ as u approaches the boundary of U_Φ,*
3. *If $E = \mathbb{Z}$, then σ is bounded and continuous on the closure of U_Φ.*

The extra condition for discrete models is necessary to control the behaviour of ergodic measures whose slope is extremal. We will demonstrate that the extra condition is automatically satisfied for all classical models.

A shift-invariant gradient measure μ is called a *minimiser* if $\mathcal{H}(\mu|\Phi) = \sigma(S(\mu)) < \infty$. Let us finish with a remark regarding the existence of minimisers, which follows directly from the previous two results.

Theorem 1.4.9. *Suppose that $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$. Then for any exposed point $u \in \bar{U}_\Phi$ of σ, there exists an ergodic gradient measure μ of slope u which is also a minimiser. In particular, if σ is strictly convex on U_Φ, then for each $u \in U_\Phi$, there is an ergodic minimiser of that slope.*

The large deviations principle and the variational principle

A *good asymptotic profile* is a bounded open set $D \subset \mathbb{R}^d$ whose boundary has zero Lebesgue measure, together with a continuous function $b : \partial D \to \mathbb{R}$ which extends to some Lipschitz function on $\mathbb{R}^d$ whose gradient lies in U_{q_ε} almost everywhere for some $\varepsilon > 0$. Let (D, b) denote a fixed good asymptotic profile.

Say that a sequence of pairs $(D_n, b_n)_{n\in\mathbb{N}}$ of subsets of $\mathbb{Z}^d$ and height functions is a *good approximation* of (D, b) if $\frac{1}{n}D_n \to D$ in the Hausdorff topology on $\mathbb{R}^d$, and if $\frac{1}{n}\operatorname{Graph} b_n|_{\partial D_n} \to \operatorname{Graph} b$ in the Hausdorff topology on $\mathbb{R}^d \times \mathbb{R}$. We furthermore require that each function b_n is q-Lipschitz whenever $E = \mathbb{Z}$, or q_ε-Lipschitz for some uniform constant $\varepsilon > 0$ whenever $E = \mathbb{R}$.

Fix K minimal subject to $Kd_1 \geq q$. Write $\operatorname{Lip}(\bar{D})$ for the set of $K\|\cdot\|_1$-Lipschitz functions on $\bar{D}$. This set is endowed with the topology of uniform convergence, denoted by $\mathcal{X}^\infty$.

We introduce a map $\mathfrak{G}_n$ to send height functions ϕ to $\operatorname{Lip}(\bar{D})$ for each $n \in \mathbb{N}$. Write first $\bar{\phi}$ for the smallest $K\|\cdot\|_1$-Lipschitz extension of ϕ to $\mathbb{R}^d$ whenever ϕ is q-Lipschitz. The function $\mathfrak{G}_n(\phi)$ is the scaled, restricted version of $\bar{\phi}$:

$$\mathfrak{G}_n(\phi) : \bar{D} \to \mathbb{R},\, x \mapsto \frac{1}{n}\bar{\phi}(nx).$$

Theorem 1.4.10. *Let $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and let $(D_n, b_n)_{n\in\mathbb{N}}$ denote a good approximation of some good asymptotic profile (D, b). Let γ_n^* denote the pushforward of $\gamma_{D_n}(\cdot, b_n)$ along the map $\mathfrak{G}_n$, for any $n \in \mathbb{N}$. Then the sequence of probability measures $(\gamma_n^*)_{n\in\mathbb{N}}$ satisfies a large deviations principle with speed n^d and rate function I on the topological space $(\operatorname{Lip}(\bar{D}), \mathcal{X}^\infty)$. Moreover, the sequence of normalising constants $(Z_n)_{n\in\mathbb{N}} := (Z_{D_n}(b_n))_{n\in\mathbb{N}}$ satisfies $-n^{-d}\log Z_n \to P_\Phi(D, g)$ as $n \to \infty$.*

In this theorem, the rate function $I(f)$ is given by the fundamental integral

$$I(f) := -P_\Phi(D, b) + \int_D \sigma(\nabla f(x))dx$$

if $f|_{\partial D} = b$, and $I(f) := \infty$ otherwise. The constant $P_\Phi(D, b)$ is the associated pressure, *which is defined precisely such that the minimum of I is zero.*

This result leads immediately to the variational principle.

Corollary 1.4.11. *Let $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and let $(D_n, b_n)_{n\in\mathbb{N}}$ denote a good approximation of some good asymptotic profile (D, b). Let γ_n^* denote the pushforward of $\gamma_{D_n}(\cdot, b_n)$ along the map $\mathfrak{G}_n$, for any $n \in \mathbb{N}$. Write f_n for the random function in γ_n^*, which—as a random object—takes values in $\mathrm{Lip}(\bar{D})$. If σ is strictly convex on U_Φ, then the random function f_n converges to the unique minimiser f^* of the rate function I, in probability in the topology of uniform convergence as $n \to \infty$. In other words, f^* is the unique minimiser of the integral*

$$\int_D \sigma(\nabla f(x))dx$$

over all Lipschitz functions $f : \bar{D} \to \mathbb{R}$ which equal b on the boundary of D. If however σ fails to be strictly convex on U_Φ, then for any neighbourhood A of the set of minimisers of the integral in the topology of uniform convergence, we have $f_n \in A$ with high probability as $n \to \infty$.

Properties of minimisers of the specific free energy

The proof relies on various properties of the minimisers of the specific free energy, which are of independent interest and stated here. These results are similar in spirit to what is derived in the weakly dependent setting. We start with a finite energy result.

Theorem 1.4.12. *Consider $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and suppose that $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ is a minimiser. Then for any $\Lambda \subset\subset \mathbb{Z}^d$, we have*

$$1_{\Omega_q}(\mu\pi_{\mathbb{Z}^d \smallsetminus \Lambda} \times \lambda^\Lambda) \ll \mu,$$

where Ω_q denotes the set of q-Lipschitz height functions.

This result is a corollary of the following result concerning the regular conditional probability distributions of μ, which holds true even in the absense of quasilocality. The result appeals to the definitions in Section 1.3.

Theorem 1.4.13. *Consider $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and suppose that $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ is a minimiser. Fix $\Lambda \subset\subset \mathbb{Z}^d$, and write μ^ϕ for the regular conditional probability distribution of μ on $(E^\Lambda, \mathcal{E}^\Lambda)$ corresponding to the projection map $\Omega \to E^{\mathbb{Z}^d \smallsetminus \Lambda}$. Then for μ-almost every $\phi \in \Omega$, we have $\mu^\phi \in \mathcal{A}_{\Lambda,\phi}$. In particular, if $\mu(\Omega_\gamma) = 1$, then μ is an almost Gibbs measure, and if $\Omega_\gamma = \Omega$, then μ is a Gibbs measure.*

Application to submodular potentials

A potential Φ is said to be *submodular* if for every $\Lambda \subset\subset \mathbb{Z}^d$, the potential function Φ_Λ satisfies the FKG lattice condition:

$$\Phi_\Lambda(\phi \wedge \psi) + \Phi_\Lambda(\phi \vee \psi) \le \Phi_\Lambda(\phi) + \Phi_\Lambda(\psi).$$

Sheffield proposes this family of potentials as a natural generalisation of simply attractive potentials, and asks if similar results as the ones proved for simply attractive potentials in [54] could be proved for finite-range submodular potentials. Specifications generated by submodular potentials are automatically monotone. We prove new results in the case that the specification is also Lipschitz.

Theorem 1.4.14. *Let Φ denote a submodular potential in $\mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$. If $E = \mathbb{R}$, then the surface tension σ is strictly convex on U_Φ. If $E = \mathbb{Z}$ and if the Lipschitz constraint q is $\mathbb{Z}^d$-invariant rather than merely $\mathcal{L}$-invariant, then the surface tension σ is strictly convex on U_Φ.*

The extra condition for discrete models is necessary to fulfill the extra condition in our general theory, and is satisfied for all commonly studied models.

Application to random graph homomorphisms from $\mathbb{Z}^d$ to the k-regular tree

It was conjectured by Menz and Tassy in [44] that the surface tension of uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree is strictly convex. In this context, a graph homomorphism is a function from $\mathbb{Z}^d$ to the k-regular tree $\mathcal{T}_k$ which also map the edges of the square lattice to the edges of the tree. The conjecture is confirmed in this thesis.

Let us first give a formal definition of the surface tension or *entropy*, as it is called in [44]. Write U for the set of slopes $u \in (\mathbb{R}^d)^*$ such that $|u(e_i)| < 1$ for each element e_i in the natural basis of $\mathbb{R}^d$. For fixed $u \in \bar{U}$, write $\phi^u : \mathbb{Z}^d \to \mathbb{Z}$ for the graph homomorphism defined by

$$\phi^u(x) := \lfloor u(x) \rfloor + \begin{cases} 0 & \text{if } d_1(0,x) \equiv \lfloor u(x) \rfloor \mod 2, \\ 1 & \text{if } d_1(0,x) \equiv \lfloor u(x) \rfloor + 1 \mod 2. \end{cases}$$

Then ϕ^u approximates u and is nearly linear, in the sense that $\|\phi^u - u|_{\mathbb{Z}^d}\|_\infty \leq 1$. Let g denote a *bi-infinite geodesic* through $\mathcal{T}_k$, that is, a $\mathbb{Z}$-indexed sequence of vertices $g = (g_n)_{n\in\mathbb{Z}} \subset \mathcal{T}_k$ such that $d_{\mathcal{T}_k}(g_n, g_m) = |m - n|$ for any $n, m \in \mathbb{Z}$. The geodesic g is thought of as a copy of $\mathbb{Z}$ in $\mathcal{T}_k$, and is used as reference frame. Write $\tilde{\phi}^u : \mathbb{Z}^d \to \mathcal{T}_k$ for the graph homomorphism defined by $\tilde{\phi}^u(x) := g_{\phi^u(x)}$ for every $x \in \mathbb{Z}^d$. It is shown in [44] that the macroscopic behaviour of uniformly random $\mathcal{T}_k$-valued graph homomorphisms is characterised by the entropy function

$$\mathrm{Ent} : \bar{U} \to [-\log k, 0],\ u \mapsto \lim_{n\to\infty} -n^{-d} \log |\{\tilde{\phi} \in \tilde{\Omega} : \tilde{\phi}_{\mathbb{Z}^d \smallsetminus \Pi_n} = \tilde{\phi}^u_{\mathbb{Z}^d \smallsetminus \Pi_n}\}|,$$

where $\tilde{\Omega}$ denotes the set of all graph homomorphisms from $\mathbb{Z}^d$ to $\mathcal{T}_k$. It is conjectured in [44] that Ent is strictly convex on U, which we prove is correct.

Theorem 1.4.15. *For any $d, k \geq 2$, the entropy function* $\mathrm{Ent} : \bar{U} \to [-\log k, 0]$ *associated to uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree, is strictly convex on U.*

1.4.3 The honeycomb dimer model in higher dimensions

Chapter 4 discusses a generalisation of the honeycomb dimer model to higher dimensions, which was introduced by Linde, Moore, and Nordahl [41]. The generalisation is valid in any dimension $d \geq 2$. For the original two-dimensional model, each sample can be interpreted in several different ways:

1. As a stack of well-supported unit cubes, where each cube at some vertex $\mathbf{x} \in \mathbb{Z}^3$ is supported by a cube at the vertex $\mathbf{x} - \mathbf{e}_i$ for all $i \in \{1, 2, 3\}$,

2. As a lozenge tiling, obtained by projecting the exposed faces of the cubes onto the hyperplane orthogonal to the vector $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$,

3. As a function from $\mathbb{Z}^2$ to $\mathbb{Z} \cup \{-\infty, \infty\}$ which gives the height of each vertical stack of cubes, and which is non-increasing in each coordinate,

4. As a height function on the triangular lattice,

5. As a dimer cover of the hexagonal lattice.

A remarkable aspect of the generalisation is that it genuinely generalises each of these five perspectives. The bijections connecting the first four perspectives are geometric in nature, and are not discussed further in this overview. We shall focus instead on the relation between the fourth and the fifth perspective, and on the main results which are described in terms of these two representations.

The generalised model

The letter $d \geq 2$ denotes the fixed dimension. Let (X^d, E^d) denote the graph obtained from the square lattice $\mathbb{Z}^{d+1}$ by identifying vertices which differ by an integer multiple of the vector $\mathbf{n} := \mathbf{e}_1 + \cdots + \mathbf{e}_{d+1}$. This graph is called the *simplicial lattice*; its vertices are equivalence classes of vertices of the square lattice. Note that (X^d, E^d) is the triangular lattice whenever $d = 2$. Let Ω denote the set of functions $f : X^d \to \mathbb{Z}$ which have the property that $f(\mathbf{0}) \in (d+1)\mathbb{Z}$ and $f([\mathbf{x} + \mathbf{e}_i]) - f([\mathbf{x}]) \in \{-d, 1\}$ for any $\mathbf{x} \in \mathbb{Z}^{d+1}$ and $1 \leq i \leq d+1$. Functions in Ω are called *height functions*. If $f \in \Omega$ and $\Lambda \subset\subset X^d$, then write $\gamma_\Lambda(\cdot, f)$ for the probability measure which draws a height function uniformly at random from all height functions which agree to f on the complement of Λ. This is a model of Lipschitz random surfaces; the differences between this setup and that of Chapter 3 are cosmetic in nature. This means that all results of that chapter apply. In particular, the surface tension associated with this model, is strictly convex.

Generalised loops and the covariance structure

Fix $f \in \Omega$ and $\Lambda \subset\subset X^d$, and write μ for the probability measure $\gamma_\Lambda(\cdot, f)$. Write g for the random function in μ. The first result concerns an identity for the covariance structure of the random function g. In two dimensions, the covariance between $g(\mathbf{x})$ and $g(\mathbf{y})$ is exactly $\frac{9}{2}$ times the expectation of the number of loops in the double dimer model in the measure $\mu \times \mu$ which surround both $\mathbf{x}$ and $\mathbf{y}$. This is due to the fact that the product measure $\mu \times \mu$ is invariant under resampling the orientation of each loop in the double dimer model uniformly at random. The same is true in higher dimensions, and this leads to the following result.

Theorem 1.4.16. *In any dimension $d \geq 2$, the covariance between $g(\mathbf{x})$ and $g(\mathbf{y})$ in the measure μ, equals $(d+1)^2/2$ times the number of generalised loops in $\mu \times \mu$ which surround both $\mathbf{x}$ and $\mathbf{y}$.*

In Chapter 4, we spend a significant amount of time on constructing and studying these generalised loops. We defer the complete, formal statement of the previous result to that chapter. The same decomposition into generalised loops leads to a strongly simplified version of Sheffield's original proof for strict convexity of the surface tension, for this specific model.

The dual lattice

Let us introduce the hypergraph (U^d, H^d) which is dual to the simplicial lattice, and generalises the hexagonal lattice. Each vertex of the simplicial lattice (X^d, E^d) has $2d+2$ neighbours; they are of the form $[\mathbf{x} \pm \mathbf{e}_i] = [\mathbf{x}] \pm \mathbf{e}_i$ for $1 \leq i \leq d+1$. A path $(\mathbf{s}_k)_{0 \leq k \leq n} \subset X^d$ of length $n = d+1$ is called a *simplicial loop* if there exists a permutation $\xi \in S_{d+1}$ such that $\mathbf{s}_k = \mathbf{s}_{k-1} + \mathbf{e}_{\xi(k)}$ for any $1 \leq k \leq d+1$. This implies that $\mathbf{s}$ is closed because $[\mathbf{x}] + \mathbf{e}_1 + \cdots + \mathbf{e}_{d+1} = [\mathbf{x} + \mathbf{n}] = [\mathbf{x}]$. In this introductory chapter, we consider two loops to be equal if they differ only by indexation, and we write U^d for this set of unrooted simplicial loops. Let us index each unrooted loop $\mathbf{s} \in U^d$ such that the increment $\mathbf{e}_{d+1}$ comes first. The loop is then completely characterised by its starting point $\mathbf{s}_0$ and the order $\xi \in S_d$ in which the remaining increments appear. We identify each unrooted simplicial loop $\mathbf{s}$ with the corresponding pair $(\mathbf{s}_0, \xi) \in X^d \times S_d$.

Write $h(e) \subset U^d$ for the set of unrooted simplicial loops traversing e, for any $e \in E^d$. Write H^d for the set $\{h(e) : e \in E^d\}$. The hypergraph (U^d, H^d) is dual to the simplicial lattice (X^d, E^d), and the map $h : E^d \to H^d$ is indeed a bijection. In dimension $d = 2$, the graph (U^d, H^d) is the hexagonal lattice. Observe that $h(e)$ contains one simplicial loop in $X^d \times \{\xi\}$ for any permutation $\xi \in S_d$. The graph (U^d, H^d) is therefore $d!$-regular and $d!$-partite, in the sense that each hyperedge contains exactly one vertex of each member of the partition $\{X^d \times \{\xi\} : \xi \in S_d\}$.

The generalised Kasteleyn theory

The final main result of Chapter 4 is that we express the normalising constant $Z_\Lambda(f)$ in terms of the Cayley hyperdeterminant of the adjacency hypermatrix of (appropriate restrictions) of the hypergraph (U^d, H^d). The normalising constant $Z_\Lambda(f)$, which is also called the *partition function*, equals the number of height functions that equal f on the complement of Λ.

We associate each height function $f \in \Omega$ with the set of edges

$$T(f) := \{\{\mathbf{x}, \mathbf{x} + \mathbf{e}_i\} \in E^d : f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x}) = -d\}.$$

This set characterises the gradient of f. Moreover, since $f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x}) \in \{1, -d\}$, it is easy to see that each simplicial loop contains exactly one edge of $T(f)$. In other words, the set $h(T(f))$ is a perfect matching of the hypergraph (U^d, H^d). We demonstrate in Chapter 4 that each perfect matching of (U^d, H^d) is in fact of the form $h(T(f))$ for some height function $f \in \Omega$. The perfect matchings are thus in bijection with the set $\{T(f) : f \in \Omega\}$.

For any map $A : \{1, \ldots, n\}^{d!} \to \mathbb{C}$, define

$$\operatorname{Det} A := \sum_{\sigma_2, \ldots, \sigma_{d!} \in S_n} \left(\left[\prod_{i=2}^{d!} \operatorname{Sign} \sigma_i \right] \left[\prod_{k=1}^{n} A(k, \sigma_2(k), \ldots, \sigma_{d!}(k)) \right] \right)$$

This expression is called the *Cayley hyperdeterminant* of A, and is the natural generalisation of the determinant of a matrix to objects of higher rank.

Let us now explain how the Cayley hyperdeterminant is related to the number of perfect matchings of (subgraphs of) the hypergraph (U^d, H^d), which is entirely analogous to the classical Kasteleyn theory. Fix an enumeration $\{\xi^1, \ldots, \xi^{d!}\} = S_d$. There is a natural bijection from the set of perfect matchings of (U^d, H^d) to the set of tuples of bijections $(\sigma_i)_{2\leq i\leq d!}$ where $\sigma_i : X^d \to X^d$ and which satisfy

$$\{(\mathbf{x}, \xi^1), (\sigma_2(\mathbf{x}), \xi^2), \ldots, (\sigma_{d!}(\mathbf{x}), \xi^{d!})\} \in H^d$$

for all $\mathbf{x} \in X^d$. In other words, these perfect matchings would correspond exactly to the nonzero terms in the sum in the definition of the determinant of the infinite hypermatrix

$$A : \prod_{i=1}^{d!} X^d \to \{0,1\},\ (\mathbf{x}^1, \ldots, \mathbf{x}^{d!}) \mapsto 1\left(\{(\mathbf{x}^1, \xi^1), \ldots, (\mathbf{x}^{d!}, \xi^{d!})\} \in H^d\right),$$

which is of course not entirely well-defined. The same idea, suitably adapted to the finite setting, yields the desired identity for the normalising constant $Z_\Lambda(f)$. In this chapter, we shall not describe exactly how boundary conditions are encoded in the choice of finite subgraph—for these details, we refer to Chapter 4.

Theorem 1.4.17. *Consider a height function f and a finite set $\Lambda \subset\subset X^d$ such that its complement is connected. Then the partition function $Z_\Lambda(f)$ equals exactly the absolute value of the Cayley hyperdeterminant of the adjacency hypermatrix of the finite subgraph of (U^d, H^d) corresponding to these boundary conditions.*

An interesting aspect of this theorem is that we do not require signs or complex numbers to go from the adjacency hypermatrix to the Kasteleyn hypermatrix. Indeed, it turns out that all nonzero terms in the sum defining the determinant have the same sign. This fact is well-known for the two-dimensional dimer model on the hexagonal lattice.

Chapter 2

Variational principle for weakly dependent random fields

Using an alternative notion of entropy introduced by Datta, the max-entropy, we present a new simplified framework to study the minimisers of the specific free energy for random fields which are weakly dependent in the sense of Lewis, Pfister, and Sullivan. The framework is then applied to derive the variational principle for the Loop $O(n)$ model and the Ising model in a random percolation environment in the nonmagnetic phase, and we explain how to extend the variational principle to similar models. To demonstrate the generality of the framework, we indicate how to naturally fit into it the variational principle for models with an absolutely summable interaction potential, and for the random-cluster model.

2.1 Introduction

2.1.1 Random fields with long-range interactions

One of the great results in statistical physics is the variational principle, which asserts that a shift-invariant infinite-volume measure is a Gibbs measure if and only if it minimises the specific free energy. The class of models which fall under the scope of the variational principle is extremely broad. Models for which the interaction potential is absolutely summable were covered by Georgii [20]. There have been numerous attempts to extend or generalise the variational principle beyond, often in relation to a study of the points where *continuity* or *quasilocality* of the specification fails. Such points are characterised by non-vanishing long-range interactions, and appear naturally in, for example, the random-cluster model [24, 53], the Loop $O(n)$ model [46], and several models in a random environment such as the Ising model in a percolation environment [15]. A non-exhaustive list of the study of the variational principle for specifications which are not quasilocal includes [14, 48, 53, 42, 15, 16, 37]. Further investigation into the variational principle was carried out in relation to renormalisation [38, 18], the large deviations principle [50, 51, 52], and projections or restrictions of Gibbs measures [43, 60]. Other works on the variational principle in the infinite-volume setting include [56, 63, 17]. The variational principle is known to fail for some models, most notably the random field Ising model [37], which was known to exhibit phase transition [3]. Despite those efforts there are still some interesting models for which it is not known if the variational principle holds true or not. Among those are various models of random fields in random environments: a noteworthy example is the Ising model on a random subgraph of the square lattice obtained from independent percolation. The inherent problem derives from the fact that the strength of the interactions between particles does not decay uniformly with the range.

This model belongs to a large, natural class of models known as *weakly dependent*: this term is due to Lewis, Pfister and Sullivan [39]. We develop a streamlined framework for studying the minimisers of the specific free energy within this class. The framework allows one to efficiently deduce the variational principle for many interesting weakly dependent models. Our discussion reviews the absolutely summable setting of [20], and the random-cluster model [53] (see [24] for a general introduction). We break new ground by proving the variational principle for the Ising model in a random environment, in the nonmagnetic phase. This significantly extends the results of [37]. We furthermore deduce the variational principle for the Loop $O(n)$ model (see [46] for a general introduction) by extension of the discussion of the random-cluster model, and we explain how these models represent any model where the nonvanishing long-range interaction is due to a potential associated with clusters, level sets, paths, or other large geometrical objects that arise from the local structure.

2.1.2 The specific free energy

The specific free energy and a suitable characterisation for it are of central importance to the study of the variational principle. A natural first question is thus to ask about restrictions on the model that guarantee that the specific free energy is well-behaved. Candidates are the previously mentioned *weakly dependent* [39], and the more general *asymptotically decoupled*. The latter was introduced by Pfister [47]. While either restriction guarantees a well-defined specific free energy, the former is more amenable to arguments involving regular conditional probability distributions, and is therefore better for studying the variational principle. Remark that we shall define the specific free energy in terms of the specification that characterises the model, unlike in [39, 47] where it is defined in terms of a reference random field. Our definition of *weakly dependent* is therefore cosmetically different.

There is a simple and natural definition of a weakly dependent specification once we introduce the max-entropy of two measures. The max-entropy of some measure μ relative to another measure ν equals

$$\mathcal{H}^\infty(\mu|\nu) := \log\inf\{\lambda \geq 0 : \mu \leq \lambda\nu\},$$

and was introduced by Datta in [8]. We call a specification *weakly dependent* if the max-entropy between any two finite-volume Gibbs measures on a box $\Lambda \subset \mathbb{Z}^d$ is of order $o(|\Lambda|)$ as Λ grows large.

The class of weakly dependent models is rich, and it is not hard to prove that the various models that were mentioned are all weakly dependent. If the model of interest is weakly dependent, then the specific free energy has all the usual properties: its level sets (which are sets of shift-invariant random fields) are compact in the topology of local convergence, and there exist shift-invariant random fields that have zero specific free energy.

2.1.3 Main results

Consider a weakly dependent specification. We call a random field a *minimiser* if it is shift-invariant and has zero specific free energy with respect to this specification. It is a corollary of the definition of the specific free energy that shift-invariant Dobrushin-Lanford-Ruelle (DLR) states are minimisers. We show that a shift-invariant random field is a minimiser if and only if it is a limit of finite-volume Gibbs measures, where we allow mixed boundary conditions. If μ is a minimiser, then we derive properties of the conditional probability distribution of μ in a box Λ, conditioned on what happens outside of Λ. If μ is supported on the points of continuity of the specification corresponding to the model, then we show that μ is a DLR state, and almost Gibbs. In general, we demonstrate that all minimisers have finite energy in the sense of Burton and Keane, so that we are able to make their case for almost sure uniqueness of the infinite cluster (if this is relevant for the model under consideration).

The variational principle asserts that the minimisers of the specific free energy coincide with the shift-invariant almost Gibbs measures. The framework provides a clear route to demonstrating its validity for weakly dependent models: it is sufficient to prove that minimisers of the specific free energy are supported on the points of continuity of the specification, and in deriving this one may assume all the properties that minimisers of the specific free energy automatically have.

We apply the framework to all models that were previously mentioned. First, we show how to fit into our framework the known variational principles for models with an absolutely summable interaction potential [20], and for the random-cluster model [53]. Then, we derive the variational principle for the Loop $O(n)$ model, and by extension we assert that the variational principle must hold true for a large class of models where the long-range interaction is due to weight on percolation clusters (such as for the random-cluster model), level sets, loops, or other large geometrical objects which arise from the local structure. Next, we derive the variational principle for the Ising model in a random percolation environment in the nonmagnetic phase. This improves upon the work of [37], where the same result is established for the phase where the random environment does not percolate. The authors believe that for a large class of models in a random environment, the proposed framework significantly reduces the complexity of determining whether or not the variational principle holds true.

Finally, it should be remarked that in all our work we shall never require the state space to be finite; the framework works for any standard Borel space, much like the setting of Georgii [20].

2.1.4 Structure of the chapter

The chapter is organised as follows. In Section 2.2 we introduce the various mathematical objects necessary to define and study the specific free energy. In Section 2.3 we give a presentation of our main results. In Section 2.4 we show how to define the specific free energy for weakly dependent specifications, and we prove some of its properties. In Section 2.5 we give a characterisation of the minimisers of the specific free energy. In Section 2.6 we show how to derive easily from our framework various versions of the variational principle.

2.2 Definitions

If $(X, \mathcal{X})$ is any measurable space, then write $\mathcal{P}(X, \mathcal{X})$ for the set of probability measures on $(X, \mathcal{X})$, and $\mathcal{M}(X, \mathcal{X})$ for the set of σ-finite measures μ with $\mu(X) > 0$. In this chapter we only consider measurable spaces that are standard Borel spaces.

We shall follow the notation of Georgii [20] wherever possible.

2.2.1 Random fields

We are concerned with the study of random fields. Fix a dimension $d \in \mathbb{N}$ and a standard Borel space $(E, \mathcal{E})$ throughout this chapter. The set $S := \mathbb{Z}^d$ is called the *parameter set*, and $(E, \mathcal{E})$ is called the *state space*. Elements of S are called *sites*. A *configuration* is a function ω that assigns to each site $x \in S$ a state $\omega_x \in E$. Write $\Omega := E^S$ for the set of configurations, and $\mathcal{F}$ for the product σ-algebra $\mathcal{E}^S$ on Ω. A *random field* is a probability measure on configurations: the set of random fields is $\mathcal{P}(\Omega, \mathcal{F})$.

Define, for each site $x \in S$, the measurable function $\pi_x : \Omega \to E$, $\omega \mapsto \omega_x$. For any $\Lambda \subset S$, we shall write $\mathcal{F}_\Lambda := \sigma(\pi_x : x \in \Lambda) \subset \mathcal{F}$. Write furthermore π_Λ for the canonical projection map $\Omega = E^S \to E^\Lambda$, and observe that π_Λ extends canonically to a bijection from $\mathcal{F}_\Lambda$ to $\mathcal{E}^\Lambda$ and to a bijection from $\mathcal{P}(\Omega, \mathcal{F}_\Lambda)$ to $\mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda)$. Define $\omega_\Lambda := \pi_\Lambda(\omega)$ for $\omega \in \Omega$, and if $\mu \in \mathcal{P}(\Omega, \mathcal{X})$ for some $\mathcal{F}_\Lambda \subset \mathcal{X} \subset \mathcal{F}$, then write $\mu_\Lambda := \pi_\Lambda(\mu) \in \mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda)$. If f is an $\mathcal{F}_\Lambda$-measurable function on Ω and g an $\mathcal{E}^\Lambda$-measurable function on E^Λ, then we shall without further notice write f for the $\mathcal{E}^\Lambda$-measurable function $f \circ \pi_\Lambda^{-1}$ on E^Λ and g for the $\mathcal{F}_\Lambda$-measurable function $g \circ \pi_\Lambda$ on Ω. Finally, if $\Lambda \subset \Delta \subset S$, then write also π_Λ for the canonical projection map $E^\Delta \to E^\Lambda$, and if $\omega \in E^\Lambda$ and $\zeta \in E^{\Delta \smallsetminus \Lambda}$, then write $\omega\zeta$ for the unique element of E^Δ such that $\pi_\Lambda(\omega\zeta) = \omega$ and $\pi_{\Delta \smallsetminus \Lambda}(\omega\zeta) = \zeta$.

Define, for every $x \in \mathbb{Z}^d$, the map $\theta_x : \mathbb{Z}^d \to \mathbb{Z}^d$, $y \mapsto y + x$. Each map θ_x is called a *shift*. Write Θ for the set of shifts, that is, $\Theta = \{\theta_x : x \in \mathbb{Z}^d\}$. If $\omega \in \Omega$ and $\theta \in \Theta$, then write $\theta\omega$ for the configuration in Ω satisfying $(\theta\omega)_x = \omega_{\theta x}$ for every $x \in S$. Similarly, define $\theta A := \{\theta\omega : \omega \in A\}$ for $A \in \mathcal{F}$. A random field $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ is called *shift-invariant* if $\mu(\theta A) = \mu(A)$ for any $A \in \mathcal{F}$ and $\theta \in \Theta$. Write $\mathcal{P}_\Theta(\Omega, \mathcal{F})$ for the collection of shift-invariant random fields.

2.2.2 Entropy and max-entropy

Consider two σ-finite measures $\mu, \nu \in \mathcal{M}(X, \mathcal{X})$ on a standard Borel space $(X, \mathcal{X})$. The *entropy* of μ relative to ν is defined by

$$\mathcal{H}(\mu|\nu) := \begin{cases} \mu(\log f) = \nu(f \log f) & \text{if } \mu \ll \nu \text{ where } f := d\mu/d\nu, \\ \infty & \text{otherwise.} \end{cases}$$

The *max-entropy* of μ relative to ν is defined by

$$\mathcal{H}^\infty(\mu|\nu) := \log \inf\{\lambda \geq 0 : \mu \leq \lambda\nu\} = \begin{cases} \operatorname{ess\,sup} \log f & \text{if } \mu \ll \nu \text{ where } f := d\mu/d\nu, \\ \infty & \text{otherwise.} \end{cases}$$

Note that both entropies are nonnegative when μ and ν are probability measures—if they are indeed probability measures, then each entropy equals zero if and only if $\mu = \nu$. If $\mathcal{Y}$ is a sub-σ-algebra of $\mathcal{X}$, then define $\mathcal{H}_\mathcal{Y}(\mu|\nu) := \mathcal{H}(\mu|_\mathcal{Y}|\nu|_\mathcal{Y})$. If $(X, \mathcal{X}) = (\Omega, \mathcal{F})$ and $\Lambda \in \mathcal{S}$, then abbreviate $\mathcal{H}_{\mathcal{F}_\Lambda}(\mu|\nu)$ to $\mathcal{H}_\Lambda(\mu|\nu)$. Introduce a similar definition for $\mathcal{H}^\infty_\Lambda(\mu|\nu)$. Finally, define the *max-diameter* of a nonempty set $\mathcal{A} \subset \mathcal{M}(X, \mathcal{X})$ by

$$\operatorname{Diam}^\infty \mathcal{A} := \sup_{\mu,\nu \in \mathcal{A}} \mathcal{H}^\infty(\mu|\nu) \geq 0,$$

where we observe equality if and only if $\mathcal{A}$ contains exactly one measure.

For probability measures $\mu, \nu \in \mathcal{P}(X, \mathcal{X})$, we always have $\mathcal{H}(\mu|\nu) \leq \mathcal{H}^\infty(\mu|\nu)$. It is possible however that $\mathcal{H}^\infty(\mu|\nu)$ is large and $\mathcal{H}(\mu|\nu)$ small, for example if the Radon-Nikodym derivative $f := d\mu/d\nu$ is large on a very small portion of $(X, \mathcal{X})$. The max-entropy should be understood as a sort of L^∞-version of the usual entropy. The max-entropy and the max-diameter prove to be efficient tools for selecting the class of models for which the theory in this chapter works. The usual entropy however, is sometimes easier to work with due to a number of standard identities that are available; see for example (2.4.2) in the proof of Lemma 2.4.1.

2.2.3 Weakly dependent specifications

A *specification* is a family $\gamma = (\gamma_\Lambda)_{\Lambda \in \mathcal{S}}$ with the following properties:

1. Each member γ_Λ is a probability kernel from $(\Omega, \mathcal{F}_{S \smallsetminus \Lambda})$ to $(\Omega, \mathcal{F})$,
2. Each member γ_Λ satisfies $\gamma_\Lambda(A, \omega) = 1(\omega \in A)$ whenever $A \in \mathcal{F}_{S \smallsetminus \Lambda}$,
3. If $\Lambda \subset \Delta \in \mathcal{S}$, then $\gamma_\Delta = \gamma_\Delta \gamma_\Lambda$.

A member γ_Λ is called *proper* if it has the second property; the family γ is called *consistent* if it has the third property. We fix a specification γ throughout this chapter. The specification γ is called *shift-invariant* if $\gamma_{\theta\Lambda}(A, \omega) = \gamma_\Lambda(\theta A, \theta\omega)$ for any $\Lambda \in \mathcal{S}$, $A \in \mathcal{F}$, $\omega \in \Omega$, $\theta \in \Theta$.

Fix $\Lambda \in \mathcal{S}$, and consider γ_Λ: this is a probability kernel from $(\Omega, \mathcal{F}_{S \smallsetminus \Lambda})$ to $(\Omega, \mathcal{F})$. Write $\hat{\gamma}_\Lambda$ for the unique probability kernel from $(\Omega, \mathcal{F}_{S \smallsetminus \Lambda})$ to $(E^\Lambda, \mathcal{E}^\Lambda)$ such that $\hat{\gamma}_\Lambda(\cdot, \omega) = \pi_\Lambda(\gamma_\Lambda(\cdot, \omega))$ for every $\omega \in \Omega$. The measure $\hat{\gamma}_\Lambda(\cdot, \omega)$ is the *finite-volume Gibbs measure* on $(E^\Lambda, \mathcal{E}^\Lambda)$ with *deterministic boundary conditions* ω. Of course, the original kernel γ_Λ can be recovered from $\hat{\gamma}_\Lambda$ through the equation $\gamma_\Lambda(\cdot, \omega) = \hat{\gamma}_\Lambda(\cdot, \omega) \times \delta_{\omega_{S \smallsetminus \Lambda}}$—this is because γ_Λ is proper. It is often more convenient to define $\hat{\gamma}_\Lambda$ than γ_Λ when describing a specific model.

Now fix a random field $\mu \in \mathcal{P}(\Omega, \mathcal{F})$, and consider the finite-volume measure $\mu\hat{\gamma}_\Lambda$. This is the *finite-volume Gibbs measure* on $(E^\Lambda, \mathcal{E}^\Lambda)$ with *mixed boundary conditions* μ. Define

$$\mathcal{A}_\Lambda(\gamma) := \{\mu\hat{\gamma}_\Lambda : \mu \in \mathcal{P}(\Omega, \mathcal{F})\} \subset \mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda) :$$

the set of all such finite-volume Gibbs measures. This set is convex because the set of all random fields is convex. For each $n \in \mathbb{N}$, we use the notation Δ_n for the box

$$\Delta_n := \{-n, \ldots, n\}^d \in \mathcal{S}.$$

The specification γ is called *weakly dependent* if γ is shift-invariant and satisfies

$$\operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma) = o(|\Delta_n|)$$

as $n \to \infty$. For technical reasons we also require that $\operatorname{Diam}^\infty \mathcal{A}_\Lambda(\gamma)$ is finite for any $\Lambda \in \mathcal{S}$; this additional condition is not restrictive. Write $\mathbb{S}$ for the collection of weakly dependent specifications.

Before proceeding, it is useful to remark that

$$\operatorname{Diam}^\infty \mathcal{A}_\Lambda(\gamma) := \sup_{\mu,\nu} \mathcal{H}^\infty(\mu\hat{\gamma}_\Lambda|\nu\hat{\gamma}_\Lambda) = \sup_{\omega,\zeta} \mathcal{H}^\infty(\hat{\gamma}_\Lambda(\cdot, \omega)|\hat{\gamma}_\Lambda(\cdot, \zeta));$$

it is sufficient to consider deterministic boundary conditions in calculating the max-diameter of $\mathcal{A}_\Lambda(\gamma)$. This can be deduced from Fubini's theorem without effort.

2.2.4 The specific free energy

Consider a shift-invariant random field μ and a weakly dependent specification γ. The *specific free energy (SFE)* of μ relative to γ is defined by

$$h(\mu|\gamma) := \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}) \in [0,\infty]$$

where $\nu \in \mathcal{P}(\Omega,\mathcal{F})$. Lemma 2.4.4 asserts that the limit exists for any ν, and that this limit is independent of the choice of ν. A shift-invariant random field μ with $h(\mu|\gamma) = 0$ is called a *minimiser* of γ. Write $h_0(\gamma)$ for the set of minimisers of γ.

Now take the perspective of a shift-invariant random field μ. The random field μ is called *weakly dependent* if $\mu \in h_0(\gamma)$ for some weakly dependent specification γ. Write $\mathbb{F}$ for the collection of weakly dependent random fields. If μ is an arbitrary shift-invariant random field and ν a weakly dependent random field, then the *specific free energy (SFE)* of μ relative to ν is defined by

$$h(\mu|\nu) := \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\Delta_n}(\mu|\nu) \in [0,\infty].$$

Lemma 2.5.10 asserts that the limit converges for any choice of μ and ν. The quantity $h(\mu|\nu)$ is also sometimes called the *entropy density* of μ with respect to ν. Write $h_0(\nu)$ for the set of shift-invariant random fields μ with $h(\mu|\nu) = 0$. Measures $\mu \in h_0(\nu)$ are called *minimisers* of ν.

2.2.5 DLR states

Now consider a random field μ and a finite set $\Lambda \in \mathcal{S}$. Write μ_Λ^ω for the regular conditional probability distribution (r.c.p.d.) on $(E^\Lambda, \mathcal{E}^\Lambda)$ of μ corresponding to the projection map $\pi_{S\smallsetminus\Lambda} : \Omega \to E^{S\smallsetminus\Lambda}$. Informally, this is the distribution of ω_Λ in μ given the states of ω outside Λ. Suppose that we are given an arbitrary specification γ. A *Dobrushin-Lanford-Ruelle (DLR) state* is a random field $\mu \in \mathcal{P}(\Omega,\mathcal{F})$ which satisfies the *DLR equation* $\mu = \mu\gamma_\Lambda$ for every $\Lambda \in \mathcal{S}$. In other words, μ is a DLR state if and only if $\mu_\Lambda^\omega = \hat{\gamma}(\cdot,\omega)$ for μ-almost every $\omega \in \Omega$, for each $\Lambda \in \mathcal{S}$. Write $\mathcal{G}(\gamma)$ for the set of DLR states, and $\mathcal{G}_\Theta(\gamma) := \mathcal{G}(\gamma) \cap \mathcal{P}_\Theta(\Omega,\mathcal{F})$ for the set of shift-invariant DLR states.

2.2.6 Topologies

The *topology of local convergence* or *$\mathcal{L}$-topology* on Ω is the coarsest topology on Ω that makes the map $\omega \mapsto \omega_x$ continuous for every $x \in \mathbb{Z}^d$, with respect to the discrete topology on E. This means that $\omega^n \to \omega$ if and only if for any $\Lambda \in \mathcal{S}$, we have $\omega^n_\Lambda = \omega_\Lambda$ for n sufficiently large.

Consider an arbitrary standard Borel space $(X,\mathcal{X})$. The *strong topology* on $\mathcal{M}(X,\mathcal{X})$ is the coarsest topology that makes the map $\mu \mapsto \mu(A)$ continuous for every $A \in \mathcal{X}$. If $\mathcal{A} \subset \mathcal{P}(X,\mathcal{X})$ is a convex set of probability measures subject to $\operatorname{Diam}^\infty \mathcal{A}$ being finite, then write $\mathcal{C}(\mathcal{A})$ for the closure of $\mathcal{A}$ in the strong topology. In Lemma 2.5.1 we present an alternative definition for $\mathcal{C}(\mathcal{A})$, which we demonstrate is equivalent.

The *topology of local convergence* or *$\mathcal{L}$-topology* on $\mathcal{P}(\Omega,\mathcal{F})$ is the coarsest topology on $\mathcal{P}(\Omega,\mathcal{F})$ that makes the map $\mu \mapsto \mu(A)$ continuous for every $A \in \cup_{\Lambda\in\mathcal{S}}\mathcal{F}_\Lambda$. This means that $\mu^n \to \mu$ in the $\mathcal{L}$-topology if and only if $\pi_\Lambda(\mu^n) \to \mu_\Lambda$ in the strong topology on $\mathcal{P}(E^\Lambda,\mathcal{E}^\Lambda)$ for every $\Lambda \in \mathcal{S}$.

Remark that we do not assume a topology on the state space E. A topology is not even implied, because the $\mathcal{L}$-topology on measures originates from the strong topology on measures. In some sense, the $\mathcal{L}$-topology thus alludes to the discrete topology on E—this holds true both when considered a topology on Ω, and when considered a topology on $\mathcal{P}(\Omega, \mathcal{F})$.

2.2.7 Limits of finite-volume Gibbs measures

Let γ be a weakly dependent specification. Write $\mathcal{W}(\gamma)$ for the set of limits of finite-volume Gibbs measures in the $\mathcal{L}$-topology, that is,

$$\mathcal{W}(\gamma) := \{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \nu^n \gamma_{\Delta_n} \to \mu \text{ in the } \mathcal{L}\text{-topology for some } (\nu^n)_{n\in\mathbb{N}} \subset \mathcal{P}(\Omega, \mathcal{F})\}.$$

It is immediate that $\mathcal{G}(\gamma) \subset \mathcal{W}(\gamma)$: if $\mu \in \mathcal{G}(\gamma)$, then $\mu\gamma_{\Delta_n} \to \mu$ and therefore $\mu \in \mathcal{W}(\gamma)$. We write $\nu^n \gamma_{\Delta_n}$ in this definition and not $\nu^n \hat{\gamma}_{\Delta_n}$ so that all measures live in the same space and convergence in the $\mathcal{L}$-topology makes sense. For simplicity the definition is in terms of the exhaustive sequence $(\Delta_n)_{n\in\mathbb{N}}$; it is straightforward to verify that the definition is the same if we replace this sequence by any other increasing exhaustive sequence. Write $\mathcal{W}_\Theta(\gamma) := \mathcal{W}(\gamma) \cap \mathcal{P}_\Theta(\Omega, \mathcal{F})$. We shall later see that $h_0(\gamma) = \mathcal{W}_\Theta(\gamma)$.

2.2.8 Continuity of the specification

Consider a weakly dependent specification γ. We are going to define more sets of finite-volume Gibbs measures, now restricting the boundary conditions that are allowed. For any $\Lambda, \Delta \in \mathcal{S}$ and $\omega \in \Omega$, define

$$\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma) := \{\mu\hat{\gamma}_\Lambda : \mu \in \mathcal{P}(\Omega, \mathcal{F}) \text{ such that } \mu_\Delta = \delta_{\omega_\Delta}\} \subset \mathcal{A}_\Lambda(\gamma).$$

The sets $\mathcal{A}_\Lambda(\gamma)$ and $\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma)$ are convex, and $\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma)$ is decreasing in Δ. Define

$$\mathcal{A}_{\Lambda,\omega}(\gamma) := \cap_{\Delta\in\mathcal{S}} \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma)) = \cap_{n\in\mathbb{N}} \mathcal{C}(\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma)).$$

Consider a measure $\mu \in \mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda)$. Then $\mu \in \mathcal{A}_{\Lambda,\omega}$ if and only if $\nu^n \hat{\gamma}_\Lambda \to \mu$ in the strong topology for some sequence of random fields $(\nu^n)_{n\in\mathbb{N}}$ converging to δ_ω in the $\mathcal{L}$-topology.

Observe that the alternative characterisation of $\mathcal{A}_{\Lambda,\omega}(\gamma)$ implies the inclusion $\delta_\omega \hat{\gamma}_\Lambda = \hat{\gamma}_\Lambda(\cdot, \omega) \in \mathcal{A}_{\Lambda,\omega}(\gamma)$. Define

$$\begin{aligned}\Omega_\gamma &:= \{\omega \in \Omega : \mathcal{A}_{\Lambda,\omega}(\gamma) = \{\hat{\gamma}_\Lambda(\cdot, \omega)\} \text{ for any } \Lambda \in \mathcal{S}\} \\ &= \{\omega \in \Omega : |\mathcal{A}_{\Lambda,\omega}(\gamma)| = 1 \text{ for any } \Lambda \in \mathcal{S}\}.\end{aligned}$$

In other words, Ω_γ is the set of configurations $\omega \in \Omega$ such that the map $\zeta \mapsto \gamma_\Lambda(\cdot, \zeta)$ is continuous—both sides endowed with the $\mathcal{L}$-topology—at ω for any $\Lambda \in \mathcal{S}$. If $\omega \in \Omega_\gamma$, then we say that the specification γ is *continuous* or *quasilocal* at ω. If $\Omega_\gamma = \Omega$, then each DLR state of γ is also called a *Gibbs measure*. If $\mu \in \mathcal{G}(\gamma)$ and $\mu(\Omega_\gamma) = 1$, then μ is called an *almost Gibbs measure*. This makes sense even if $\Omega_\gamma \neq \Omega$.

2.3 Main results

2.3.1 The specific free energy

Consider a weakly dependent specification γ. We prove that for any shift-invariant random field μ, the SFE

$$h(\mu|\gamma) := \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}) \in [0,\infty]$$

is well-defined, and independent of the choice of $\nu \in \mathcal{P}(\Omega, \mathcal{F})$ (Lemma 2.4.4). Moreover, we show that the level sets of the SFE—given by $\{h(\cdot|\gamma) \leq C\} \subset \mathcal{P}_\Theta(\Omega, \mathcal{F})$ for $C \in [0,\infty)$—are compact in the topology of local convergence, and that $h_0(\gamma) = \{h(\cdot|\gamma) = 0\}$ is nonempty (Lemma 2.4.10). We prove the first half of the variational principle, which asserts that $\mathcal{G}_\Theta(\gamma) \subset h_0(\gamma)$ (Corollary 2.4.9).

2.3.2 Minimisers of the specific free energy

Next, we focus on the set of minimisers $h_0(\gamma)$ of the weakly dependent specification γ. We find some alternative characterisations for the set of minimisers. In particular, if μ is a shift-invariant random field, then the following are equivalent:

1. $\mu \in h_0(\gamma)$, that is, μ is a minimiser of γ,
2. $\mu \in \mathcal{W}(\gamma)$, that is, μ is a limit of finite-volume Gibbs measures,
3. $\mu_{\Delta_n} \in \mathcal{C}(\mathcal{A}_{\Delta_n}(\gamma))$ for each $n \in \mathbb{N}$;

see Lemma 2.5.6 and Corollary 2.5.5. Moreover, if μ is a minimiser, then we demonstrate that

1. μ is almost Gibbs if $\mu(\Omega_\gamma) = 1$,
2. $\mu_\Lambda^\omega \in \mathcal{A}_{\Lambda,\omega}$ for μ-almost every ω, for each $\Lambda \in \mathcal{S}$,
3. μ has finite energy, in the sense of Burton and Keane.

The first statement follows almost immediately from the second, see Lemma 2.5.7 and Corollary 2.5.8. The third statement requires a short argument, see Corollary 2.5.9.

2.3.3 The relation between $\mathbb{F}$ and $\mathbb{S}$

Now take a more abstract viewpoint, and consider the set of all weakly dependent random fields $\mathbb{F}$. Choose a weakly dependent specification $\gamma \in \mathbb{S}$ and a minimiser $\nu \in \mathbb{F}$ of γ. First, we prove that $h(\mu|\nu)$ is well-defined and equal to $h(\mu|\gamma)$ for any shift-invariant random field μ (Lemma 2.5.10). This implies in particular that $h_0(\nu) = h_0(\gamma)$. For $\mu, \nu \in \mathbb{F}$, we declare $\mu \sim \nu$ if $h(\mu|\nu) = 0$. We prove that $\sim$ is an equivalence relation. Write $\mathbb{F}^*$ for the partition of $\mathbb{F}$ into equivalence classes. This provides a canonical way to partition the set of specifications $\mathbb{S}$ as well: define the map

$$\Xi : \mathbb{S} \to \mathbb{F}^*, \gamma \mapsto h_0(\gamma),$$

and write $\mathbb{S}^*$ for the partition of $\mathbb{S}$ into the level sets of Ξ. This makes Ξ into a bijection from $\mathbb{S}^*$ to $\mathbb{F}^*$—the original map Ξ was surjective by definition a weakly dependent random field.

2.3.4 The variational principle in the weakly dependent setting

Consider a weakly dependent specification γ. The previous results provide efficient machinery for attacking the variational principle. Consider an arbitrary shift-invariant random field μ. The variational principle asserts that

$$\mu \in h_0(\gamma) \iff \mu \text{ is almost Gibbs with respect to } \gamma. \tag{2.3.1}$$

To prove the variational principle for the model of interest, we must always derive two results. First, we must show that the specification γ corresponding to the model is indeed weakly dependent. Second, one must show that $\mu(\Omega_\gamma) = 1$ for any minimiser μ of γ. The variational principle then follows from Corollaries 2.4.9 and 2.5.8.

Once weak dependence of the specification has been established, the systematic study of the minimisers of the SFE provides a number of useful properties that minimisers of the SFE automatically have—see Subsection 2.3.2. This usually makes it easier to prove that $\mu(\Omega_\gamma) = 1$ for arbitrary minimisers μ.

We chose to formulate the variational principle with respect to the standard entropy functional $\mathcal{H}$. It is also possible to use the max-entropy $\mathcal{H}^\infty$ for this purpose. To that end, simply replace the set $h_0(\gamma)$ in (2.3.1) with the set

$$h_0^\infty(\gamma) := \left\{\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F}) : |\Delta_n|^{-1}\mathcal{H}^\infty_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}) \to 0\right\}.$$

We shall derive that $h_0^\infty(\gamma) = h_0(\gamma)$ for any weakly dependent specification γ. The inclusion $h_0^\infty(\gamma) \subset h_0(\gamma)$ follows from the fact that $\mathcal{H}^\infty(\mu|\nu) \geq \mathcal{H}(\mu|\nu)$ for any $\mu, \nu \in \mathcal{P}(X, \mathcal{X})$. The other inclusion follows from Lemma 2.5.1 and Corollary 2.5.5, jointly with the definition of a weakly dependent specification.

2.3.5 Applications

The weakly dependent setting is very general: it contains most nonpathological non-gradient models that do not have some form of combinatorial exclusion (such as for example the dimer models, which have a non-gradient interpretation but which are not weakly dependent). We start by showing how to naturally fit two known variational principles into our framework. Then we derive the variational principle for the Loop $O(n)$ model, and finally we derive new results for the Ising model in a random percolation environment.

In Subsection 2.6.1, we show how to efficiently derive the variational principle for models that are defined in terms of an absolutely summable interaction potential. This setting is treated in the classical work of Georgii [20]. For such models we find that $\Omega = \Omega_\gamma$, meaning that all almost Gibbs measures are in fact Gibbs. In Subsection 2.6.2, we show how to derive the variational principle for the random-cluster model. The original proof is due to Seppäläinen [53]. The proofs (the one of Seppäläinen and the one presented here) rely on the finite energy of minimisers of the SFE, which implies that there is at most one infinite cluster almost surely with respect to such measures (see Burton and Keane [4]). In Subsection 2.6.3, we discuss how to derive the variational principle for the Loop $O(n)$ model, by analogy with the random-cluster model. This result is new. We also discuss how to derive the variational principle for similar models. In Subsection 2.6.4, we prove the variational principle for the Ising model in a random percolation environment, in the nonmagnetic phase. Moreover, we demonstrate that the minimiser of the SFE is unique. This is a new result. The variational principle was previously derived for the subcritical percolation phase in [37].

2.4 The specific free energy

This section has two main goals. The first goal is to prove Lemma 2.4.4, which asserts that the SFE is well-defined for weakly dependent specifications. It also provides some useful identities. As an immediate corollary we observe that DLR states minimise the SFE. The second goal is to prove Lemma 2.4.10, which asserts that the level sets of the SFE are compact in the $\mathcal{L}$-topology, and that there exist measures with zero SFE.

2.4.1 Consistency of the definition

The definition of the SFE relies on two key lemmas. Lemma 2.4.1 concerns super-additivity of a useful quantity. Lemma 2.4.3 bounds the difference of two relative entropies in terms of the max-entropy.

Lemma 2.4.1. *Let γ denote any specification and μ a random field. Consider a finite pairwise disjoint family of finite sets $(\Lambda_k)_{1\leq k\leq n} \subset \mathcal{S}$, and write $\Lambda := \cup_k \Lambda_k \in \mathcal{S}$. Then*

$$\inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_\Lambda(\mu|\rho\gamma_\Lambda) \geq \sum_k \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Lambda_k}(\mu|\rho\gamma_{\Lambda_k}).$$

Proof. Fix $\nu \in \mathcal{P}(\Omega,\mathcal{F})$, and replace ν by $\nu\gamma_\Lambda$ if the two are not equal. We must demonstrate that

$$\mathcal{H}_\Lambda(\mu|\nu) \geq \sum_k \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Lambda_k}(\mu|\rho\gamma_{\Lambda_k}).$$

By induction, it is sufficient to consider the case $n = 2$. We have

$$\mathcal{H}_\Lambda(\mu|\nu) = \mathcal{H}_{\Lambda_1}(\mu|\nu) + \int_{E^{\Lambda_1}} \mathcal{H}_{\Lambda_2}(\mu^\zeta|\nu^\zeta) d\mu_{\Lambda_1}(\zeta), \tag{2.4.2}$$

where μ^ζ and ν^ζ denote the r.c.p.d. on $(\Omega,\mathcal{F})$ of μ and ν respectively corresponding to the projection map $\Omega \to E^{\Lambda_1}$. Recall that $\nu = \nu\gamma_\Lambda$. For the first term on the right in (2.4.2), consistency of γ implies that $\nu = \nu\gamma_{\Lambda_1}$ and

$$\mathcal{H}_{\Lambda_1}(\mu|\nu) = \mathcal{H}_{\Lambda_1}(\mu|\nu\gamma_{\Lambda_1}) \geq \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Lambda_1}(\mu|\rho\gamma_{\Lambda_1}).$$

The goal is to obtain a similar lower bound for the integral in (2.4.2). Assume in the sequel that $\mathcal{H}_{\Lambda_1}(\mu|\nu)$ is finite; the lemma follows from (2.4.2) if it is not. This means in particular that $\mu_{\Lambda_1} \ll \nu_{\Lambda_1}$. Formally, μ^ζ and ν^ζ are probability kernels from $(E^{\Lambda_1}, \mathcal{E}^{\Lambda_1})$ to $(\Omega,\mathcal{F})$, which may be measured by μ_{Λ_1}. Moreover, these kernels satisfy $\pi_{\Lambda_1}(\mu^\zeta) = \pi_{\Lambda_1}(\nu^\zeta) = \delta_\zeta$. First we assert that

$$\int_{E^{\Lambda_1}} \mathcal{H}_{\Lambda_2}(\mu^\zeta|\nu^\zeta) d\mu_{\Lambda_1}(\zeta) = \mathcal{H}_\Lambda(\mu_{\Lambda_1}\mu^\zeta|\mu_{\Lambda_1}\nu^\zeta).$$

It is straightforward to see that this holds true: an expansion of the expression on the right in this display similar to the expansion in (2.4.2) yields the integral on the left plus the entropy term $\mathcal{H}(\mu_{\Lambda_1},\mu_{\Lambda_1}) = 0$. It is clear that $\mu_{\Lambda_1}\mu^\zeta = \mu$. For the other kernel, we observe that $\nu^\zeta = \nu^\zeta\gamma_{\Lambda_2}$ by consistency for ν_{Λ_1}-almost every ζ, and therefore also for μ_{Λ_1}-almost every ζ. In particular, this means that

$$\mathcal{H}_\Lambda(\mu_{\Lambda_1}\mu^\zeta|\mu_{\Lambda_1}\nu^\zeta) = \mathcal{H}_\Lambda(\mu|\mu_{\Lambda_1}\nu^\zeta\gamma_{\Lambda_2}) \geq \mathcal{H}_{\Lambda_2}(\mu|\mu_{\Lambda_1}\nu^\zeta\gamma_{\Lambda_2}) \geq \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Lambda_2}(\mu|\rho\gamma_{\Lambda_2}).$$

□

Lemma 2.4.3. *Let $(X, \mathcal{X})$ denote a measurable space, and consider $\mathcal{A} \subset \mathcal{M}(X, \mathcal{X})$ with $\operatorname{Diam}^\infty \mathcal{A}$ finite. Then for any finite measure $\mu \in \mathcal{M}(X, \mathcal{X})$ and for any $\nu, \nu' \in \mathcal{A}$, we have*

$$|\mathcal{H}(\mu|\nu) - \mathcal{H}(\mu|\nu')| \leq \mu(X) \operatorname{Diam}^\infty \mathcal{A},$$

where we interpret $|\infty - \infty|$ as 0.

Proof. Note that $\mu \ll \nu$ if and only if $\mu \ll \nu'$. Write $f := d\mu/d\nu$ and $f' := d\mu/d\nu'$. Then μ-almost everywhere $d\nu/d\nu' = f'/f$ and $|\log f' - \log f| \leq \operatorname{Diam}^\infty \mathcal{A}$. In particular,

$$|\mathcal{H}(\mu|\nu) - \mathcal{H}(\mu|\nu')| = |\mu(\log f) - \mu(\log f')| \leq \mu(|\log f - \log f'|) \leq \mu(X) \operatorname{Diam}^\infty \mathcal{A}.$$

□

Lemma 2.4.4. *The specific free energy functional $h(\cdot|\gamma) : \mathcal{P}_\Theta(\Omega, \mathcal{F}) \to [0, \infty]$ satisfies*

$$\begin{aligned} h(\mu|\gamma) &:= \lim_{n\to\infty} |\Delta_n|^{-1} \mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}) && (2.4.5)\\ &= \sup_{n\in\mathbb{N}} |\Delta_n|^{-1} (\mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}) - \operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma)) && (2.4.6)\\ &= \lim_{n\to\infty} |\Delta_n|^{-1} \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Delta_n}(\mu|\rho\gamma_{\Delta_n}) && (2.4.7)\\ &= \sup_{n\in\mathbb{N}} |\Delta_n|^{-1} \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Delta_n}(\mu|\rho\gamma_{\Delta_n}) && (2.4.8) \end{aligned}$$

for any weakly dependent specification γ and for any $\nu \in \mathcal{P}(\Omega, \mathcal{F})$.

Proof. Together, Lemma 2.4.1 of the current chapter and Lemma 15.11 of [20] assert that the sequence in (2.4.7) converges, with limit (2.4.8). Lemma 2.4.3 and weak dependence of γ imply that for any $\nu \in \mathcal{P}(\Omega, \mathcal{F})$,

$$\left|\mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n}) - \inf_{\rho\in\mathcal{P}(\Omega,\mathcal{F})} \mathcal{H}_{\Delta_n}(\mu|\rho\gamma_{\Delta_n})\right| \leq \operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma) = o(|\Delta_n|)$$

as $n \to \infty$. This means that (2.4.5) and (2.4.7) are the same. The inequality in the display implies that each term in the supremum in (2.4.6) is bounded from above by the corresponding term in (2.4.8), and therefore the supremum in (2.4.6) is bounded from above by the supremum in (2.4.8). However, the asymptotic bound on $\operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma)$ implies that the supremum in (2.4.6) equals at least the limit in (2.4.5). Conclude that (2.4.5), (2.4.6), (2.4.7) and (2.4.8) are all equal. □

Corollary 2.4.9. *We have $\mathcal{G}_\Theta(\gamma) \subset h_0(\gamma)$ whenever γ is weakly dependent.*

Proof. Consider $\mu \in \mathcal{G}_\Theta(\gamma)$, and apply the previous lemma with $\nu = \mu$. □

2.4.2 Minimisers and level sets

Lemma 2.4.10. *Let γ denote a weakly dependent specification. Then $\{h(\cdot|\gamma) \leq C\}$ is nonempty and compact in the $\mathcal{L}$-topology for any $C \in [0, \infty)$. In particular, $h_0(\gamma)$ is nonempty and compact in the $\mathcal{L}$-topology.*

Proof. The motivation for this lemma is standard; we include a proof for completeness. Fix a measure $\nu \in \mathcal{P}_\Theta(\Omega, \mathcal{F})$ and a constant $C \in [0, \infty)$. Level sets of relative entropy are compact: in our setting

$$\mathcal{P}_{n,C} := \{\mu \in \mathcal{P}(E^{\Delta_n}, \mathcal{E}^{\Delta_n}) : \mathcal{H}(\mu|\nu\hat{\gamma}_{\Delta_n}) \leq |\Delta_n| C + \operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma)\}$$

is compact in the strong topology on $\mathcal{P}(E^{\Delta_n}, \mathcal{E}^{\Delta_n})$ for any $n \in \mathbb{N}$. Equation 2.4.6 of Lemma 2.4.4 says that

$$\{h(\cdot|\gamma) \le C\} = \cap_{n\in\mathbb{N}}\{\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F}) : \mu_{\Delta_n} \in \mathcal{P}_{n,C}\}.$$

Let $(\mu^m)_{m\in\mathbb{N}} \subset \mathcal{P}(\Omega, \mathcal{F})$ denote a sequence of random fields—not necessarily shift-invariant—such that for any fixed $n \in \mathbb{N}$, we have $\pi_{\Delta_n}(\mu^m) \in \mathcal{P}_{n,C}$ for m sufficiently large. By compactness of each set $\mathcal{P}_{n,C}$, a standard diagonalisation argument, and the Kolmogorov extension theorem, we obtain a subsequential limit $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ of $(\mu^m)_{m\in\mathbb{N}}$ in the $\mathcal{L}$-topology with the property that $\mu_{\Delta_n} \in \mathcal{P}_{n,C}$ for each $n \in \mathbb{N}$.

For the lemma, it suffices to prove that $\{h(\cdot|\gamma) \le C\}$ is compact and that $h_0(\gamma)$ is nonempty. Start with the former. Suppose that $(\mu^m)_{m\in\mathbb{N}} \subset \{h(\cdot|\gamma) \le C\}$. Then $\pi_{\Delta_n}(\mu^m) \in \mathcal{P}_{n,C}$ for any $n, m \in \mathbb{N}$. Apply the previous argument to obtain a subsequential limit $\mu \in \mathcal{P}(\Omega, \mathcal{F})$. Then μ must be shift-invariant because each μ^m is shift-invariant. The argument says moreover that $\mu_{\Delta_n} \in \mathcal{P}_{n,C}$ for each $n \in \mathbb{N}$, that is, $h(\mu|\gamma) \le C$. This proves that the level set $\{h(\cdot|\gamma) \le C\}$ is compact. Finally, we prove that $h_0(\gamma)$ is nonempty. Set C to 0, and define

$$\mu^m := \frac{1}{|\Delta_m|} \sum_{x\in\Delta_m} \nu\gamma_{\Delta_{2m}+x} = \frac{1}{|\Delta_m|} \sum_{x\in\Delta_m} \theta_x \nu\gamma_{\Delta_{2m}}.$$

The two measures are equal because ν is shift-invariant, and it is clear that any subsequential limit of $(\mu^m)_{m\in\mathbb{N}}$ is also shift-invariant. Moreover, $\mu^m\gamma_{\Delta_n} = \mu^m$ whenever $m \ge n$ because $\Delta_n \subset \Delta_m \subset \Delta_{2m} + x$ for any $x \in \Delta_m$. This means that $\pi_{\Delta_n}(\mu^m) \in \mathcal{P}_{n,0}$ for m sufficiently large, for each fixed $n \in \mathbb{N}$. The sequence thus has a subsequential limit μ in the $\mathcal{L}$-topology. This limit μ must satisfy $\mu_{\Delta_n} \in \mathcal{P}_{n,0}$ for any n. Conclude that $\mu \in h_0(\gamma)$, that is, $h_0(\gamma)$ is nonempty. □

2.5 Minimisers of the specific free energy

2.5.1 Mazur's lemma

Lemma 2.5.1. *Let $(X, \mathcal{X})$ denote a standard Borel space and $\mathcal{A}$ a convex subset of $\mathcal{P}(X, \mathcal{X})$ subject to* $\mathrm{Diam}^\infty \mathcal{A}$ *being finite. Then the set*

$$\mathcal{C} := \mathcal{C}(\mathcal{A}) := \{\mu \in \mathcal{P}(X, \mathcal{X}) : \inf_{\nu\in\mathcal{A}} \mathcal{H}(\mu|\nu) = 0\}$$

is compact in the strong topology on $\mathcal{P}(X, \mathcal{X})$, satisfies $\mathrm{Diam}^\infty \mathcal{C} = \mathrm{Diam}^\infty \mathcal{A}$, *and equals*

1. *The closure of $\mathcal{A}$ in the total variation topology,*
2. *The closure of $\mathcal{A}$ in the strong topology.*

This lemma is close to trivial when E is finite, which is the case for many, but certainly not all, interesting models. It is this lemma that makes the theory work also for models where $(E, \mathcal{E})$ is a general standard Borel space. The lemma comes down to a straightforward application of Mazur's lemma: a well-known result from functional analysis. Remark that there is a confusing difference in naming conventions for topologies between functional analysis and measure theory, when the topologies are on sets of (probability) measures.

Proof of Lemma 2.5.1. Fix a measure $\lambda \in \mathcal{A}$; this measure will serve as reference measure. Write $f_\mu := d\mu/d\lambda$ for any σ-finite measure μ on $(X, \mathcal{X})$ that is absolutely continuous with respect to λ. For example, if $\mu \in \mathcal{A}$, then λ-almost everywhere $|\log f_\mu| \leq \operatorname{Diam}^\infty \mathcal{A}$. In particular, the map $\mu \mapsto f_\mu$ injects $\mathcal{A}$ into $L^1(\lambda)$—the image of $\mathcal{A}$ under this map is also convex. Write f^- for the lattice infimum of the family $\{f_\mu : \mu \in \mathcal{A}\}$; this is the largest $\mathcal{X}$-measurable function such that λ-almost everywhere $f^- \leq f_\mu$ for each $\mu \in \mathcal{A}$. See Lemma 2.6 in [27] for existence and uniqueness of $f^- \in L^1(\lambda)$. Similarly, write f^+ for the lattice supremum of $\{f_\mu : \mu \in \mathcal{A}\}$. Observe that λ-almost everywhere $f^- \leq 1 \leq f^+$ and

$$0 \leq \log \frac{f^+}{f^-} \leq \operatorname{Diam}^\infty \mathcal{A};$$

the former because $\lambda \in \mathcal{A}$, the latter follows from the definition of the diameter. In particular,

$$e^{-\operatorname{Diam}^\infty \mathcal{A}} \leq \operatorname{ess\,inf}_\lambda f^\pm \leq \operatorname{ess\,sup}_\lambda f^\pm \leq e^{\operatorname{Diam}^\infty \mathcal{A}}.$$

Define the measures $\lambda^\pm := f^\pm \lambda$—these should be considered the lattice infimum and supremum of the set $\mathcal{A}$, and are independent of the choice of reference measure $\lambda \in \mathcal{A}$. A measure $\mu \in \mathcal{P}(X, \mathcal{X})$ must satisfy $\lambda^- \leq \mu \leq \lambda^+$ if either $\mu \in \mathcal{C}$, or if μ is in the closure of $\mathcal{A}$ in the total variation topology, or if μ is in the closure of $\mathcal{A}$ in the strong topology. This also implies that λ-almost everywhere $f^- \leq f_\mu \leq f^+$.

We first show that $\operatorname{Diam}^\infty \mathcal{C} = \operatorname{Diam}^\infty \mathcal{A}$. The previous observation implies that

$$\operatorname{Diam}^\infty \mathcal{C} \leq \operatorname{Diam}^\infty \{\mu \in \mathcal{P}(X, \mathcal{X}) : \lambda^- \leq \mu \leq \lambda^+\} = \mathcal{H}^\infty(\lambda^+|\lambda^-) = \operatorname{Diam}^\infty \mathcal{A}.$$

Now $\mathcal{A} \subset \mathcal{C}$ and therefore $\operatorname{Diam}^\infty \mathcal{A} \leq \operatorname{Diam}^\infty \mathcal{C}$: we conclude that $\operatorname{Diam}^\infty \mathcal{C} = \operatorname{Diam}^\infty \mathcal{A}$.

Fix a probability measure μ subject to $\lambda^- \leq \mu \leq \lambda^+$; the goal is to show that $\mu \in \mathcal{C}$ if and only if μ is contained in the closure of $\mathcal{A}$ in the total variation topology. Fix a sequence $(\nu_n)_{n \in \mathbb{N}} \subset \mathcal{A}$. Observe that $d\mu/d\nu_n = f_\mu / f_{\nu_n}$, and that

$$\mathcal{H}(\mu|\nu_n) = \nu_n \left(\tfrac{f_\mu}{f_{\nu_n}} \log \tfrac{f_\mu}{f_{\nu_n}} \right) = \nu_n \left(\Xi \left(\tfrac{f_\mu}{f_{\nu_n}} \right) \right),$$

where $\Xi : (0, \infty) \to [0, \infty)$ is defined by $\Xi(x) := 1 - x + x \log x$. The function Ξ is convex and attains its minimum 0 at $x = 1$ only. We observe that, as $n \to \infty$,

$$\begin{aligned} \mathcal{H}(\mu|\nu_n) \to 0 &\iff \nu_n(\Xi(f_\mu/f_{\nu_n})) \to 0 \\ &\iff \lambda(\Xi(f_\mu/f_{\nu_n})) \to 0 && (2.5.2) \\ &\iff f_{\nu_n} \to f_\mu \qquad \text{in } L^1(\lambda) && (2.5.3) \\ &\iff \nu_n \to \mu \qquad \text{in total variation.} && (2.5.4) \end{aligned}$$

The equivalence in (2.5.2) is due to the fact that $e^{-\operatorname{Diam}^\infty \mathcal{A}} \lambda \leq \nu_n \leq e^{\operatorname{Diam}^\infty \mathcal{A}} \lambda$ for each $n \in \mathbb{N}$, and nonnegativity of Ξ. Equivalence in (2.5.3) is due to said properties of the function Ξ, and the fact that all functions f_μ and f_{ν_n} are uniformly bounded away from zero and infinity. Equivalence in (2.5.4) is straightforward as $\lambda(|f_{\nu_n} - f_\mu|)$ equals the total variation distance from ν_n to μ. We have now proven that $\mathcal{C}$ equals the closure of $\mathcal{A}$ in the total variation topology.

Claim that the closure of $\mathcal{A}$ in the total variation topology equals the closure of $\mathcal{A}$ in the strong topology. The map $\mu \mapsto f_\mu$ is a bijection from the closure of $\mathcal{A}$ in the total variation topology to the closure of $\{f_\mu : \mu \in \mathcal{A}\}$ in the norm topology on $L^1(\lambda)$.

The map $\mu \mapsto f_\mu$ is also a bijection from the closure of $\mathcal{A}$ in the strong topology to the closure of $\{f_\mu : \mu \in \mathcal{A}\}$ in the weak topology on $L^1(\lambda)$. The set $\{f_\mu : \mu \in \mathcal{A}\}$ is convex, and therefore Mazur's lemma asserts that the closure of $\{f_\mu : \mu \in \mathcal{A}\}$ in $L^1(\lambda)$ is the same for the norm topology and for the weak topology.

The set $\mathcal{C}$ is compact in the strong topology because it is closed in the strong topology and has finite max-diameter: it is a subset of the compact set $\{\mu \in \mathcal{P}(X, \mathcal{X}) : \mathcal{H}(\mu|\lambda) \leq \mathrm{Diam}^\infty \mathcal{C}\}$. □

Corollary 2.5.5. *Consider a weakly dependent specification γ, and a shift-invariant random field μ. Then $\mu \in h_0(\gamma)$ if and only if $\mu_{\Delta_n} \in \mathcal{C}(\mathcal{A}_{\Delta_n}(\gamma))$ for each $n \in \mathbb{N}$.*

Proof. This is due to (2.4.8) of Lemma 2.4.4 in combination with Lemma 2.5.1. □

2.5.2 Limits of finite-volume Gibbs measures

Lemma 2.5.6. *If γ is a weakly dependent specification, then $h_0(\gamma) = \mathcal{W}_\Theta(\gamma)$.*

Proof. If $\mu \in \mathcal{W}_\Theta(\gamma)$, then $\mu_{\Delta_n} \in \mathcal{C}(\mathcal{A}_{\Delta_n}(\gamma))$ by definition of $\mathcal{W}(\gamma)$, and therefore $\mu \in h_0(\gamma)$ by Corollary 2.5.5. Now consider $\mu \in h_0(\gamma)$. For the lemma, it suffices to prove that $\mu \in \mathcal{W}_\Theta(\gamma)$. Again, Corollary 2.5.5 says that $\mu_{\Delta_n} \in \mathcal{C}(\mathcal{A}_{\Delta_n}(\gamma))$ for each $n \in \mathbb{N}$. Write $d(\cdot, \cdot)$ for total variation distance. Lemma 2.5.1 implies that there exists a sequence of measures $(\nu^n)_{n\in\mathbb{N}} \subset \mathcal{P}(\Omega, \mathcal{F})$ such that

$$d(\mu_{\Delta_n}, \nu^n \hat{\gamma}_{\Delta_n}) \leq 1/n$$

for each $n \in \mathbb{N}$. Now for any $m \geq n$, we observe that

$$d(\mu_{\Delta_n}, \pi_{\Delta_n}(\nu^m \gamma_{\Delta_m})) \leq d(\mu_{\Delta_m}, \pi_{\Delta_m}(\nu^m \gamma_{\Delta_m})) = d(\mu_{\Delta_m}, \nu^m \hat{\gamma}_{\Delta_m}) \leq 1/m.$$

In particular, $\pi_{\Delta_n}(\nu^m \gamma_{\Delta_m})$ approaches μ_{Δ_n} in the total variation topology as $m \to \infty$, and therefore also in the strong topology. Conclude that $\nu^m \gamma_{\Delta_m} \to \mu$ in the topology of local convergence as $m \to \infty$. In other words, $\mu \in \mathcal{W}_\Theta(\gamma)$. □

2.5.3 Regular conditional probability distributions

Recall that μ_Λ^ω denotes the r.c.p.d. on $(E^\Lambda, \mathcal{E}^\Lambda)$ of μ corresponding to the projection map $\pi_{S\smallsetminus\Lambda} : \Omega \to E^{S\smallsetminus\Lambda}$, where $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ is an arbitrary random field, and $\Lambda \in \mathcal{S}$. Recall also that we use the notation $\mathcal{A}_{\Lambda,\omega}(\gamma)$ for the set

$$\mathcal{A}_{\Lambda,\omega}(\gamma) := \cap_{\Delta\in\mathcal{S}} \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma)) = \cap_{n\in\mathbb{N}} \mathcal{C}(\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma)).$$

Lemma 2.5.7. *Let γ be a weakly dependent specification, and fix a minimiser $\mu \in h_0(\gamma)$ and a finite set $\Lambda \in \mathcal{S}$. Then the r.c.p.d. of μ satisfies $\mu_\Lambda^\omega \in \mathcal{A}_{\Lambda,\omega}(\gamma)$ for μ-almost every ω.*

Proof. Fix an arbitrary set $\Delta \in \mathcal{S}$ that contains Λ. For the lemma it suffices to show that $\mu_\Lambda^\omega \in \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma))$ for μ-almost every ω. Write μ_n^ω for the r.c.p.d. of μ on $(E^\Lambda, \mathcal{E}^\Lambda)$ corresponding to the natural projection map $\pi_{\Delta_n\smallsetminus\Lambda} : \Omega \to E^{\Delta_n\smallsetminus\Lambda}$; we are only interested in n so large that $\Delta_n \supset \Delta$. For such n, we claim that

$$\mu_n^\omega \in \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma))$$

almost surely (that is: for μ-almost every ω). Equation 2.4.8 of Lemma 2.4.4 implies that

$$\inf_{\rho\in\mathcal{A}_{\Delta_n}(\gamma)} \mathcal{H}(\mu_{\Delta_n}|\rho) = 0.$$

This implies that

$$\inf_{\rho\in\mathcal{A}_{\Delta_n}(\gamma)}\left(\mathcal{H}(\mu_{\Delta_n\smallsetminus\Lambda}|\rho_{\Delta_n\smallsetminus\Lambda})+\int_{E^{\Delta_n\smallsetminus\Lambda}}\mathcal{H}(\mu_n^\omega|\rho^\omega)d\mu_{\Delta_n\smallsetminus\Lambda}(\omega)\right)=0,$$

where ρ^ω is the r.c.p.d. of ρ on $(E^\Lambda,\mathcal{E}^\Lambda)$ corresponding to the projection map $E^{\Delta_n}\to E^{\Delta_n\smallsetminus\Lambda}$. Remark that $\rho^\omega\in\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma)$ almost surely because $\rho\in\mathcal{A}_{\Delta_n}(\gamma)$ and by consistency of γ. This means that

$$\inf_{\rho^\omega\in\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma)}\mathcal{H}(\mu_n^\omega|\rho^\omega)=0,$$

and therefore $\mu_n^\omega\in\mathcal{C}(\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma))$, almost surely. But $\mathcal{C}(\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma))\subset\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma))$ because $\Delta\subset\Delta_n$, which proves the claim.

For any $A\in\mathcal{E}^\Lambda$, the bounded martingale convergence theorem says that almost surely

$$\mu_n^\omega(A)\to\mu_\Lambda^\omega(A).$$

The set $\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma))$ is compact in the strong topology, and therefore almost surely $\mu_n^\omega\to\mu_\Lambda^\omega\in\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}(\gamma))$. □

Corollary 2.5.8. *If γ is a weakly dependent specification and $\mu\in h_0(\gamma)$ satisfies $\mu(\Omega_\gamma)=1$, then μ is almost Gibbs.*

Proof. By the previous lemma, $\mu_\Lambda^\omega\in\mathcal{A}_{\Lambda,\omega}(\gamma)=\{\hat\gamma_\Lambda(\cdot,\omega)\}$ for μ-a.e. ω, proving that μ is a DLR state. □

Corollary 2.5.9. *Let γ denote a weakly dependent specification, and fix a measure $\lambda\in\mathcal{A}_{\{0\}}(\gamma)$. We pretend that λ is a probability measure on the state space $(E,\mathcal{E})$. Then there exists a constant $\varepsilon>0$ such that, for any minimiser $\mu\in h_0(\gamma)$ and for any $\Lambda\in\mathcal{S}$, we have $\mu_\Lambda^\omega\geq(\varepsilon\lambda)^\Lambda$ for μ-almost every ω. In other words, μ has finite energy.*

In particular, if E is finite and every state $e\in E$ has positive probability with respect to λ, then one may replace λ by the counting measure on E, which possibly has the effect of forcing us to take ε smaller. By doing so, we obtain the original finite energy formulation of Burton and Keane [4].

Proof of Corollary 2.5.9. Consider a weakly dependent specification γ, and fix a probability measure $\lambda\in\mathcal{A}_{\{0\}}(\gamma)$. The definition of a weakly dependent specification and Lemma 2.5.1 imply that $\operatorname{Diam}^\infty\mathcal{C}(\mathcal{A}_{\{0\}}(\gamma))$ is finite, and therefore there exists an $\varepsilon>0$ such that $\mu\geq\varepsilon\lambda$ for any $\mu\in\mathcal{C}(\mathcal{A}_{\{0\}}(\gamma))$. (In fact, it is easy to see that the choice $\varepsilon:=\exp-\operatorname{Diam}^\infty\mathcal{C}(\mathcal{A}_{\{0\}}(\gamma))$ suffices for this purpose.)

Claim that $\mu\geq(\varepsilon\lambda)^\Lambda$ for any $\mu\in\mathcal{A}_\Lambda(\gamma)$, for fixed $\Lambda\in\mathcal{S}$. Write $\mu=\nu\hat\gamma_\Lambda$ for some $\nu\in\mathcal{P}(\Omega,\mathcal{F})$. Without loss of generality, we suppose that $\nu=\nu\gamma_\Lambda$, so that $\mu=\nu_\Lambda$. We also have $\nu=\nu\prod_{x\in\Lambda}\gamma_{\{x\}}$. By induction,

$$\nu=\nu\prod\nolimits_{x\in\Lambda}\gamma_{\{x\}}\geq(\varepsilon\lambda)^\Lambda\times\nu_{S\smallsetminus\Lambda}.$$

This proves the claim. The claim also proves that $\mu\geq(\varepsilon\lambda)^\Lambda$ for any $\mu\in\mathcal{C}(\mathcal{A}_\Lambda(\gamma))$, which implies the corollary due to Lemma 2.5.7. □

2.5.4 Duality between random fields and specifications

Lemma 2.5.10. *Let γ denote a weakly dependent specification and ν a minimiser of γ. Then for any shift-invariant random field μ, we have*

$$h(\mu|\gamma) = h(\mu|\nu) := \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\Delta_n}(\mu|\nu).$$

Proof. We observe that $|\mathcal{H}_{\Delta_n}(\mu|\nu) - \mathcal{H}_{\Delta_n}(\mu|\nu\gamma_{\Delta_n})| \leq \mathrm{Diam}^\infty\,\mathcal{C}(\mathcal{A}_{\Delta_n}(\gamma)) = o(|\Delta_n|)$ as $n \to \infty$. □

Let us now investigate the relation between $\mathbb{S}$ and $\mathbb{F}$. Define the relation $\sim$ on $\mathbb{F}$ by declaring that $\mu \sim \nu$ whenever $\mu \in h_0(\nu)$.

Lemma 2.5.11. *The relation $\sim$ is an equivalence relation on $\mathbb{F}$ with $h_0(\mu)$ the equivalence class of $\mu \in \mathbb{F}$.*

Proof. Fix $\nu \in \mathbb{F}$. Clearly $\nu \sim \nu$, because $h(\nu|\nu) = 0$. It suffices to show that $h_0(\mu) = h_0(\nu)$ whenever $\mu \sim \nu$. Suppose that $\mu \sim \nu$. As $\nu \in \mathbb{F}$, there exists a specification $\gamma \in \mathbb{S}$ such that $\nu \in h_0(\gamma)$. The previous lemma implies that $h_0(\nu) = h_0(\gamma)$, that is, $\mu \in h_0(\gamma)$, and therefore also $h_0(\mu) = h_0(\gamma)$. This proves that $h_0(\mu) = h_0(\nu)$. □

This is sufficient for the conclusions that were drawn in Subsection 2.3.3.

2.6 Applications

Most of the classical results on the variational principle follows directly from our new setting. In this section we will give several examples of this fact. We derive new results for the Loop $O(n)$ model and for the Ising model in a random percolation environment, which is also called the Griffiths singularity random field.

2.6.1 Models with an absolutely summable interaction potential

In this subsection we show how to derive naturally from our work the variational principle for absolutely summable potential as described in [20] or [49]. The model of interest is described by a reference measure and a shift-invariant absolutely summable potential. Write λ for the *reference measure*, which is a probability measure on the state space $(E, \mathcal{E})$. This measure informs us of the most random distribution of the state of an isolated vertex in the absence of any interaction. Write $\Phi = (\Phi_A)_{A\in\mathcal{S}}$ for the interaction potential. The potential encodes the interactions that exist between the states at different sites. Formally, an *interaction potential* $\Phi = (\Phi_A)_{A\in\mathcal{S}}$ is a family of functions such that $\Phi_A : \Omega \to \mathbb{R} \cup \{\infty\}$ is $\mathcal{F}_A$-measurable. The potential Φ is called *shift-invariant* if $\Phi_{\theta A}(\omega) = \Phi_A(\theta\omega)$ for any $A \in \mathcal{S}$, $\theta \in \Theta$, $\omega \in \Omega$. The potential Φ is called *absolutely summable* if

$$\|\Phi\| := \sum_{A\in\mathcal{S},\, 0\in A} \|\Phi_A\|_\infty < \infty,$$

where $\|\cdot\|_\infty$ denotes the supremum norm. It is thus assumed that Φ is shift-invariant and absolutely summable.

The potential induces a Hamiltonian. For $\Lambda \in \mathcal{S}$ and $\Delta \subset \mathbb{Z}^d$, define

$$H_{\Lambda,\Delta} := \sum_{A\in\mathcal{S},\, A\cap\Lambda\neq\varnothing,\, A\subset\Delta} \Phi_A.$$

In particular, the *Hamiltonians* are the functions of the form $H_\Lambda := H_{\Lambda,S}$, where $\Lambda \in \mathcal{S}$. The reference measure λ and the potential Φ generate a *Gibbs specification* $\gamma = (\gamma_\Lambda)_{\Lambda\in\mathcal{S}}$ defined by

$$\gamma_\Lambda(A,\omega) := \frac{1}{Z_\Lambda^\omega}\int_{E^\Lambda} 1_A(\zeta\omega_{S\smallsetminus\Lambda})e^{-H_\Lambda(\zeta\omega_{S\smallsetminus\Lambda})}d\lambda^\Lambda(\zeta)$$

for any $\Lambda \in \mathcal{S}$, $\omega \in \Omega$, and $A \in \mathcal{F}$, where Z_Λ^ω is the normalising constant

$$Z_\Lambda^\omega := \int_{E^\Lambda} e^{-H_\Lambda(\zeta\omega_{S\smallsetminus\Lambda})}d\lambda^\Lambda(\zeta). \tag{2.6.1}$$

The Hamiltonian H_Λ is always bounded by $|\Lambda|\cdot\|\Phi\|$. Moreover, for absolutely summable potentials, the strength of the interaction decreases with the range. We aim to show two things: that the specification γ is weakly dependent, and that $\Omega_\gamma = \Omega$. In that case, Corollary 2.4.9 and Corollary 2.5.8 prove the variational principle, where all almost Gibbs measures are Gibbs measures. For the analysis it is convenient to define, for $\Lambda, \Delta \in \mathcal{S}$,

$$\varepsilon_{\Lambda,\Delta} := \sum_{A\in\mathcal{S},\, A\cap\Lambda\neq\varnothing,\, A\not\subset\Delta} \|\Phi_A\|_\infty.$$

Compare this to the definition of $H_{\Lambda,\Delta}$—the construction implies the inequality $\|H_\Lambda - H_{\Lambda,\Delta}\|_\infty \le \varepsilon_{\Lambda,\Delta}$. The constants $\varepsilon_{\Lambda,\Delta}$ contain precisely all the information that we need for proving weak dependence and that $\Omega_\gamma = \Omega$. To see this, we first prove the following lemma.

Lemma 2.6.2. *For any $\omega \in \Omega$ and $\Lambda, \Delta \in \mathcal{S}$, we have* $\mathrm{Diam}^\infty\, \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}) \le 4\varepsilon_{\Lambda,\Delta}$.

Proof. Fix $\omega', \omega'' \in \Omega$ such that $\omega_\Delta = \omega'_\Delta = \omega''_\Delta$. Choose $\zeta \in E^\Lambda$. Then we know that $H_{\Lambda,\Delta}(\zeta\omega'_{S\smallsetminus\Lambda}) = H_{\Lambda,\Delta}(\zeta\omega''_{S\smallsetminus\Lambda})$, and the triangular inequality implies that

$$\begin{aligned}&|H_\Lambda(\zeta\omega'_{S\smallsetminus\Lambda}) - H_\Lambda(\zeta\omega''_{S\smallsetminus\Lambda})| \le \\ &\qquad |H_\Lambda(\zeta\omega'_{S\smallsetminus\Lambda}) - H_{\Lambda,\Delta}(\zeta\omega'_{S\smallsetminus\Lambda})| + |H_\Lambda(\zeta\omega''_{S\smallsetminus\Lambda}) - H_{\Lambda,\Delta}(\zeta\omega''_{S\smallsetminus\Lambda})| \le 2\varepsilon_{\Lambda,\Delta}.\end{aligned}$$

This inequality and (2.6.1)—the definition of Z_Λ^ω—imply that

$$|\log Z_\Lambda^{\omega'} - \log Z_\Lambda^{\omega''}| \le 2\varepsilon_{\Lambda,\Delta}.$$

The definition of the specification implies that $\hat\gamma_\Lambda(\cdot,\omega) = \frac{1}{Z_\Lambda^\omega}e^{-H_\Lambda(\cdot\omega_{S\smallsetminus\Lambda})}\lambda^\Lambda$, and therefore we deduce that $\mathcal{H}^\infty(\hat\gamma_\Lambda(\cdot,\omega'),\hat\gamma_\Lambda(\cdot,\omega'')) \le 4\varepsilon_{\Lambda,\Delta}$ from the inequalities in the previous two displays. Conclude that

$$\begin{aligned}&\mathrm{Diam}^\infty\, \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\omega}) = \\ &\qquad \mathrm{Diam}^\infty\, \mathcal{A}_{\Lambda,\Delta,\omega} = \sup_{\omega',\omega''\in\Omega,\, \omega_\Delta=\omega'_\Delta=\omega''_\Delta} \mathcal{H}^\infty(\hat\gamma_\Lambda(\cdot,\omega'),\hat\gamma_\Lambda(\cdot,\omega'')) \le 4\varepsilon_{\Lambda,\Delta}.\end{aligned}$$

This is the desired inequality. □

We now simply employ the bound provided by the lemma, in order to arrive at the variational principle. To deduce the variational principle with Gibbs measures, we must prove that the specification γ is weakly dependent, and that $\Omega_\gamma = \Omega$. By the lemma, we know that

1. $\mathrm{Diam}^\infty\, \mathcal{A}_{\Delta_n}(\gamma) \le 4\varepsilon_{\Delta_n,\Delta_n}$,

2. $\mathrm{Diam}^\infty \mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma) \le 4\varepsilon_{\Lambda,\Delta_n}$ for any $\omega \in \Omega$.

To prove weak dependence, it is therefore sufficient to show that $\varepsilon_{\Delta_n,\Delta_n} = o(|\Delta_n|)$ as $n \to \infty$. Similarly, to prove that $\Omega_\gamma = \Omega$, it is sufficient to show that $\varepsilon_{\Lambda,\Delta_n} \to 0$ as $n \to \infty$ for any $\Lambda \in \mathcal{S}$, as this would imply that

$$\mathrm{Diam}^\infty \mathcal{A}_{\Lambda,\omega}(\gamma) \le \inf_{n\in\mathbb{N}} \mathrm{Diam}^\infty \mathcal{C}(\mathcal{A}_{\Lambda,\Delta_n,\omega}(\gamma)) = 0.$$

Start with the latter. It is immediate from the definition of $\varepsilon_{\Lambda,\Delta_n}$ that

$$\varepsilon_{\Lambda,\Delta_n} \le \sum_{x\in\Lambda} \varepsilon_{\{x\},\Delta_n} = \sum_{x\in\Lambda} \varepsilon_{\{0\},\Delta_n - x} \to 0$$

as $n \to \infty$, because $|\Lambda|$ and $||\Phi||$ are both finite. This proves that $\Omega_\gamma = \Omega$. For weak dependence, decompose

$$\begin{aligned}
\varepsilon_{\Delta_n,\Delta_n} &\le \sum\nolimits_{x\in\Delta_n} \varepsilon_{\{x\},\Delta_n} = \sum\nolimits_{x\in\Delta_n} \varepsilon_{\{0\},\Delta_n - x} \\
&= \sum\nolimits_{x\in\Delta_{n-\lfloor\log n\rfloor}} \varepsilon_{\{0\},\Delta_n - x} + \sum\nolimits_{x\in\Delta_n\smallsetminus\Delta_{n-\lfloor\log n\rfloor}} \varepsilon_{\{0\},\Delta_n - x} \\
&\le |\Delta_{n-\lfloor\log n\rfloor}| \cdot \varepsilon_{\{0\},\Delta_{\lfloor\log n\rfloor}} + |\Delta_n \smallsetminus \Delta_{n-\lfloor\log n\rfloor}| \cdot ||\Phi|| = o(|\Delta_n|)
\end{aligned}$$

as $n \to \infty$.

2.6.2 The random-cluster model

Let us introduce the random-cluster model. Fix an *edge-weight* $p \in (0,1)$ and a *cluster-weight* $q \in (0,\infty)$. The idea of the random-cluster model is to perform independent bond percolation (with parameter p) on (a subset of) the square lattice $\mathbb{Z}^d$, and subsequently weight each configuration by q raised to the number of percolation clusters in the resulting random graph. To cast the random-cluster model into the formalism of this chapter, we must first choose a suitable state space $(E,\mathcal{E})$ for the vertices $x \in \mathbb{Z}^d$, which allows us to encode for each edge if it is open or not. There exists a natural way to do this: with each vertex x we associate the d edges of the form $\{x, x+e_i\}$ with $1 \le i \le d$. The state space that we choose is

$$E = \{0,1\}^{\{1,\dots,d\}},$$

where for $\omega_x \in E$ the i-th coordinate is a 1 if the edge $\{x, x+e_i\}$ is open and 0 if it is closed. For $e \in E$ we define $|e| := |\{1 \le i \le d : e_i = 1\}|$, the number of open edges encoded in e. If $\omega \in E^\Lambda$ for some $\Lambda \in \mathcal{S}$, then write $\|\omega\| := \sum_{x\in\Lambda} |\omega_x|$. If $\omega \in \Omega$ and $\Lambda \in \mathcal{S}$, then define

$$C(\omega,\Lambda) := \begin{array}{l}\text{the number of open clusters of } \omega \text{ that intersect} \\ \Lambda \text{ or contain a vertex adjacent to } \Lambda.\end{array}$$

It is important to observe that

$$|C(\omega,\Lambda) - C(\zeta,\Lambda)| \le 2|\partial\Lambda| \tag{2.6.3}$$

if $\omega_\Lambda = \zeta_\Lambda$, where $\partial\Lambda$ denotes the *edge boundary* of Λ, that is, set of edges of the square lattice with exactly one endpoint in Λ. We now introduce the specification $\gamma = (\gamma_\Lambda)_{\Lambda\in\mathcal{S}}$ corresponding to the random-cluster model. For any $\omega \in \Omega$, $\Lambda \in \mathcal{S}$, and $\zeta \in E^\Lambda$, we define the *weight function*

$$w(\zeta,\omega,\Lambda) := p^{\|\zeta\|}(1-p)^{d|\Lambda| - \|\zeta\|} q^{C(\zeta\omega_{S\smallsetminus\Lambda},\Lambda)}.$$

The probability kernel $\hat{\gamma}_\Lambda$ corresponding to the random-cluster model is now defined by

$$\hat{\gamma}_\Lambda(\zeta, \omega) := \frac{1}{Z_\Lambda^\omega} w(\zeta, \omega, \Lambda),$$

where Z_Λ^ω is a suitable normalisation constant. The complete, nonrestricted probability kernel γ_Λ is given by $\gamma_\Lambda(\cdot, \omega) = \hat{\gamma}_\Lambda(\cdot, \omega) \times \delta_{\omega_{S \smallsetminus \Lambda}}$. Let us now prove that the resulting specification $\gamma = (\gamma_\Lambda)_{\Lambda \in \mathcal{S}}$ is weakly dependent. From (2.6.3) and the definition of w it is clear that

$$\left| \log \frac{w(\zeta, \omega, \Lambda)}{w(\zeta, \omega', \Lambda)} \right| \leq 2|\partial \Lambda| |\log q|$$

for all possible ζ, ω, ω', and Λ. As a direct consequence

$$\left| \log \frac{Z_\Lambda^\omega}{Z_\Lambda^{\omega'}} \right| \leq 2|\partial \Lambda| |\log q|, \quad \text{and} \quad \left| \log \frac{\hat{\gamma}_\Lambda(\zeta, \omega)}{\hat{\gamma}_\Lambda(\zeta, \omega')} \right| \leq 4|\partial \Lambda| |\log q|.$$

The right inequality implies that

$$\operatorname{Diam}^\infty \mathcal{A}_{\Delta_n}(\gamma) \leq 4|\partial \Delta_n| |\log q| = o(|\Delta_n|)$$

as $n \to \infty$, which proves that the specification γ is weakly dependent.

The goal is to prove the variational principle, which asserts the equivalence in Equation 2.3.1 for any shift-invariant random field μ. Weak dependence of γ gives us access to the framework that is developed in this chapter. In particular, we have the following three results:

1. There exists at least one shift-invariant measure μ such that $h(\mu|\gamma) = 0$,
2. If μ is a shift-invariant DLR state, then $\mu \in h_0(\gamma)$,
3. If $\mu \in h_0(\gamma)$ and $\mu(\Omega_\gamma) = 1$, then μ is almost Gibbs.

To arrive at the variational principle, it is now sufficient to prove that $\mu(\Omega_\gamma) = 1$ whenever $\mu \in h_0(\gamma)$.

Define

$$\Omega' := \left\{ \omega \in \Omega : \begin{array}{l} \text{if } \zeta \in \Omega \text{ is any other configuration that equals } \omega \text{ up to} \\ \text{finitely many edges, then } \zeta \text{ has at most one infinite com-} \\ \text{ponent} \end{array} \right\}.$$

It follows from the well-known argument of Burton and Keane [4] that $\mu(\Omega') = 1$ whenever μ is a shift-invariant random field with finite energy. Minimisers of the specific free energy have finite energy due to Corollary 2.5.9. Thus, in order to deduce the variational principle for the random-cluster model, it suffices to demonstrate that $\Omega' \subset \Omega_\gamma$.

Fix $\omega \in \Omega'$, and claim that $\omega \in \Omega_\gamma$. This is well-known for the random-cluster model, but perhaps not in the language of this chapter; we give a concise proof. Fix $\Lambda \in \mathcal{S}$. We make the stronger claim that

$$\mathcal{A}_{\Lambda, \Delta, \omega}(\gamma) = \{\hat{\gamma}_\Lambda(\cdot, \omega)\}$$

for Δ sufficiently large. In other words, we claim that for some appropriate choice of Δ, the measure $\hat{\gamma}_\Lambda(\cdot, \omega)$ is invariant under changing ω on the complement of Δ. The point is that the dependence of $\hat{\gamma}_\Lambda(\cdot, \omega)$ on ω is through the way that the percolation structure encoded in ω connects the vertices in the boundary of Λ with paths through the complement of Λ. Choose $\Delta \in \mathcal{S}$ such that

1. Δ contains Λ,
2. If x is adjacent to Λ and part of a finite ω-cluster, then Δ contains that entire finite ω-cluster and all vertices adjacent to it,
3. If x and y are adjacent to Λ and contained in the infinite ω-cluster, then Δ contains an open path from x to y through the complement of Λ.

The choice $\omega \in \Omega'$ guarantees that the open path from x to y through the complement of Λ exists. The merit of this choice of Δ is of course that

$$C(\xi, \Lambda) = C(\xi', \Lambda),$$

whenever $\xi, \xi' \in \Omega$ are chosen such that $\xi_\Delta = \xi'_\Delta$ and $\xi_{\Delta \smallsetminus \Lambda} = \xi'_{\Delta \smallsetminus \Lambda} = \omega_{\Delta \smallsetminus \Lambda}$. In particular, this implies that

$$w(\zeta, \omega, \Lambda) = w(\zeta, \omega', \Lambda)$$

for any $\zeta \in E^\Lambda$ and for any $\omega' \in \Omega$ such that $\omega'_\Delta = \omega_\Delta$. Conclude that $\hat{\gamma}_\Lambda(\cdot, \omega') = \hat{\gamma}_\Lambda(\cdot, \omega)$ for such $\omega' \in \Omega$, which implies the claim.

2.6.3 The Loop $O(n)$ model

The arguments for the variational principle for the random-cluster model work for any weakly dependent model in which the long-range interaction is due to weight on percolation clusters, level sets, paths, or other large geometrical objects which arise from the local structure (for the random-cluster model this was the cluster-weight q). The variational principle holds true for all such models. Consider, for example, the Loop $O(n)$ model. In this model, one draws disjoint loops on the hexagonal lattice; the probability of a certain configuration depends on the number of loops and on the number of loop edges in that configuration. It is thus a two-parameter model, much like the random-cluster model. See the work of Peled and Spinka [46] for a detailed introduction. The Loop $O(n)$ model may be formalised as follows: it is a model of random functions from the faces of the hexagonal lattice to $E = \{0, 1\}$. The number of level sets of these functions corresponds to the number of loops in the Loop $O(n)$ model, and the number of edges on which the function is not constant corresponds to the number of edges that are contained in a loop. Remark that in this case the Burton and Keane argument tells us that there is at most one infinite level set on which the function equals 0, and at most one infinite level set on which the function equals 1. If both infinite level sets are present, then they are clearly distinguished by their type.

2.6.4 The Griffiths singularity random field

The Griffiths singularity random field was introduced by Van Enter, Maes, Schonmann, and Shlosman [15]. They study the model in relation to the phenomenon of so-called *Griffiths singularities.* The model depends on two parameters: the *percolation parameter* $p \in (0, 1)$, and the *inverse temperature* $\beta \in \mathbb{R}$; both are fixed throughout the discussion. We take $\beta \geq 0$ without loss of generality, which corresponds to the ferromagnetic setting. To draw from the Griffiths singularity random field $K_{p,\beta}$, one first samples independent site percolation with parameter p; then, on each percolation cluster, one samples an independent Ising model with parameter β. The Griffiths singularity random field is thus an Ising model in a random environment.

First, we introduce some notation. A natural choice for the state space is $E = \{-1, 0, 1\}$. The state 0 indicates a closed vertex, while the state ± 1 indicates an open vertex of that spin. Write $\mathcal{E}$ for the powerset of E, a σ-algebra, and $\mathcal{E}_0$ for the σ-algebra on E generated by the function 1_0. Let $\mathcal{F}^0$ denote the product σ-algebra $\mathcal{E}_0^S$. If $\omega \in \Omega$ or $\omega \in E^\Lambda$ for some $\Lambda \subset S$, then write $\Pi(\omega) \subset \mathbb{Z}^d$ for the set of open vertices. We consider each configuration $\omega \in \Omega$ to be a function from $\mathbb{Z}^d$ to $\{-1, 0, 1\}$, and in that light we treat $|\omega|$, $-\omega$, and 1_Λ as configurations in Ω for any $\omega \in \Omega$ or $\Lambda \subset \mathbb{Z}^d$. There is a natural ordering $\leq$ on Ω; write $\omega^1 \leq \omega^2$ whenever $\omega^1_x \leq \omega^2_x$ for any $x \in \mathbb{Z}^d$. If $\mu_1, \mu_2 \in \mathcal{P}(\Omega, \mathcal{F})$, then write $\mu_1 \preceq \mu_2$ if μ_1 is stochastically dominated by μ_2, that is, if there exists a coupling between μ_1 and μ_2 such that $\omega^1 \leq \omega^2$ almost surely. Finally, the square lattice $\mathbb{Z}^d$ has naturally associated to it an edge set; write xy (juxtaposition) for an unordered pair of neighbouring vertices $x, y \in \mathbb{Z}^d$ in this graph. Write $\partial\Lambda$ for the edge boundary of any set $\Lambda \in \mathcal{S}$, as in the analysis of the random-cluster model.

The Ising model on a finite graph

For finite sets $\Lambda \in \mathcal{S}$, the *Ising model* in Λ is a probability measure on E^Λ defined by

$$\alpha_\Lambda(\omega) \propto \prod_{xy \subset \Lambda} e^{-\beta\omega_x\omega_y}$$

if $\omega_x = \pm 1$ for every $x \in \Lambda$, and $\alpha_\Lambda(\omega) = 0$ otherwise. The following key identity is a corollary of the definition:

$$\alpha_\Lambda(\omega) = \frac{1}{Z} \cdot f_{\Lambda,\Delta}(\omega) \cdot \alpha_{\Lambda\cap\Delta}(\omega_{\Lambda\cap\Delta}) \cdot \alpha_{\Lambda\smallsetminus\Delta}(\omega_{\Lambda\smallsetminus\Delta}) \tag{2.6.4}$$

for any $\Lambda, \Delta \in \mathcal{S}$ and $\omega \in E^\Lambda$, where

$$f_{\Lambda,\Delta}(\omega) := \prod\nolimits_{xy\subset\Lambda,\, xy\in\partial\Delta} e^{-\beta\omega_x\omega_y} \qquad \text{and} \qquad Z = \int_{E^\Lambda} f_{\Lambda,\Delta} d(\alpha_{\Lambda\cap\Delta} \times \alpha_{\Lambda\smallsetminus\Delta}).$$

In particular, if $\Lambda \in \mathcal{S}$ and Δ a connected component of Λ, then (2.6.4) implies that $\alpha_\Lambda = \alpha_\Delta \times \alpha_{\Lambda\smallsetminus\Delta}$.

If $\Lambda \in \mathcal{S}$ and $\omega \in E^\Delta$ for some $\Lambda \subset \Delta \subset S$, then we sometimes write $\alpha_\Lambda(\omega)$ for $\alpha_\Lambda(\omega_\Lambda)$.

The Ising model on an infinite graph

The Ising model on infinite subgraphs of $\mathbb{Z}^d$ is introduced in terms of the associated specification, which is denoted by $\kappa = (\kappa_\Lambda)_{\Lambda\in\mathcal{S}}$. Consider arbitrary $\Lambda \in \mathcal{S}$ and $\omega \in \Omega$. Informally, the measure $\kappa_\Lambda(\cdot, \omega) \in \mathcal{P}(\Omega, \mathcal{F})$ is the Ising model in the graph $\Pi(\omega) \cap \Lambda$—the edges inherited from the square lattice—subject to boundary conditions provided by the configuration ω. Formally, $\kappa_\Lambda(\cdot, \omega)$ is the unique random field such that

$$\kappa_\Lambda(\zeta, \omega) \propto \prod_{xy \,\subset\, \Lambda \text{ or } xy \,\in\, \partial\Lambda} e^{-\beta\zeta_x\zeta_y}$$

for any $\zeta \in \Omega$ such that $\zeta_{S\smallsetminus\Lambda} = \omega_{S\smallsetminus\Lambda}$ and $\Pi(\zeta) = \Pi(\omega)$, and $\kappa_\Lambda(\zeta, \omega) = 0$ for all other ζ. Of course, the only edges xy that contribute to the product in the display are the ones that are also contained in $\Pi(\zeta) = \Pi(\omega)$. As per usual, we abbreviate $\hat\kappa_\Lambda(\cdot, \omega) := \pi_\Lambda(\kappa_\Lambda(\cdot, \omega))$, and we observe that $\alpha_\Lambda = \hat\kappa_\Lambda(\cdot, 1_\Lambda)$ in this notation.

The interest is however in the Ising model in the entire graph induced by $\Pi(\omega)$. By monotonicity, the sequence of random fields $(\kappa_{\Delta_n}(\cdot,|\omega|))_{n\in\mathbb{N}}$ is decreasing with respect to $\preceq$, and therefore tends to a limit in the $\mathcal{L}$-topology as $n\to\infty$. Write $\kappa^+(\cdot,\omega)$ for this limit, and similarly write $\kappa^-(\cdot,\omega)$ for the limit of the increasing sequence $(\kappa_{\Delta_n}(\cdot,-|\omega|))_{n\in\mathbb{N}}$. Remark that both $\kappa^-(\cdot,\omega)$ and $\kappa^+(\cdot,\omega)$ depend on the percolation structure $\Pi(\omega)$ of ω only, and not on the spins of the open sites. In other words, κ^+ and κ^- are probability kernels from $(\Omega,\mathcal{F}^0)$ to $(\Omega,\mathcal{F})$. A monotonicity argument implies that $\kappa^-(\cdot,\omega)\preceq\kappa^+(\cdot,\omega)$. If the two measures are distinct, then it is said that the Ising model *magnetises* on $\Pi(\omega)$. Write $M\subset\Omega$ for the collection of configurations ω such that the Ising model magnetises on $\Pi(\omega)$. The set M is measurable with respect to $\mathcal{F}^0$. It is also measurable with respect to $\mathcal{F}^0_{S\smallsetminus\Lambda}$, for any $\Lambda\in\mathcal{S}$. In other words, M is tail measurable. If $\zeta\in\Omega-M$, then another monotonicity argument implies that $\kappa^+(\cdot,\zeta)$ is the unique random field such that almost surely $\Pi(\omega)=\Pi(\zeta)$ and which is invariant under each probability kernel κ_Λ. We finally state an important proposition, which also follows from monotonicity.

Proposition 2.6.5. *The map $\omega\mapsto\kappa^+(\cdot,\omega)$ is continuous—both sides endowed with the $\mathcal{L}$-topology—at some $\zeta\in\Omega$ if and only if $\zeta\notin M$.*

The random percolation environment

Write P_p for the percolation measure with parameter p, that is, the measure in which each vertex takes value 1 with probability p, and value 0 with probability $1-p$, independently of all other vertices. Note that we have a zero-one law for the tail-measurable event M in P_p. We therefore distinguish three phases at most: one phase of subcritical percolation, one phase of supercritical percolation but with $P_p(M)=0$, and one phase of supercritical percolation with $P_p(M)=1$. Clearly $P_p(M)=0$ in the subcritical percolation regime as there are no infinite clusters almost surely and therefore the infinite Ising model decomposes into the product of infinitely many finite cluster Ising models. The interesting regime is therefore the supercritical percolation regime. Our goal is to prove the variational principle for the nonmagnetic phase—both in the subcritical and supercritical percolation regime.

Below critical percolation

Let us for now assume that we are in the subcritical percolation regime $p<p_c$, so that we avoid the presence of an infinite percolation cluster altogether. The Griffiths singularity random field $K_{p,\beta}$ is simply defined by the equation $K_{p,\beta}:=P_p\kappa^+$. To sample from $K_{p,\beta}$, one first samples the percolation structure ζ from P_p, then one draws the final sample ω from the Ising model $\kappa^+(\cdot,\zeta)$, which decomposes into a product of Ising models on the finite clusters of $\Pi(\zeta)$ almost surely.

Fix $\Lambda\in\mathcal{S}$. Observe that $K_{p,\beta}$ is invariant under the kernel which first resamples the percolation structure on Λ, then resamples the Ising model on each percolation cluster that intersects Λ. This motivates the definition of a natural specification associated to $K_{p,\beta}$. First, consider those $\omega\in\Omega$ for which there is no infinite percolation cluster. For any $\Lambda\in\mathcal{S}$, write $\Gamma(\omega,\Lambda)\subset\mathbb{Z}^d$ for the union of ω-open clusters that contain a vertex that is in or adjacent to Λ. Also write $\|\omega_\Lambda\|$ for the number of ω-open vertices in Λ. For such ω and Λ, we define the probability measure $\hat\gamma_\Lambda(\cdot,\omega)$ by

$$\hat\gamma_\Lambda(\zeta,\omega):=\frac{1}{Z^\omega_\Lambda}p^{\|\zeta\|}(1-p)^{|\Lambda|-\|\zeta\|}\alpha_{\Gamma(\zeta\omega_{S\smallsetminus\Lambda},\Lambda)}(\zeta\omega_{S\smallsetminus\Lambda}),\tag{2.6.6}$$

where Z^ω_Λ is a suitable normalisation constant, and ζ ranges over E^Λ. As per usual, the full kernel γ_Λ is recovered through the equation $\gamma_\Lambda(\cdot,\omega) = \hat\gamma_\Lambda(\cdot,\omega)\times\delta_{\omega_{S\smallsetminus\Lambda}}$. It follows from this definition and the intuitive picture that $K_{p,\beta} = K_{p,\beta}\gamma_\Lambda$ for every $\Lambda\in\mathcal{S}$, even though we have not yet defined $\gamma_\Lambda(\cdot,\omega)$ for those ω with an infinite percolation cluster.

Let us now rewrite the previous definition of $\hat\gamma_\Lambda(\cdot,\omega)$ into an expression that is less intuitive but more useful for the analysis. First, write $\xi := \zeta\omega_{S\smallsetminus\Lambda}$ and $\Gamma := \Gamma(\xi,\Lambda)$. Use (2.6.4) to obtain

$$\alpha_\Gamma(\xi) = \frac{f_{\Gamma,\Lambda}(\xi)\cdot\alpha_{\Gamma\cap\Lambda}(\zeta)\cdot\alpha_{\Gamma\smallsetminus\Lambda}(\omega)}{(\alpha_{\Gamma\cap\Lambda}\times\alpha_{\Gamma\smallsetminus\Lambda})(f_{\Gamma,\Lambda})}$$

Note that $\Gamma\cap\Lambda = \Pi(\zeta)$. The set $\Gamma(\zeta\omega_{S\smallsetminus\Lambda},\Lambda)-\Lambda$ depends on $\omega_{S\smallsetminus\Lambda}$ only, and therefore $\alpha_{\Gamma\smallsetminus\Lambda}(\omega)$ is independent of ζ. We may therefore combine $\alpha_{\Gamma\smallsetminus\Lambda}(\omega)$ with the normalisation constant in (2.6.6) to obtain

$$\hat\gamma_\Lambda(\zeta,\omega) = \frac{1}{Z^\omega_\Lambda}p^{\|\zeta\|}(1-p)^{|\Lambda|-\|\zeta\|}\alpha_{\Pi(\zeta)}(\zeta)\frac{f_{\Gamma,\Lambda}(\xi)}{(\alpha_{\Pi(\zeta)}\times\alpha_{\Gamma\smallsetminus\Lambda})(f_{\Gamma,\Lambda})};$$

now with a different normalisation constant. If we write f_Λ for the function

$$f_\Lambda(\omega) := \prod_{xy\in\partial\Lambda} e^{-\beta\omega_x\omega_y},$$

then the previous equation simplifies to

$$\hat\gamma_\Lambda(\zeta,\omega) = \frac{1}{Z^\omega_\Lambda}p^{\|\zeta\|}(1-p)^{|\Lambda|-\|\zeta\|}\alpha_{\Pi(\zeta)}(\zeta)\frac{f_\Lambda(\xi)}{(\hat\kappa_\Lambda(\cdot,1_{\Pi(\zeta)})\times\pi_{S\smallsetminus\Lambda}(\kappa^+(\cdot,1_{\Pi(\omega)-\Lambda})))(f_\Lambda)}. \tag{2.6.7}$$

This probability kernel is well-defined for any ω, even if ω has infinite clusters or if the Ising model magnetises on $\Pi(\omega)$. We shall take (2.6.7) as a definition for each kernel γ_Λ. The family $\gamma = (\gamma_\Lambda)_{\Lambda\in\mathcal{S}}$ so produced is a specification. The long-range interaction derives exclusively from the appearance of the measure $\kappa^+(\cdot,1_{\Pi(\omega)-\Lambda})$ in the denominator in the fraction on the right in (2.6.7). Recall that M is tail measurable: the Ising model magnetises on $\Pi(\omega)$ if and only if the Ising model magnetises on $\Pi(\omega)-\Lambda$. This leads to the following crucial observation.

Proposition 2.6.8. *Consider $\zeta\in\Omega$. If $\zeta\notin M$, then the map $\omega\mapsto\gamma_\Lambda(\cdot,\omega)$ is continuous—both sides endowed with the $\mathcal{L}$-topology—at ζ for any $\Lambda\in\mathcal{S}$. In other words, Ω_γ contains $\Omega - M$.*

We claim that the specification γ is weakly dependent. The reasoning is similar to the discussion of the random-cluster model. The dependence on ω in (2.6.7) is only through its appearance in the fraction on the right, and its effect on the normalisation constant Z^ω_Λ. But the definition of f_Λ implies that $|\log f_\Lambda|\leq|\partial\Lambda||\beta|$. The logarithm of the fraction in (2.6.7) is therefore bounded by $2|\partial\Lambda||\beta|$. Much like for the random-cluster model, this implies that

$$\left|\log\frac{Z^\omega_\Lambda}{Z^{\omega'}_\Lambda}\right|\leq 4|\partial\Lambda||\beta|\quad\text{and}\quad\left|\log\frac{\hat\gamma_\Lambda(\zeta,\omega)}{\hat\gamma_\Lambda(\zeta,\omega')}\right|\leq 8|\partial\Lambda||\beta|,$$

and we conclude with the asymptotic bound

$$\operatorname{Diam}^\infty\mathcal{A}_{\Delta_n}(\gamma)\leq 8|\partial\Delta_n||\beta| = o(|\Delta_n|)$$

as $n\to\infty$; the specification γ is weakly dependent. Note that the argument for weak dependence of γ works for any choice of parameters $p\in(0,1)$ and $\beta\geq 0$, regardless of the phase that we work in.

Below magnetisation

For the remainder of the theory, it is no longer necessary to require $p < p_c$. Instead, we fix the percolation parameter p and inverse temperature β subject only to $P_p(M) = 0$. Of course, the Griffiths singularity random field $K_{p,\beta}$ is defined by the equation $K_{p,\beta} := P_p\kappa^+ = P_p\kappa^-$. This measure is a DLR state of the specification γ as defined in (2.6.7). Moreover, we observe that $K_{p,\beta}(M) = P_p(M) = 0$, and therefore $K_{p,\beta}$ is supported on Ω_γ. In other words, $K_{p,\beta}$ is almost Gibbs with respect to γ. Our final goal is to prove the following theorem.

Theorem 2.6.9. *If the parameters p and β are such that $P_p(M) = 0$, then $h_0(\gamma) = \{K_{p,\beta}\}$.*

This statement is stronger than the variational principle, it also implies that $K_{p,\beta}$ is the unique DLR state of γ, and that $K_{p,\beta}$ is the unique minimiser of γ.

Proof of Theorem 2.6.9. Fix $\mu \in h_0(\gamma)$. Then $\mu \in h_0(K_{p,\beta})$. Remark that $K_{p,\beta}|_{\mathcal{F}^0} = P_p|_{\mathcal{F}^0}$; sampling the Ising model on the percolation clusters alters the spins on those clusters, but not the percolation structure itself. Observe that

$$\begin{aligned} h(\mu|K_{p,\beta}) &= \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\mathcal{E}^{\Delta_n}}(\mu_{\Delta_n}|\pi_{\Delta_n}(K_{p,\beta})) \\ &\geq \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\mathcal{E}_0^{\Delta_n}}(\mu_{\Delta_n}|\pi_{\Delta_n}(K_{p,\beta})) = \lim_{n\to\infty} |\Delta_n|^{-1}\mathcal{H}_{\mathcal{E}_0^{\Delta_n}}(\mu_{\Delta_n}|\pi_{\Delta_n}(P_p)). \end{aligned}$$

What we read on the last line in this display is exactly the SFE of $\mu|_{\mathcal{F}^0}$ with respect to $P_p|_{\mathcal{F}^0}$. But $P_p|_{\mathcal{F}^0}$ is a Gibbs measure with respect to an independent specification, which has a unique minimiser. We chose μ such that $h(\mu|K_{p,\beta}) = 0$, which now implies that $\mu|_{\mathcal{F}^0} = P_p|_{\mathcal{F}^0}$. We observe in particular that $\mu(M) = 0$, and consequently $\mu(\Omega_\gamma) = 1$. Therefore μ is almost Gibbs with respect to γ. Finally, we observe that $\gamma_\Lambda = \gamma_\Lambda\kappa_\Lambda$. This implies that μ is also a DLR state of the specification κ. But the Ising model is nonmagnetising on $\Pi(\omega)$ for μ-almost every ω, and therefore μ is also invariant under the probability kernel κ^+. This kernel is $\mathcal{F}^0$-measurable; conclude that $\mu = (\mu|_{\mathcal{F}^0})\kappa^+ = P_p\kappa^+ = K_{p,\beta}$. $\square$

Chapter 3

Macroscopic behaviour of Lipschitz random surfaces

The motivation for this chapter is to derive strict convexity of the surface tension for Lipschitz random surfaces, that is, for models of random Lipschitz functions from $\mathbb{Z}^d$ to $\mathbb{Z}$ or $\mathbb{R}$. An essential innovation is that random surface models with long- and infinite-range interactions are included in the analysis. More specifically, we cover at least: uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree for any $k \geq 2$ and Lipschitz potentials which satisfy the FKG lattice condition. The latter includes perturbations of dimer- and six-vertex models and of Lipschitz simply attractive potentials introduced by Sheffield. The main result is that we prove strict convexity of the surface tension—which implies uniqueness for the limiting macroscopic profile—if the model of interest is monotone in the boundary conditions. This solves a conjecture of Menz and Tassy, and answers a question posed by Sheffield. Auxiliary to this, we prove several results which may be of independent interest, and which do not rely on the model being monotone. This includes existence and topological properties of the specific free energy, as well as a characterisation of its minimisers. We also prove a general large deviations principle which describes both the macroscopic profile and the local statistics of the height functions. This work is inspired by, but independent of, *Random Surfaces* by Sheffield.

3.1 Introduction

3.1.1 Preface

We study the macroscopic behaviour of models of *Lipschitz random surfaces*, that is, random Lipschitz functions from $\mathbb{Z}^d$ to $\mathbb{Z}$ or $\mathbb{R}$. Examples of such models include height functions of dimer models and six-vertex models and uniformly random K-Lipschitz functions. One studies in particular the local Gibbs measures, subject to boundary conditions. It is generally expected that the macroscopic limit of a random surface under the influence of boundary conditions is governed by a *variational principle.* This variational principle asserts that, under suitable boundary conditions on a bounded domain $D \subset \mathbb{R}^d$, the asymptotic macroscopic profile f^* must concentrate on any neighbourhood of the set of minimisers of the integral

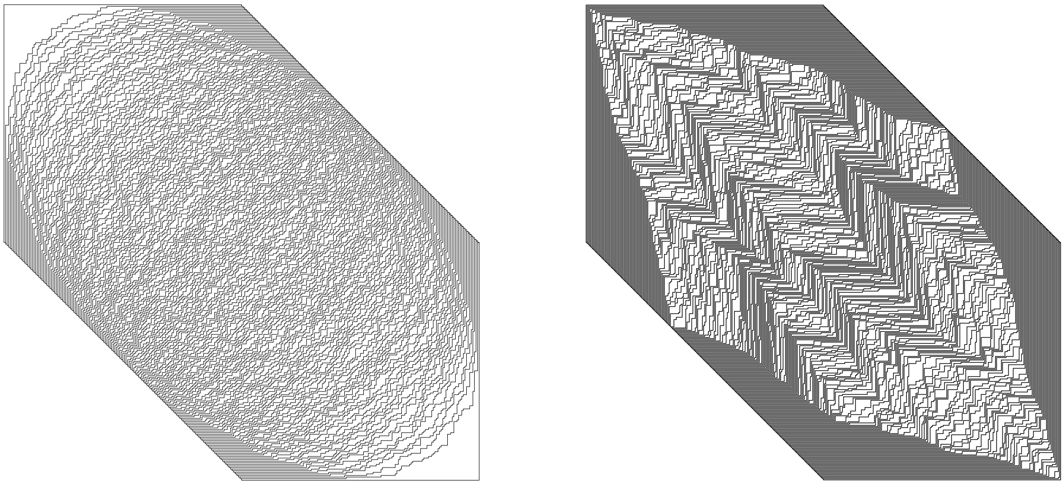
$$\int_D \sigma(\nabla f(x))dx \tag{3.1.1}$$

over all those functions f that match these boundary conditions.

The convex function σ, which is called the *surface tension*, is specific to the model and encodes the free energy density of gradient Gibbs measures which are constrained to a certain slope. Sheffield proves in his seminal work *Random Surfaces* [54] that this variational principle can be generalised into a large deviation principle that governs not only the macroscopic profile, but also the local statistics of a random surface over macroscopic regions. These results apply to a significant number of models. The fundamental integral in (3.1.1) connects the large deviations principle and the variational principle: it appears as the rate function in the large deviations principle, which implies the asserted concentration. When σ is strictly convex, the rate function of the large deviations principle has a unique minimiser f^* and the random functions concentrate around this unique minimiser (see [9] for a proof that strict convexity of σ implies uniqueness of the minimiser of the integral). This also implies that the model is stable under microscopic changes in the boundary conditions. On the other hand, when σ fails to be strictly convex, simulations have suggested that microscopic changes to boundary conditions might have macroscopic effects, and (more generally) that random surfaces might be macroscopically disordered. To illustrate this point, we refer to Figure 3.1 for two samples from the five-vertex model, one with parameters which make σ strictly convex, and one with parameters for which σ is not strictly convex. The difference in the macroscopic appearance of these two figures is striking. This dichotomy underlines the pivotal role played by the surface tension in the study of the asymptotic behaviour of random surfaces.

In the last thirty years, there have been various models in statistical physics for which strict convexity of the surface tension has been derived. The two most famous are probably the dimer model [5] for $\mathbb{Z}$-valued random surfaces and the Ginzburg-Landau $\nabla\phi$-interface under suitable conditions [19, 10] for $\mathbb{R}$-valued random surfaces. In either case, the strategy employed to demonstrate strict convexity of the surface tension relies heavily on particular properties of the model under consideration. For dimer models, one is able to calculate σ due to exact integrability of the model [5]; for the Ginzburg-Landau $\nabla\phi$-interface, the strategy relies on the fact that the potentials considered are almost Gaussian [19]. A decisive breakthrough was made in [54] in the pursuit of a more general approach. In this work, Sheffield proves that statistical physics models associated with simply attractive potentials—that is, convex potentials for which the interactions are exclusively between pairs of points—must have a strictly convex surface tension. Beyond the surprising generality of the result, this work also distinguishes itself by the method that was used to prove strict convexity of the surface tension. Rather than using direct computational arguments, the author reasons by contradiction: if there is a line segment on which the surface tension is affine, then the minimising measures corresponding to either endpoint are used to construct a new measure which minimises the specific free energy, but is not a Gibbs measure. This is then shown to be impossible.

Despite this significant progress, the techniques used in [54] rely heavily on the interactions being between pairs of points only—they cannot capture what happens for models with interactions involving larger clusters of points. The purpose of this chapter is to dramatically increase the class of models for which strict convexity of the surface tension can be derived. We do so by providing a new approach which does not rely on a particular formalism of the model in terms of a potential, but instead on *stochastic monotonicity*. Notably, the new class includes all Lipschitz models for which the interaction potential satisfies the Fortuin-Kasteleyn-Ginibre (FKG) lattice condition. Such potentials are also called *submodular*, and form a natural



Monotone parameters　　　　Non-monotone parameters

Figure 3.1: Limiting behaviour of the five-vertex model for different parameters

generalisation of the class of simply attractive potentials. Moreover, the new class also covers interaction potentials which assign a weight to each level set of the height function, in the spirit of the random-cluster model. Such models have infinite-range interactions, and we use them to derive strict convexity of the surface tension for the tree-valued graph homomorphisms studied in [44].

There are several ideas which suggest that stochastic monotonicity is a suitable starting point for studying the macroscopic behaviour of random surfaces. First, for general percolation models, such as independent percolation and Fortuin-Kasteleyn percolation, the FKG inequality is essential to the understanding of the macroscopic behaviour of the model: most, if not all, modern techniques in percolation theory rely on this crucial observation. It appears that stochastic monotonicity is the most general equivalent of the FKG inequality in the context of random height functions. Second, when the height functions of interest are also Lipschitz, the Azuma-Hoeffding inequality implies immediately that the random surface concentrates in some precise sense; the picture on the right in Figure 3.1 is therefore instantaneously ruled out. Third, it turns out that for this five-vertex model, stochastic monotonicity (which depends on the choice of parameters), is in fact *equivalent* to strict convexity of σ.

Finally, stochastic monotonicity does not depend on any formalism of potentials. This is a significant difference with the class of simply attractive models in [54], which depends on a particular representation of the model in terms of an underlying interaction potential. Stochastic monotonicity is thus practical: it suffices to check the Holley criterion. For discrete finite-range models, this is particularly efficient, as it amounts to evaluating a finite number of cases.

3.1.2 Description of the main results

Let us now broadly describe the main results of this chapter. Precise statements of the corresponding theorems are to be found in Section 3.4. Write Ω for the set of *height functions*, that is, functions ϕ from $\mathbb{Z}^d$ to E, where the choice of d and $E \in \{\mathbb{Z}, \mathbb{R}\}$ depends on the model of interest. Write $\Lambda \subset\subset \mathbb{Z}^d$ if the former is a finite subset of the latter; the model of interest is formalised in terms of a specification $\gamma = (\gamma_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^d}$ which allows one to forget about the values of ϕ on Λ and resample those values according to the model. The measure $\gamma_\Lambda(\cdot,\phi)$ is also called the *local Gibbs measure* in Λ with boundary conditions ϕ. This model must be invariant by some full-rank sublattice $\mathcal{L}$ of $\mathbb{Z}^d$ if any convergent macroscopic behaviour is to be expected. We impose two key restrictions on γ for the main results to apply: that $\gamma_\Lambda(\cdot,\phi)$ is supported on height functions which are suitably Lipschitz whenever ϕ is Lipschitz, and that $\gamma_\Lambda(\cdot,\phi) \preceq \gamma_\Lambda(\cdot,\psi)$ whenever $\phi \leq \psi$. Models satisfying the former condition are called *Lipschitz*, if they satisfy the latter then they are called *stochastically monotone*. Finally, for the thermodynamical formalism, we require that the specification γ is generated by some *interaction potential* Φ which encodes the interactions of the values of ϕ at different vertices. We shall see that the heart of the proof does not rely on the formalism of potentials as it is expressed directly in terms of the specification. As a consequence, we are able to incorporate potentials Φ belonging to a very large class which is described in detail in Section 3.3. Informally, we allow any potential Φ which decomposes as the sum of two potentials Ψ and Ξ, where Ψ is a potential of finite range which enforces the Lipschitz property (by assigning infinite potential to functions which are not Lipschitz), and where Ξ is potentially an infinite-range potential whose intensity decays fast enough for the specific free energy to be well-defined.

While the finite-range part Ψ of the potential encompasses all common finite-range models in statistical physics, the infinite-range part Ξ is tailored to fit long-range interaction potentials such as those associated with the random-cluster model or the Loop $O(n)$ model. We demonstrate in Subsection 3.13.3 that this formalism can even be used to prove a conjecture on the limiting behaviour of uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree for $k \geq 2$.

Let us now introduce a few notions before describing the main results. Write $\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ for the collection of $\mathcal{L}$-invariant gradient measures on Ω. Any measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ has an associated *slope* $S(\mu)$ which is the unique linear functional $u \in (\mathbb{R}^d)^*$ such that

$$u(x) = \mu(\phi(x) - \phi(0))$$

for all $x \in \mathcal{L}$. The *specific free energy* of μ is defined by the limit

$$\mathcal{H}(\mu|\Phi) := \lim_{n\to\infty} n^{-d}\mathcal{H}_{\Pi_n}(\mu|\Phi),$$

where $\Pi_n \subset\subset \mathbb{Z}^d$ denotes a box of sides n, and where $\mathcal{H}_\Lambda(\mu|\Phi)$ denotes the *free energy* of μ over Λ with respect to the interior Hamiltonian generated by Φ; this quantity is introduced formally in Section 3.2. The *surface tension* is the function $\sigma : (\mathbb{R}^d)^* \to \mathbb{R} \cup \{\infty\}$ defined by

$$\sigma(u) := \inf_{\mu \,\in\, \mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla) \text{ with } S(\mu)\,=\,u} \mathcal{H}(\mu|\Phi).$$

This function is automatically convex as $S(\cdot)$ and $\mathcal{H}(\cdot|\Phi)$ are affine over $\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$—as will be shown—and we write U_Φ for the topological interior of the set $\{\sigma < \infty\} \subset (\mathbb{R}^d)^*$. Finally, call a shift-invariant measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ a *minimiser* if μ satisfies the equation

$$\mathcal{H}(\mu|\Phi) = \sigma(S(\mu)) < \infty.$$

Let us start with the motivating result of this chapter.

Theorem (strict convexity of the surface tension)**.** *Let Φ denote a potential which decomposes as described above, and such that the induced specification γ^Φ is monotone.*

1. *If $E = \mathbb{R}$, then σ is strictly convex on U_Φ.*
2. *If $E = \mathbb{Z}$, then σ is strictly convex on U_Φ if for any affine map $h : (\mathbb{R}^d)^* \to \mathbb{R}$ with $h \leq \sigma$, the set $\{h = \sigma\} \cap \partial U_\Phi$ is convex. In particular, σ is strictly convex on U_Φ if $E = \mathbb{Z}$ and at least one of the following conditions is satisfied:*
   (a) σ is affine on ∂U_Φ, but not on $\bar{U}_\Phi$,
   (b) σ is not affine on $[u_1, u_2]$ for any distinct $u_1, u_2 \in \partial U_\Phi$ such that $[u_1, u_2] \not\subset \partial U_\Phi$.

See Theorem 3.4.12 for the formal statement of this theorem. The extra condition for $E = \mathbb{Z}$ is necessary to control the behaviour of ergodic measures whose slope is extremal. It is shown in the last part of this chapter that this condition holds true for all classical models. What happens in general is that measures whose slope lies in ∂U_Φ have zero combinatorial entropy, which makes it straightforward to derive the inequalities required for satisfying the extra condition. However, it is possible to design exotic models for which it is not known if the condition holds true or not, and consequently we cannot rule out the existence of an affine part of the surface tension for such exotic models.

Our second main result concerns a characterisation of minimisers, for potentials which decompose as described above. This generalises the results of Chapter 2 to the gradient setting. It is valid even if γ^Φ fails to be monotone, and if σ fails to be strictly convex. However, if σ is strictly convex, then there exists an ergodic minimiser of slope u for any $u \in U_\Phi$.

Theorem (minimisers of the specific free energy)**.** *Consider a potential Φ which decomposes as described, as well as a minimiser $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$. Then μ has finite energy in the sense of Burton and Keane, which means that any local configuration that is Lipschitz, has a positive density (if $E = \mathbb{R}$) or probability (if $E = \mathbb{Z}$) of occurring. Moreover, if the specification γ^Φ is quasilocal, then μ is a Gibbs measure, and if γ^Φ is not quasilocal but if μ is supported on its points of quasilocality, then μ is an almost Gibbs measure—which implies in particular that $\mu = \mu\gamma^\Phi_\Lambda$ for any $\Lambda \subset\subset \mathbb{Z}^d$. Finally, if μ is not supported on the points of quasilocality of γ^Φ, then we obtain results on the regular conditional probability distributions of μ which are similar in spirit to those obtained in Chapter 2.*

See Theorem 3.4.4 for the formal statement of this theorem.

The third main result of this chapter is a large deviations principle. This large deviations principle concerns both the macroscopic profile of a height function, as well as the local statistics of the height function within a region of macroscopic size. Its formal description requires a significant amount of technical constructions, for which we refer to Sections 3.4 and 3.11. One can also consider the large deviations principle on macroscopic profiles only, and the rate function so appearing is given by (3.1.1) up to an additive constant so that its minimum equals zero. This immediately implies the classical variational principle of [5]. The formal statements are included in Theorem 3.4.10, Corollary 3.4.11, and Theorem 3.11.5.

Theorem (variational principle)**.** *Consider a potential Φ which decomposes as above. Let $(D_n, b_n)_{n\in\mathbb{N}}$ denote a sequence of pairs of discrete regions $D_n \subset\subset \mathbb{Z}^d$ and boundary conditions $b_n \in \Omega$ which, after rescaling, suitably approximates some continuous region $D \subset \mathbb{R}^d$ endowed with some boundary function $b : \partial D \to \mathbb{R}$. Then the random function*

f_n obtained by sampling a configuration from $\gamma^\Phi_{D_n}(\cdot, b_n)$ and rescaling, is contained with high probability as $n \to \infty$ in any neighbourhood of the set of minimisers f^* of the integral

$$\int_D \sigma(\nabla f(x))dx$$

over all functions $f : \bar{D} \to \mathbb{R}$ which equal b on ∂D. If σ is strictly convex, then this minimiser f^* is unique, in which case $f_n \to f^*$ in probability as $n \to \infty$.

In the final part of this chapter, we provide several applications of our results.

Sheffield conjectured that similar results to those obtained in [54] apply to finite-range submodular potentials, that is, finite-range potentials which satisfy the FKG lattice condition. We prove that our framework applies to submodular Lipschitz potentials, and we prove that the extra condition for $E = \mathbb{Z}$ is automatically satisfied if the model of interest is $\mathcal{L}$-invariant for $\mathcal{L}$ equal to the full lattice $\mathbb{Z}^d$. In fact, we do not even require that the submodular potential of interest has finite range. See Theorem 3.4.14 for the corresponding formal statements.

We furthermore consider the model of uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree. Remark that k-regular trees are also Cayley graphs of finitely generated free groups. We confirm the conjecture in [44], which asserts that the surface tension associated with this model is strictly convex: see Theorem 3.4.15. This is remarkable because our theory is phrased in terms of $\mathbb{R}$- or $\mathbb{Z}$-valued functions only.

3.1.3 Ideas and strategy of the proof

The proof of the main results splits into two parts. The first part develops a range of thermodynamical machinery for the class of potentials under consideration. The line of thought motivating these results and proofs was already present in the literature, most notably in the work of Georgii [20] and Sheffield [54], as well as in Chapter 2. However, it requires significant effort to adapt these existing tools to the generality of our setting. The second part provides a proof of strict convexity of the surface tension, if the potential of interest furthermore induces a specification that is stochastically monotone. This is where we break new ground. Sheffield [54] proves that the surface tension is strictly convex by employing the following general strategy:

1. Suppose that σ is affine on a line segment $[u_1, u_2]$ for $u_1, u_2 \in U_\Phi$ distinct,
2. Construct a shift-invariant gradient measure in $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ of slope $u = (u_1 + u_2)/2$ with minimal specific free energy and which does not have finite energy,
3. Conclude that this contradicts the characterisation of the minimisers of the specific free energy, as mentioned earlier in this introduction.

The same strategy is employed here, but the construction of the gradient measure, as well as the heuristic that this construction is based on, are entirely original. The remainder of this subsection gives an overview of this construction.

First, the surface tension $\sigma(u)$ at some slope u can be expressed in terms of the asymptotic behaviour of the partition function of $\gamma_{\Pi_n}(\cdot, \phi^u)$ where ϕ^u approximates u in some precise sense: this is a consequence of the large deviations principle. We then consider the product measure $\mu := \gamma_{\Pi_n}(\cdot, \phi^u) \times \gamma_{\Pi_n}(\cdot, \phi^u)$; write (ϕ_1, ϕ_2) for the random pair of height functions in μ, and write f for the difference $\phi_1 - \phi_2$.

One can use the fact that σ is affine on the line segment $[u_1, u_2]$ to derive that the function f deviates macroscopically—that is, at scale n—from 0 with log probability of order $o(n^d)$ as $n \to \infty$. We then use monotonicity of the specification γ to compare the probability of a macroscopic deviation of f to the probability that the set $\{f \in [a, b]\} \subset \Pi_n \subset\subset \mathbb{Z}^d$ has many large connected components for fixed $0 < a < b < \infty$. This requires the development of an essential and original geometrical construction. The connected components of $\{f \in [a, b]\}$ of interest are called *moats*. Finally, we randomly shift the functions ϕ_1 and ϕ_2 by a vector in $\Pi_n \cap \mathcal{L}$ and take limits to produce a shift-invariant measure on the product space, such that each marginal has slope u. The two lower bounds on probabilities imply an upper bound on the specific free energy of this product measure. We show that the moats—the large connected components of $\{f \in [a, b]\}$—grow to be distinct infinite components in this limiting procedure. This contradicts that for a shift-invariant measure with finite energy, the random set $\{f \in [a, b]\}$ cannot have more than one infinite component due to the argument of Burton and Keane: the desired contradiction.

Let us finally elaborate briefly on the geometrical construction involving moats. The goal is to find a lower bound on the probability that $\{f \in [a, b]\}$ has many large level set, in terms of the probability that f deviates macroscopically from 0. Write $c_n := (\lfloor n/2 \rfloor, \dots, \lfloor n/2 \rfloor) \in \Pi_n$ for the centre vertex of Π_n, and suppose, by means of illustration, that $f(c_n) > \varepsilon n$ for some $\varepsilon > 0$. If ϕ_1 and ϕ_2 are K-Lipschitz for some $K \in (0, \infty)$, then f is $2K$-Lipschitz. Choose $a = 4K$ and $b = 8K$. Since $f(c_n)$ is large and since f equals 0 on the complement of Π_n, we observe that $\{f \in [a, b]\}$ must contain a connected component which is contained in Π_n and surrounds the vertex c_n in some precise sense. This connected component is called a *moat*. Now fix an arbitrary connected set $M \subset \Pi_n$, and condition on the event that M is a moat, and that f is larger than b directly inside M. Equipped with monotonicity, it is straightforward to demonstrate that it is more likely (in this conditioned measure) that $f(c_n) \leq -\varepsilon n + 10K$, than that $f(c_n) \geq \varepsilon n$. But if $f(c_n) \leq -\varepsilon n + 10K$ and if f is larger than b directly on the inside of M, then $\{f \in [a, b]\}$ must have another connected component which surrounds c_n, and which is in turn surrounded by the original moat M. One can continue this procedure to generate a sequence of moats of length $\lfloor \varepsilon n / 10K \rfloor$, such that each moat surrounds the moat that succeeds it. It is important that the union of all moats occupy a uniformly positive proportion of Π_n as $n \to \infty$, so that they do not disappear in the limiting procedure after rerandomising the position of the origin; this is indeed the case because of the lower bound on the number of moats.

3.1.4 Open questions

The first natural question which is left open in this work is to decide if it is possible to drop the requirement that random functions are Lipschitz. We believe that it is indeed the case, a significant clue being that this requirement does not appear in [54]. Finding a way around this restriction would open the main result to a whole new class of interactions. However, the geometrical construction involving the moats relies heavily on the Lipschitz property.

Secondly, it would be interesting to study how the requirement of stochastic monotonicity can be relaxed. Results on strict convexity of the surface tension have been obtained for some non-monotone models for a class of non-convex potentials [7, 6, 1], and for small non-monotone perturbations of dimer models [21, 22]. In the simulation on the right in Figure 3.1, macroscopic disorder is explained by a heuristic.

For this simulation, the parameters of the model are chosen such that straight lines are much preferred over corners. This means that the random surface is able to build *momentum*: deviations from the mean reinforce each other. This is the exact opposite of stochastic monotonicity. However, there are more subtle (and potentially more local) ways in which stochastic monotonicity might fail. A simple example would be to consider random 1-Lipschitz functions from $\mathbb{Z}^d$ to $\mathbb{Z}$, where the potential discourages neighbouring vertices from taking the exact same value. It is easy to show that this model is not monotone, but there is no heuristic of momentum building which would imply macroscopic disorder. Perhaps it would be possible to prove that this model is stochastically monotone in some relaxed sense, in which case the results on moats could be adapted to fit this model.

3.1.5 Structure of the chapter

Section 3.2 provides an overview of the objects which play a role in the study of random surfaces. These definitions are standard, and derive mainly from the work of Georgii [20] and Sheffield [54]. Section 3.3 presents the class of models which fall under the scope of this chapter. Formal statements of the main results are contained in Section 3.4. The heart of the chapter is contained in Section 3.5, which is independent of the rest of the chapter. It only refers to some of the most basic constructions in Section 3.2. Sections 3.6–3.11 develop the thermodynamical machinery necessary for understanding random surfaces, and culminate in the large deviations principle in Section 3.11. Section 3.6 contains a number of observations concerning the geometry induced by the Lipschitz condition. Section 3.7 studies properties of the specific free energy. Section 3.8 contains a characterisation of the minimisers of the specific free energy, and essentially adapts the arguments of Chapter 2 to the setting of random surfaces. In Section 3.9 we state and adapt some standard results on ergodic decompositions from [20]. Sections 3.10 and 3.11 extend Chapters 6 and 7 of [54] to the infinite-range Lipschitz setting. Once the large deviations principle has been established, we combine it with the theory of moats from Section 3.5 to prove that the surface tension is strictly convex. These are the contents of Section 3.12. Section 3.13 finally contains a number of significant applications of our theory.

3.2 The thermodynamical formalism

The interest is in distributions of the random function ϕ which assigns a value $\phi(x)$ from E to each vertex $x \in \mathbb{Z}^d$, where $d \geq 2$ and—depending on the model of interest—E denotes either $\mathbb{Z}$ or $\mathbb{R}$. Such distributions are studied in relation to an underlying model, which encodes the interactions that exist between the function values of ϕ at different vertices in $\mathbb{Z}^d$. At the very least, the underlying model must give rise to a functional, which assigns a real number—the specific free energy—to any shift-invariant distribution of ϕ. In the non-gradient setting there are at least three ways to characterise the model of interest:

1. Through a reference measure on E and an interaction potential,

2. Through a reference distribution of ϕ,

3. Directly through the specification.

Each formulation has slightly different properties, but they all generate a suitable entropy functional whenever the correct conditions are imposed. See Chapter 2 for an overview. In the gradient setting of this chapter we must be more careful, and it seems that only the first formulation generates a suitable entropy functional. The goal of this section is to efficiently describe the standard objects for the formal framework of gradient models on $\mathbb{Z}^d$.

Subsection 3.2.1 introduces the necessary objects and symmetries for the shift-invariant gradient setting. The same subsection also introduces the key restrictions on the model: that the specification is monotone, and that it produces Lipschitz functions. Subsection 3.2.2 describes the formalism of potentials. Subsection 3.2.3 introduces the specific free energy and the surface tension. The specific free energy is well-defined for all potentials Φ in the class $\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$ which is introduced in Section 3.3; we prove existence of the specific free energy in Section 3.7. All definitions in the current section are standard.

3.2.1 The gradient formalism

Height functions

We are interested in distributions of the random function ϕ, which assigns values from the measure space $(E, \mathcal{E}, \lambda)$ to the vertices of the square lattice $\mathbb{Z}^d$. Here E refers to either $\mathbb{Z}$ or $\mathbb{R}$, depending on the context, $\mathcal{E}$ is the Borel σ-algebra, and λ denotes the counting measure (if $E = \mathbb{Z}$) or the Lebesgue measure (if $E = \mathbb{R}$). The choice of E is considered fixed throughout the entire work. The set of all functions ϕ from $\mathbb{Z}^d$ to E is denoted by Ω. Functions in Ω are called *samples* or *height functions*. For $\Lambda \subset \mathbb{Z}^d$ and $\phi \in \Omega$, write $\phi_\Lambda \in E^\Lambda$ for the restriction $\phi|_\Lambda$. If furthermore $\Delta \subset \mathbb{Z}^d$ and $\psi \in \Omega$ with Λ and Δ disjoint, then write $\phi_\Lambda\psi_\Delta \in E^{\Lambda\cup\Delta}$ for the unique function that restricts to ϕ on Λ and to ψ on Δ.

Subsets of $\mathbb{Z}^d$

Write $\Lambda \subset\subset \mathbb{Z}^d$ if Λ is a finite subset of $\mathbb{Z}^d$. Throughout this chapter, we shall reserve the notation $(\Pi_n)_{n\in\mathbb{N}}$ for the sequence of subsets of $\mathbb{Z}^d$ defined by $\Pi_n := [0, n)^d \subset\subset \mathbb{Z}^d$ for each $n \in \mathbb{N}$. Remark that $|\Pi_n| = n^d$ for any $n \in \mathbb{N}$.

Next, introduce two notions of boundary for subsets Λ of $\mathbb{Z}^d$. Write $\partial\Lambda$ for set of the vertices which are adjacent to Λ in the square lattice. Write $\partial^n\Lambda$ for the set of vertices in Λ which are at d_1-distance at most n from $\mathbb{Z}^d \smallsetminus \Lambda$, for any $n \in \mathbb{Z}_{\geq 0}$; here d_1 is the graph metric corresponding to the square lattice. Write also Λ^{-n} for $\Lambda \smallsetminus \partial^n\Lambda$. If $D \subset \mathbb{R}^d$, then write $\Lambda(D) := D \cap \mathbb{Z}^d$ and $\Lambda^{-n}(D) := (\Lambda(D))^{-n}$.

Now let $(\Lambda_n)_{n\in\mathbb{N}}$ denote a sequence of subsets of $\mathbb{Z}^d$. If all sets Λ_n are finite with $|\Lambda_n| \to \infty$ and $|\partial\Lambda_n|/|\Lambda_n| \to 0$ as $n \to \infty$, then $(\Lambda_n)_{n\in\mathbb{N}}$ is called a *Van Hove sequence*. We write $(\Lambda_n)_{n\in\mathbb{N}} \uparrow \mathbb{Z}^d$ to mean that $(\Lambda_n)_{n\in\mathbb{N}}$ is a Van Hove sequence. The sequence $(\Pi_n)_{n\in\mathbb{N}}$ is an example of a Van Hove sequence.

σ-Algebras and random fields

If $(X, \mathcal{X})$ is any measurable space, then write $\mathcal{P}(X, \mathcal{X})$ for the set of probability measures on it, and $\mathcal{M}(X, \mathcal{X})$ for the set of σ-finite measures. Define the following

σ-algebras on Ω for any $\Lambda \subset \mathbb{Z}^d$:

$$\begin{aligned}
\mathcal{F} &:= \sigma(\phi(x) : x \in \mathbb{Z}^d), & \mathcal{F}_\Lambda &:= \sigma(\phi(x) : x \in \Lambda), \\
\mathcal{F}^\nabla &:= \sigma(\phi(y) - \phi(x) : x, y \in \mathbb{Z}^d), & \mathcal{F}_\Lambda^\nabla &:= \sigma(\phi(y) - \phi(x) : x, y \in \Lambda).
\end{aligned}$$

A *random field* is a probability measure in $\mathcal{P}(\Omega, \mathcal{A})$ for some σ-algebra $\mathcal{A} \subset \mathcal{F}$. We introduce the gradient σ-algebra $\mathcal{F}^\nabla$ because it is often not possible to measure the height $\phi(x)$ directly; only the height differences $\phi(y) - \phi(x)$ are measurable. Note that, with the above definitions, $\mathcal{F}_\Lambda^\nabla = \mathcal{F}^\nabla \cap \mathcal{F}_\Lambda$. For $\Lambda \subset \mathbb{Z}^d$, write π_Λ for the natural probability kernel from $(\Omega, \mathcal{F})$ to $(E^\Lambda, \mathcal{E}^\Lambda)$ which restricts random fields to Λ.

A *cylinder set* is a measurable subset of Ω which is contained in $\mathcal{F}_\Lambda$ for some $\Lambda \subset\subset \mathbb{Z}^d$; a *cylinder function* is a function $\Omega \to \mathbb{R}$ which is $\mathcal{F}_\Lambda$-measurable for some $\Lambda \subset\subset \mathbb{Z}^d$. A cylinder function is called *continuous* if it is continuous with respect to the topology of uniform convergence on Ω. Note that all cylinder functions are continuous whenever $E = \mathbb{Z}$.

Define the further σ-algebras on Ω for any $\Lambda \subset \mathbb{Z}^d$:

$$\mathcal{T}_\Lambda := \mathcal{F}_{\mathbb{Z}^d \smallsetminus \Lambda}, \quad \mathcal{T} := \cap_{\Delta \subset\subset \mathbb{Z}^d} \mathcal{T}_\Delta, \quad \mathcal{T}_\Lambda^\nabla := \mathcal{T}_\Lambda \cap \mathcal{F}^\nabla, \quad \mathcal{T}^\nabla := \mathcal{T} \cap \mathcal{F}^\nabla.$$

Sets in $\mathcal{T}$ are called *tail-measurable.*

The topology of (weak) local convergence

The *topology of local convergence* is the coarsest topology on $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$ that makes the map $\mu \mapsto \mu(f)$ continuous for any bounded cylinder function f. The *topology of weak local convergence* is the coarsest topology on $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$ that makes the map $\mu \mapsto \mu(f)$ continuous for any bounded continuous cylinder function f. Note that the two topologies coincide whenever $E = \mathbb{Z}$. Section 3.10 uses a particular basis $\mathcal{B}$ for the topology of weak local convergence on $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$. This basis $\mathcal{B}$ is defined such that it contains exactly all sets $B \subset \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ which can be written as finite intersections of open sets of the form $\{\mu : a < \mu(f) < b\}$, where $a, b \in \mathbb{R}$ and where f is a continuous bounded cylinder function.

Shift-invariance and ergodicity

To see convergence of the model at a macroscopic scale it is important that the model exhibits shift-invariance. For $x \in \mathbb{Z}^d$, write $\theta_x : \mathbb{Z}^d \to \mathbb{Z}^d$, $y \mapsto y + x$. Throughout this chapter, the letter $\mathcal{L}$ denotes a fixed full-rank sublattice of $\mathbb{Z}^d$, and $\Theta = \Theta(\mathcal{L}) = \{\theta_x : x \in \mathcal{L}\}$ is the corresponding group of translations of $\mathbb{Z}^d$. If $\phi \in \Omega$ and $\theta \in \Theta$, then $\theta\phi$ denotes the unique height function satisfying $(\theta\phi)(x) = \phi(\theta x)$ for all x. Similarly, define

$$\theta A := \{\theta\phi : \phi \in A\}, \quad \theta\mathcal{A} := \{\theta A : A \in \mathcal{A}\}, \quad \theta\mu : \theta\mathcal{A} \to [0, \infty], \theta\mu(\theta A) \mapsto \mu(A)$$

for $A \subset \Omega$, for $\mathcal{A}$ a sub-σ-algebra of $\mathcal{F}$, and for μ a measure on $\mathcal{A}$. Any of these three objects is called *$\mathcal{L}$-invariant* if they are invariant under θ for any $\theta \in \Theta$. If $\mathcal{A}$ is an $\mathcal{L}$-invariant σ-algebra on Ω, then write $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{A})$ for the collection of $\mathcal{L}$-invariant probability measures on $(\Omega, \mathcal{A})$. Note that $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{A})$ is the set of probability measures on $(\Omega, \mathcal{A})$ such that ϕ and $\theta\phi$ have the same distribution for any $\theta \in \Theta$.

Define finally

$$\mathcal{I}_\mathcal{L} := \{A \in \mathcal{F} : A = \theta A \text{ for all } \theta \in \Theta\}, \qquad \mathcal{I}_\mathcal{L}^\nabla := \mathcal{I}_\mathcal{L} \cap \mathcal{F}^\nabla.$$

A gradient measure $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ is called *ergodic* if μ is $\mathcal{L}$-invariant and trivial on $\mathcal{I}_\mathcal{L}^\nabla$. Write $\mathrm{ex}\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ for the set of all such ergodic gradient measures. Write $e(\mathrm{ex}\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla))$ for the smallest σ-algebra that makes the map $A \mapsto \mu(A)$ measurable for all $A \in \mathcal{F}^\nabla$.

Specifications

A *specification* is a family $\gamma = (\gamma_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^d}$ of probability kernels, such that

1. γ_Λ is a probability kernel from $(\Omega, \mathcal{T}_\Lambda)$ to $(\Omega, \mathcal{F})$ for each $\Lambda \subset\subset \mathbb{Z}^d$,
2. $\mu\gamma_\Lambda(A) = \mu(A)$ for any $\Lambda \subset\subset \mathbb{Z}^d$, $A \in \mathcal{T}_\Lambda$, and $\mu \in \mathcal{P}(\Omega, \mathcal{F})$,
3. $\gamma_\Lambda\gamma_\Delta = \gamma_\Lambda$ for any $\Delta \subset \Lambda \subset\subset \mathbb{Z}^d$.

The specification defines the local behaviour of the model, and we think of $\gamma_\Lambda(\cdot, \phi)$ as the *local Gibbs measure* in $\Lambda \subset\subset \mathbb{Z}^d$ with boundary conditions $\phi \in \Omega$. A specification γ is called *$\mathcal{L}$-invariant* if $\gamma_\Lambda(\cdot, \theta\phi) = \theta\gamma_{\theta\Lambda}(\cdot, \phi)$ for any $\Lambda \subset\subset \mathbb{Z}^d$, $\phi \in \Omega$, and $\theta \in \Theta$. Call γ a *gradient specification* if the distribution of $\psi + a$ in $\gamma_\Lambda(\cdot, \phi)$ equals that of ψ in $\gamma_\Lambda(\cdot, \phi + a)$ for any $\Lambda \subset\subset \mathbb{Z}^d$, $\phi \in \Omega$, and $a \in E$, where ψ denotes the random height function in each local Gibbs measure. Note that each kernel γ_Λ restricts to a kernel from $(\Omega, \mathcal{T}_\Lambda^\nabla)$ to $(\Omega, \mathcal{F}^\nabla)$ whenever γ is a gradient specification.

Monotonicity

An event $A \in \mathcal{F}$ is called *increasing* if $\phi \in A$ and $\psi \geq \phi$ implies $\psi \in A$. Consider two measures $\mu_1, \mu_2 \in \mathcal{P}(\Omega, \mathcal{F})$. Say that μ_2 *stochastically dominates* μ_1, and write $\mu_1 \preceq \mu_2$, if $\mu_1(A) \leq \mu_2(A)$ for any increasing event A. This is equivalent to asking that there exists a coupling between the two measures such that $\phi_1 \leq \phi_2$ almost surely, where the distributions of ϕ_1 and ϕ_2 are prescribed by the measures μ_1 and μ_2 respectively. A specification γ is called *monotone* if for each $\Lambda \subset\subset \mathbb{Z}^d$, the kernel γ_Λ preserves the partial order $\preceq$ on $\mathcal{P}(\Omega, \mathcal{F})$. Now consider a fixed measurable set $A \in \mathcal{F}$, and use—in this definition—the shorthand $\mathcal{P}_A$ for the set $\{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \mu(A) = 1\}$. The specification γ is called *monotone over* A if $\mu\gamma_\Lambda \in \mathcal{P}_A$ for any $\Lambda \subset\subset \mathbb{Z}^d$ and $\mu \in \mathcal{P}_A$, and if γ_Λ preserves the partial order $\preceq$ on $\mathcal{P}_A$. The assumption that γ is monotone over a suitable set of Lipschitz functions is crucial to the proof of strict convexity of the surface tension.

The Lipschitz property

Consider some fixed constant $K \in [0, \infty)$. A height function is called *K-Lipschitz* if that height function is K-Lipschitz with respect to the graph metric d_1 on the square lattice $\mathbb{Z}^d$. A measure is called *K-Lipschitz* if it is supported on K-Lipschitz functions. The Lipschitz property is further refined in Subsection 3.3.1.

The slope

Consider $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$. If $\phi(y) - \phi(x)$ is μ-integrable for any $x, y \in \mathbb{Z}^d$, then μ is said to have *finite slope*. If μ has finite slope, then shift-invariance of μ implies that the function

$$\mathcal{L} \to \mathbb{R},\, x \mapsto \mu(\phi(x) - \phi(0))$$

is additive. In particular, this means that there is a unique linear functional $u \in (\mathbb{R}^d)^*$ such that

$$u(x) = \mu(\phi(x) - \phi(0))$$

for any $x \in \mathcal{L} \subset \mathbb{R}^d$. This linear functional u is called the *slope* of μ, and we write $S(\mu)$ for it. The map S is affine: it is clear that $S((1-t)\mu+t\nu) = (1-t)S(\mu)+tS(\nu)$ for any $t \in [0,1]$ and for any $\mu, \nu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with finite slope.

If we restrict to K-Lipschitz measures in $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ for fixed $K \in [0, \infty)$, then all measures have finite slope, and the map $\mu \mapsto S(\mu)$ is then continuous with respect to the topology of (weak) local convergence.

3.2.2 Interaction potentials, reference measures, and specifications

Interaction potentials

All models are formalised in terms of an *interaction potential* $\Phi = (\Phi_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^d}$, which is a family of *potential functions* $\Phi_\Lambda : \Omega \to \mathbb{R} \cup \{\infty\}$ where each function Φ_Λ is required to be measurable with respect to $\mathcal{F}_\Lambda$. The potential Φ is called a *gradient potential* if each function Φ_Λ is in addition $\mathcal{F}_\Lambda^\nabla$-measurable. The potential Φ is furthermore called *$\mathcal{L}$-invariant* or *periodic* if $\Phi_{\theta\Lambda}(\phi) = \Phi_\Lambda(\theta\phi)$ for all $\theta \in \Theta$ and for any $\phi \in \Omega$. In the sequel, Φ shall always denote a fixed periodic gradient potential. It is always conventionally assumed that $\Phi_\Lambda \equiv 0$ whenever Λ is a singleton or empty because the σ-algebra $\mathcal{F}_\Lambda^\nabla$ is then trivial.

Next, introduce the Hamiltonian. For $\Lambda \subset\subset \mathbb{Z}^d$ and $\Delta \subset \mathbb{Z}^d$ containing Λ, let $H_{\Lambda,\Delta}$ denote the $\mathcal{F}_\Delta^\nabla$-measurable function from Ω to $\mathbb{R} \cup \{\infty\}$ defined by

$$H_{\Lambda,\Delta} := \sum\nolimits_{\Gamma \subset\subset \mathbb{Z}^d \text{ with } \Gamma \subset \Delta \text{ and with } \Gamma \text{ intersecting } \Lambda} \Phi_\Gamma.$$

In particular, we write $H_\Lambda := H_{\Lambda,\mathbb{Z}^d}$ and $H_\Lambda^0 := H_{\Lambda,\Lambda}$. We shall soon introduce further conditions on Φ which ensure that the sum in the display is always well-defined and bounded below. The function H_Λ is called the *Hamiltonian* of Λ and H_Λ^0 is called the *interior Hamiltonian* of Λ. We add a superscript Φ to this notation whenever multiple interaction potentials are considered and confusion might possibly arise.

Reference measures

For any fixed nonempty $\Lambda \subset\subset \mathbb{Z}^d$, there exist natural reference measures on the measurable spaces $(\Omega, \mathcal{F}_\Lambda)$ and $(\Omega, \mathcal{F}_\Lambda^\nabla)$, in terms of the previously introduced reference measure λ on $(E, \mathcal{E})$. In the non-gradient setting this is straightforward: the map $\phi \mapsto \phi_\Lambda$ extends to a bijection from $\mathcal{F}_\Lambda$ to $\mathcal{E}^\Lambda$, and λ^Λ is a measure on $(E^\Lambda, \mathcal{E}^\Lambda)$. With only slight abuse of notation, we write also λ^Λ for the unique measure on $(\Omega, \mathcal{F}_\Lambda)$ that makes the map $\phi \mapsto \phi_\Lambda$ into a measure-preserving projection from $(\Omega, \mathcal{F}_\Lambda, \lambda^\Lambda)$ to $(E^\Lambda, \mathcal{E}^\Lambda, \lambda^\Lambda)$. We must be more subtle in the gradient setting: we cannot measure the height of ϕ directly, and so we cannot pullback the measure λ^Λ. Fix therefore some reference point $x \in \Lambda$ and set $\Lambda' := \Lambda \smallsetminus \{x\}$, and consider instead the map $\phi \mapsto \phi_{\Lambda'} - \phi(x)$. This map extends to a bijection from $\mathcal{F}_\Lambda^\nabla$ to $\mathcal{E}^{\Lambda'}$. Abuse notation again by writing $\lambda^{\Lambda-1}$ for the unique measure on $(\Omega, \mathcal{F}_\Lambda^\nabla)$ that turns the map $\phi \mapsto \phi_{\Lambda'} - \phi(x)$ into a measure-preserving projection from $(\Omega, \mathcal{F}_\Lambda^\nabla, \lambda^{\Lambda-1})$ to $(E^{\Lambda'}, \mathcal{E}^{\Lambda'}, \lambda^{\Lambda'})$. The notation $\lambda^{\Lambda-1}$ bears no reference to the choice of $x \in \Lambda$, as the resulting measure $\lambda^{\Lambda-1}$ is indeed independent of this arbitrary choice. The gradient reference measures $\lambda^{\Lambda-1}$ are not used in the definition of the specification that Φ generates; they will first appear in the definition of the specific free energy.

The specification generated by a potential

The potential Φ generates a specification $\gamma^\Phi = (\gamma^\Phi_\Lambda)_{\Lambda\subset\subset\mathbb{Z}^d}$ defined by

$$\gamma^\Phi_\Lambda(A,\phi) := \frac{1}{Z^\Phi_\Lambda(\phi)}\int_{E^\Lambda} 1_A(\psi\phi_{\mathbb{Z}^d\smallsetminus\Lambda})e^{-H^\Phi_\Lambda(\psi\phi_{\mathbb{Z}^d\smallsetminus\Lambda})}d\lambda^\Lambda(\psi),$$

for any $\Lambda\subset\subset\mathbb{Z}^d$, $\phi\in\Omega$, and $A\in\mathcal{F}$, where $Z^\Phi_\Lambda(\phi)$ is the normalising constant

$$Z^\Phi_\Lambda(\phi) := \int_{E^\Lambda} e^{-H^\Phi_\Lambda(\psi\phi_{\mathbb{Z}^d\smallsetminus\Lambda})}d\lambda^\Lambda(\psi).$$

We drop the superscript Φ in this notation unless the choice of potential is ambiguous. Of course, $\gamma_\Lambda(\cdot,\phi)$ is a well-defined probability measure on $(\Omega,\mathcal{F})$ only if $Z_\Lambda(\phi)\in(0,\infty)$. Say that ϕ has *finite energy* if $\Phi_\Lambda(\phi)<\infty$ for any $\Lambda\subset\subset\mathbb{Z}^d$, and say that ϕ is *admissible* if it has finite energy and $Z_\Lambda(\phi)\in(0,\infty)$ for any $\Lambda\subset\subset\mathbb{Z}^d$. To draw a sample ψ from $\gamma_\Lambda(\cdot,\phi)$, set first ψ equal to ϕ on the complement of Λ, then sample ψ_Λ proportional to $e^{-H_\Lambda}\lambda^\Lambda$. Similarly, if μ is a probability measure on $(\Omega,\mathcal{T}_\Lambda)$ supported on admissible height functions, then $\mu\gamma_\Lambda$ is a probability measure on $(\Omega,\mathcal{F})$; to sample from $\mu\gamma_\Lambda$ one first obtains an auxiliary sample ϕ from μ; then one draws the final sample ψ from $\gamma_\Lambda(\cdot,\phi)$.

It is important to observe that γ is a gradient specification. This is due to the fact that Φ is a gradient potential which makes H_Λ measurable with respect to $\mathcal{F}^\nabla$, and because the reference measures λ and λ^Λ are invariant under translations.

3.2.3 The surface tension

Relative entropy

Recall first the relative entropy. If $(X,\mathcal{X},\nu)$ is an arbitrary σ-finite measure space and μ another probability measure on $(X,\mathcal{X})$, then the *relative entropy* of μ with respect to ν is defined by

$$\mathcal{H}(\mu|\nu) := \begin{cases}\mu(\log f)=\nu(f\log f) & \text{if } \mu\ll\nu \text{ where } f=d\mu/d\nu,\\ \infty & \text{otherwise.}\end{cases}$$

Remark that $\mathcal{H}(\mu|\nu)\in\mathbb{R}\cup\{-\infty,\infty\}$ in general, and that $\mathcal{H}(\mu|\nu)\geq-\log\nu(X)$. If ν is a finite measure, then we have equality if and only if μ is a scalar multiple of ν. If $\mathcal{A}$ is a sub-σ-algebra of $\mathcal{X}$, then use the shorthand $\mathcal{H}_\mathcal{A}(\mu|\nu)$ for $\mathcal{H}(\mu|_\mathcal{A}|\nu|_\mathcal{A})$.

The free energy

We are now ready to introduce the free energy. This already requires the presence of some gradient potential Φ, although we do not yet impose any condition on it. Consider also some gradient random field $\mu\in\mathcal{P}(\Omega,\mathcal{F}^\nabla)$, and some finite set $\Lambda\subset\subset\mathbb{Z}^d$. Then the *free energy* of μ in Λ with respect to Φ is defined by

$$\mathcal{H}_\Lambda(\mu|\Phi) := \mathcal{H}_{\mathcal{F}^\nabla_\Lambda}(\mu|e^{-H^{0,\Phi}_\Lambda}\lambda^{\Lambda-1}) = \mathcal{H}_{\mathcal{F}^\nabla_\Lambda}(\mu|\lambda^{\Lambda-1})+\mu(H^{0,\Phi}_\Lambda).$$

The free energy is sometimes decomposed into the *entropy* and the *energy* of μ in Λ—the two terms in the rightmost expression in the display respectively. (For the final equality, we adopt the convention that $\infty-\infty=\infty$.)

The specific free energy

The *specific free energy* of a shift-invariant random field $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with respect to Φ is defined by the limit

$$\mathcal{H}(\mu|\Phi) := \lim_{n\to\infty} n^{-d}\mathcal{H}_{\Pi_n}(\mu|\Phi).$$

The specific free energy thus describes the asymptotic of the normalised free energy of μ with respect to Φ over a large box. In Section 3.7 we prove that the limit converges for all Φ in the class $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ which is described in Section 3.3. It is also shown in Section 3.7 that $\mathcal{H}(\cdot|\Phi)$ is affine and bounded below.

The surface tension

Consider a potential Φ in our class $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, which implies that the specific free energy is well-defined, affine and bounded below. The *surface tension* is the function $\sigma : (\mathbb{R}^d)^* \to \mathbb{R} \cup \{\infty\}$ defined by

$$\sigma(u) := \inf_{\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla) \text{ with } S(\mu) = u} \mathcal{H}(\mu|\Phi).$$

The function σ must be convex because both $S(\cdot)$ and $\mathcal{H}(\cdot|\Phi)$ are affine. We shall write U_Φ for the interior of the convex set $\{\sigma < \infty\} \subset (\mathbb{R}^d)^*$. Slopes in U_Φ are called *allowable*. The major contribution of this chapter is that we show that σ is strictly convex on U_Φ whenever γ^Φ is monotone over the set of admissible height functions and if Φ is in our class $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ (and under an additional condition whenever $E = \mathbb{Z}$).

3.3 The class of models under consideration

In the following four subsections, we describe the conditions which are imposed on the model of interest: these are specific to this work, and this is where we broaden the class of models for which strict convexity of the surface tension can be derived. Subsection 3.3.1 describes the Lipschitz setting in more detail. We take great care in formulating the Lipschitz condition: this is not necessary for the arguments to work, but it rather minimises the restrictions imposed on the class of models. Let us now consider the potential which generates the model. The potential of the model of interest must decompose as the sum of two potentials, where the first component is a strong, local potential which—at the very least—enforces the Lipschitz condition (Subsection 3.3.2), and where the second component is a weak interaction of infinite range (Subsection 3.3.3). The word *weak* here is only relative to the word *strong* that was used to describe the first potential: in particular, we do not mean to imply that the second component demonstrates any sort of decay over long distances. It is the second potential that allows us to assign energy to large geometric objects, such as level sets. Subsection 3.3.4 finally gives an overview of the objects describing the model of interest, and which are considered fixed throughout most of the analysis.

3.3.1 Local Lipschitz constraints

We require that a height function has finite energy if and only if it is Lipschitz with respect to the correct quasimetric. We shall allow quasimetrics (subject to certain necessary constraints) in order to be as general as possible. The Lipschitz constraint must be enforced locally by the potential, due to the nature of the arguments that we use to derive the main result. This means that for each vertex $x \in \mathbb{Z}^d$ we are allowed to enforce a Lipschitz constraint between x and only finitely many other vertices $y \in \mathbb{Z}^d$. In other words, what we have in mind is a set $A \subset \mathbb{Z}^d \times \mathbb{Z}^d \times \mathbb{R}$, such that a height function ϕ is Lipschitz if and only if $\phi(y) - \phi(x) \leq a$ for any $(x, y, a) \in A$, and such that A becomes a finite set once we identify each triple of the form (x, y, a) with all triples of the form $(\theta x, \theta y, a)$ as θ ranges over Θ. The local Lipschitz constraint also enforces that the functions are globally Lipschitz with respect to the correct quasimetric. This is formalised as follows.

Definition 3.3.1 (local Lipschitz constraint)**.** Call an edge set $\mathbb{A}$ on $\mathbb{Z}^d$ an *admissible graph* if $\mathbb{A}$ is $\mathcal{L}$-invariant and makes $(\mathbb{Z}^d, \mathbb{A})$ a connected graph of bounded degree. Call a function $q : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}$ an *admissible quasimetric* if

1. $q(x, x) = 0$ for any $x \in \mathbb{Z}^d$,
2. $q(x, y) + q(y, x) > 0$ for any $x, y \in \mathbb{Z}^d$ distinct,
3. $q(x, z) \leq q(x, y) + q(y, z)$ for any $x, y, z \in \mathbb{Z}^d$,
4. $q(\theta x, \theta y) = q(x, y)$ for any $x, y \in \mathbb{Z}^d$ and $\theta \in \Theta$.

Such a function is called *integral* if it takes integral values. A *local Lipschitz constraint* is a pair $(\mathbb{A}, q)$ where

1. $\mathbb{A}$ is an admissible graph,
2. q is an admissible quasimetric,
3. q is maximal among all admissible quasimetrics that equal q on $\mathbb{A}$, in the sense that $p \leq q$ for any admissible quasimetric p with $p(x, y) \leq q(x, y)$ for all $\{x, y\} \in \mathbb{A}$.

If $(\mathbb{A}, q)$ is a local Lipschitz constraint and $\varepsilon \geq 0$ a sufficiently small constant, then write q_ε for the largest admissible quasimetric subject to $q_\varepsilon(x, y) \leq q(x, y) - \varepsilon$ for all $\{x, y\} \in \mathbb{A}$. (It is demonstrated in Proposition 3.6.5 that this is indeed well-defined for $\varepsilon > 0$ sufficiently small.) Note that the resulting pair $(\mathbb{A}, q_\varepsilon)$ is also a local Lipschitz constraint.

Remarks.

1. The last condition in the definition of a local Lipschitz constraint guarantees that q is fully determined by its values on the edges in $\mathbb{A}$.
2. We shall sometimes omit the reference to $\mathbb{A}$ and simply call q the *local Lipschitz constraint*. If $(\mathbb{A}, q)$ is a local Lipschitz constraint and $\mathbb{B}$ another admissible graph on $\mathbb{Z}^d$, then the pair $(\mathbb{A} \cup \mathbb{B}, q)$ is also a local Lipschitz constraint producing the same quasimetric q. We shall always assume, without loss of generality, that $\mathbb{A}$ contains the edges of the square lattice.
3. If q is a local Lipschitz constraint, then there is a constant $K < \infty$ such that $K d_1 \geq q$.
4. We do not impose that q takes values in $[0, \infty)$. This restriction is not necessary to make the arguments work.

From now on, we shall always have in mind a fixed local Lipschitz constraint $(\mathbb{A}, q)$.

Definition 3.3.2 (q-Lipschitz)**.** A function $\phi : \mathbb{Z}^d \to \mathbb{R}$ is called *q-Lipschitz* if, for every $x, y \in \mathbb{Z}^d$,

$$\phi(y) - \phi(x) \leq q(x, y).$$

The function ϕ is called *q-Lipschitz at $z \in \mathbb{Z}^d$* if this inequality is satisfied for any edge $\{x, y\} \in \mathbb{A}$ containing z. Naturally extend these definitions to cover the cases that $\phi : \Lambda \to \mathbb{R}$ for some $\Lambda \subset \mathbb{Z}^d$. Write Ω_q for the collection of q-Lipschitz height functions. A measure is called *q-Lipschitz* if it is supported on Ω_q. A specification is called *q-Lipschitz* if it maps q-Lipschitz measures to q-Lipschitz measures. Finally, a function is called *strictly q-Lipschitz* if it is q_ε-Lipschitz for $\varepsilon > 0$ sufficiently small.

We now construct a number of objects which derive from q. These are necessary to state the main results, which address the macroscopic behaviour of Lipschitz surfaces.

Definition 3.3.3 (U_q, $\|\cdot\|_q$)**.** By a *slope* we simply mean an element u in the dual space $(\mathbb{R}^d)^*$ of $\mathbb{R}^d$. Write U_q for the interior of the set of slopes u such that $u|_{\mathcal{L}}$ is q-Lipschitz. The set U_q is nonempty and convex—this follows from the definition of a local Lipschitz constraint; see Lemma 3.6.1. Introduce furthermore the function $\|\cdot\|_q : \mathbb{R}^d \to \mathbb{R}$ defined by

$$\|x\|_q := \sup\{u(x) : u \in U_q\}.$$

The function $\|\cdot\|_q$ is positive homogeneous: we have $\|ax\|_q = a\|x\|_q$ for $a \in [0, \infty)$ and $x \in \mathbb{R}^d$. It also satisfies the triangle inequality, in the sense that $\|x + y\|_q \leq \|x\|_q + \|y\|_q$ for any $x, y \in \mathbb{R}^d$.

Definition 3.3.4 ($\|\cdot\|$-Lipschitz)**.** If $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ is any positive homogeneous function satisfying the triangle inequality, then any other function $f : D \to \mathbb{R}$ defined on a subset D of $\mathbb{R}^d$ is called *$\|\cdot\|$-Lipschitz* if $f(y) - f(x) \leq \|y - x\|$ for any $x, y \in D$. The function f is called *strictly $\|\cdot\|_q$-Lipschitz* if it is $\|\cdot\|_{q_\varepsilon}$-Lipschitz for some $\varepsilon > 0$. If D is open, then f is called *locally strictly $\|\cdot\|_q$-Lipschitz* if $f|_K$ is strictly $\|\cdot\|_q$-Lipschitz for all compact sets $K \subset D$.

For example, U_q is the interior of the set of slopes $u \in (\mathbb{R}^d)^*$ which are $\|\cdot\|_q$-Lipschitz.

3.3.2 Strong interactions

Let Ψ denote an arbitrary periodic gradient potential. The potential Ψ is called *positive* if $\Psi_\Lambda \geq 0$ for any $\Lambda \subset\subset \mathbb{Z}^d$. The potential Ψ is said to have *finite range* if $\Psi_\Lambda \equiv 0$ whenever the diameter of Λ—in the graph metric d_1 on the square lattice—exceeds some fixed constant $R \in \mathbb{N}$; in that case the smallest such R is called the *range* of Ψ. The potential Ψ is called *Lipschitz* if there exists a local Lipschitz constraint $(\mathbb{A}, q)$ such that $\Psi_\Lambda(\phi) = \infty$ if and only if $\Lambda = \{x, y\} \in \mathbb{A}$ and $\phi(y) - \phi(x) > q(x, y)$ for some $x, y \in \mathbb{Z}^d$. If $E = \mathbb{R}$ and Ψ Lipschitz with constraint $(\mathbb{A}, q)$, then Ψ is called *locally bounded* if for any $\varepsilon > 0$ sufficiently small, there exists a fixed constant $C_\varepsilon < \infty$, such that

$$H^\Psi_{\{x\}}(\phi) \leq C_\varepsilon$$

for any $x \in \mathbb{Z}^d$ and for any $\phi \in \Omega$ which is q_ε-Lipschitz at x.

Definition 3.3.5 (strong interaction, $\mathcal{S}_\mathcal{L}$)**.** A potential Ψ is called a *strong interaction* if Ψ has all of the above properties, that is, if Ψ is a positive Lipschitz periodic gradient potential of finite range, and if it is locally bounded in the case that $E = \mathbb{R}$. We shall write $\mathcal{S}_\mathcal{L}$ for the collection of strong interactions.

The class $\mathcal{S}_\mathcal{L}$ includes all so-called *Lipschitz simply attractive potentials*. These are convex Lipschitz nearest-neighbour interactions, see [54].

3.3.3 Weak interactions

Let Ξ denote an arbitrary periodic gradient potential.

Definition 3.3.6 (summability). The potential Ξ is called *summable* if it has finite norm

$$\|\Xi\| := \sup_{(x,\phi)\in\mathbb{Z}^d\times\Omega} \sum_{\Lambda\subset\subset\mathbb{Z}^d \text{ with } x\in\Lambda} |\Xi_\Lambda(\phi)|.$$

This requirement is significantly weaker than the *absolutely summable* setting of Georgii [20].

Definition 3.3.7 (amenability). By an *amenable function* we mean a function f which assigns a number in $[0,\infty)$ to each finite subset of $\mathbb{Z}^d$, such that:

1. $f(\Lambda)=f(\theta\Lambda)$ for all $\Lambda\subset\subset\mathbb{Z}^d$ and for any $\theta\in\Theta$,
2. $f(\Lambda\cup\Delta)\leq f(\Lambda)+f(\Delta)$ for all $\Lambda,\Delta\subset\subset\mathbb{Z}^d$ disjoint,
3. $f(\Lambda_n)=o(|\Lambda_n|)$ as $n\to\infty$ for any $(\Lambda_n)_{n\in\mathbb{N}}\uparrow\mathbb{Z}^d$.

Definition 3.3.8 (lower exterior bound). Let us now turn back to the potential Ξ and define, for any $\Lambda\subset\subset\mathbb{Z}^d$,

$$e^-(\Lambda) := \sup_{\phi\in\Omega} \sum_{\Delta\subset\subset\mathbb{Z}^d \text{ with } \Delta \text{ intersecting both } \Lambda \text{ and } \mathbb{Z}^d\smallsetminus\Lambda} |\Xi_\Delta(\phi)|.$$

The function $e^-(\cdot)$ is called the *lower exterior bound* of Ξ.

The key property of the function $e^-(\cdot)$ is that $|H_\Lambda^\Xi - H_\Lambda^{0,\Xi}|\leq e^-(\Lambda)$. The lower exterior bound satisfies Properties 1 and 2 from the definition of an amenable function; this is immediate from the definition.

Definition 3.3.9 (weak interaction, $\mathcal{W}_\mathcal{L}$). A *weak interaction* is a summable periodic gradient potential for which the lower exterior bound is amenable. Write $\mathcal{W}_\mathcal{L}$ for the collection of weak interactions.

It is straightforwardly verified that amenability of $e^-(\cdot)$ is equivalent to asking that $e^-(\Pi_n)=o(n^d)$ as $n\to\infty$. Remark that $(\mathcal{W}_\mathcal{L},\|\cdot\|)$ is a Banach space.

3.3.4 Overview

Let us fix a number of notations, in order to avoid an excessive number of declarations. We notify the reader of any deviation from this notation. We had already agreed that the choices for $d\geq 2$ and $E\in\{\mathbb{Z},\mathbb{R}\}$ are fixed, and that $\mathcal{L}$ denotes a fixed full-rank sublattice of $\mathbb{Z}^d$ with corresponding translation group Θ. The letter Φ denotes a fixed potential in $\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$, and we fix some pair $(\Psi,\Xi)\in\mathcal{S}_\mathcal{L}\times\mathcal{W}_\mathcal{L}$ such that $\Phi=\Psi+\Xi$. This decomposition is not unique, but this is never a problem. The specification generated by Φ is denoted $\gamma=\gamma^\Phi$. The pair $(\mathbb{A},q)$ always denotes the local Lipschitz constraint corresponding to Ψ, and the range of Ψ is denoted by R. If $E=\mathbb{Z}$, then q is always assumed to be integral. The function $e^-(\cdot)$ denotes the lower exterior bound of Ξ. Finally, let $K\in(0,\infty)$ denote the smallest constant such that $Kd_1\geq q$, and let $N\in\mathbb{N}$ denote the smallest positive integer such that $N\cdot\mathbb{Z}^d\subset\mathcal{L}$.

Definition 3.3.10. The potential $\Phi\in\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$ is called *monotone* if the induced specification $\gamma=\gamma^\Phi$ is monotone over Ω_q.

3.4 Main results

The motivation for writing this chapter was to demonstrate that the surface tension is strictly convex on U_Φ if the potential of interest is in the class $\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$ and monotone. If $E=\mathbb{Z}$, then we require an extra condition to be met, but we also demonstrate that this condition is satisfied for many natural models. This section contains an overview of the main results, including several results and applications which are of independent interest. The results are presented roughly in the order in which they appear in the chapter.

3.4.1 The specific free energy and its minimisers

The specific free energy functional plays a fundamental role in the analysis. The following result is therefore of independent interest in the study of Lipschitz random surfaces; it is a direct extension of a result of Sheffield [54] to the setting of this chapter.

Theorem 3.4.1 (specific free energy)**.** *If $\Phi \in \mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$, then the specific free energy functional*

$$\mathcal{H}(\cdot|\Phi):\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)\to\mathbb{R}\cup\{\infty\},\ \mu\mapsto\lim_{n\to\infty}n^{-d}\mathcal{H}_{\Pi_n}(\mu|\Phi)$$

is well-defined, affine, bounded below, lower-semicontinuous, and for each $C\in\mathbb{R}$ its lower level set

$$M_C:=\{\mu\in\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla):\mathcal{H}(\mu|\Phi)\leq C\}$$

is a compact Polish space, with respect to the topology of (weak) local convergence. In fact, the two topologies coincide on each set M_C.

A measure $\mu\in\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ is called a *minimiser of the specific free energy*, or simply a *minimiser*, if it satisfies the equation

$$\mathcal{H}(\mu|\Phi)=\sigma(S(\mu))<\infty.$$

For the purpose of deriving the main result, all that we require is that such minimisers have finite energy, in a sense which is similar to the notion of finite energy in the original paper of Burton and Keane [4]. There is a canonical way to translate the concept of finite energy to the gradient Lipschitz setting: we shall see that the following result fits our arguments. Recall that Ω_q denotes the set of q-Lipschitz height functions, and that π_Λ is the kernel which restrict measures to Λ, for any $\Lambda\subset\mathbb{Z}^d$.

Theorem 3.4.2 (finite energy)**.** *Consider $\Phi\in\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$, and suppose that $\mu\in\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ is a minimiser. Then for any $\Lambda\subset\subset\mathbb{Z}^d$, we have*

$$1_{\Omega_q}(\mu\pi_{\mathbb{Z}^d\smallsetminus\Lambda}\times\lambda^\Lambda)\ll\mu.$$

In [54], finite energy follows from the *variational principle*, which asserts that shift-invariant measures μ which satisfy $\mathcal{H}(\mu|\Phi)=\sigma(S(\mu))$ must also be Gibbs measures with respect to the specification $\gamma=\gamma^\Phi$ induced by the potential Φ—which has finite range. In the infinite-range setting one cannot hope for such a statement, because the specification γ is not necessarily quasilocal. This pathology, and its relation to the variational principle, is discussed extensively in Chapter 2. One of the key observations in that article is that minimisers of the specific free energy must have finite energy, even if the concept of a Gibbs measure is not well-defined because the specification fails to be quasilocal. There, finite energy is an immediate corollary of a result (Lemma 5.4) which is not quite equivalent to the variational principle, but it is "as close as one expects to get" to it in the non-quasilocal setting. We shall follow the same strategy here: the following theorem states the strongest result on minimisers of the specific free energy, implies directly that such minimisers have finite energy, and is a direct translate of Lemma 2.5.7 in Chapter 2 to the Lipschitz gradient setting. Let us first introduce the necessary definitions for the analysis of quasilocality.

Definition 3.4.3 (quasilocality, almost Gibbs measure)**.** Consider two finite sets $\Lambda \subset \Delta \subset\subset \mathbb{Z}^d$. Denote by $\mathcal{A}_{\Lambda,\Delta,\phi}$ the set of probability measures on $(E^\Lambda, \mathcal{E}^\Lambda)$ of the form $\mu\gamma_\Lambda\pi_\Lambda$, where μ is any measure in $\mathcal{P}(\Omega, \mathcal{F})$ subject only to $\mu\pi_\Delta = \delta_{\phi_\Delta}$. In other words, $\mathcal{A}_{\Lambda,\Delta,\phi}$ is the set of local Gibbs measures in Λ (and restricted to Λ) given (mixed) boundary conditions which match ϕ on Δ. Write $\mathcal{C}(\mathcal{A})$ for the closure of any $\mathcal{A} \subset \mathcal{P}(E^\Lambda, \mathcal{E}^\Lambda)$ in the strong topology, and define

$$\mathcal{A}_{\Lambda,\phi} := \cap_{\Delta\subset\subset\mathbb{Z}^d}\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi}).$$

A height function $\phi \in \Omega$ is called a *point of quasilocality* if $\mathcal{A}_{\Lambda,\phi} = \{\delta_\phi\gamma_\Lambda\pi_\Lambda\} = \{\gamma_\Lambda(\cdot,\phi)\pi_\Lambda\}$ for any $\Lambda \subset\subset \mathbb{Z}^d$. Write Ω_γ for the set of points of quasilocality. A measure $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ is called an *almost Gibbs measure* whenever $\mu(\Omega_\gamma) = 1$ and $\mu = \mu\gamma_\Lambda$ for any $\Lambda \subset\subset \mathbb{Z}^d$. The definition of an almost Gibbs measure is the same for gradient measures $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$—noting that $\Omega_\gamma \in \mathcal{F}^\nabla$ as γ is a gradient specification. Almost Gibbs measures are also called *Gibbs measures* whenever $\Omega_\gamma = \Omega$.

Let us now state the strongest result on minimisers, which is of independent interest.

Theorem 3.4.4 (minimisers of the specific free energy)**.** *Consider $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and suppose that $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ is a minimiser. Fix $\Lambda \subset\subset \mathbb{Z}^d$, and write μ^ϕ for the regular conditional probability distribution of μ on $(\Omega, \mathcal{F})$ corresponding to the projection map $\Omega \to E^{\mathbb{Z}^d\smallsetminus\Lambda}$. Then for μ-almost every $\phi \in \Omega$, we have $\mu^\phi\pi_\Lambda \in \mathcal{A}_{\Lambda,\phi}$. In particular, if $\mu(\Omega_\gamma) = 1$, then μ is an almost Gibbs measure, and if $\Omega_\gamma = \Omega$, then μ is a Gibbs measure.*

We shall furthermore demonstrate that in each of our applications, all minimisers are indeed (almost) Gibbs measures. We finally derive the following result.

Theorem 3.4.5 (existence of ergodic minimisers)**.** *Suppose that $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$. Then for any exposed point $u \in \bar{U}_\Phi$ of σ, there exists an ergodic gradient measure μ of slope u which is also a minimiser. In particular, if σ is strictly convex on U_Φ, then for each $u \in U_\Phi$, there is an ergodic minimiser of that slope.*

Theorem 3.4.1 is proven in Section 3.7. Theorems 3.4.2 and 3.4.4 are proven in Section 3.8. Theorem 3.4.5 is proven in Section 3.9.

3.4.2 Large deviations principle and variational principle

In Section 3.11 we prove a large deviations principle (LDP) of similar strength to the one stated in Chapter 7 of [54], with the noteworthy difference that we express it directly in terms of the Gibbs specification. This LDP captures both the macroscopic profile of each sample, as well as its local statistics. In this subsection however, we shall state a simpler LDP: one that captures only the macroscopic profile. By doing so we deliver on the premise that limit shapes are characterised by a variational principle, without spending many pages discussing the exact topology for the LDP with local statistics. However, the full LDP is also of independent interest, and we refer the interested reader to Subsection 3.11.1. Before stating the LDP, we must first describe how a sequence of discrete boundary conditions can approximate a continuous boundary profile, and we must also introduce a topology which captures the macroscopic profile of each sample. Let Φ denote a fixed potential throughout this subsection, and adopt the standard notation from Subsection 3.3.4.

Definition 3.4.6 (asymptotic boundary profile)**.** A *domain* is a nonempty bounded open subset of $\mathbb{R}^d$ such that its boundary has zero Lebesgue measure. An *asymptotic boundary profile* is a pair (D, b) where D is a domain and b a $\|\cdot\|_q$-Lipschitz function on ∂D. If $E = \mathbb{R}$, then call an asymptotic boundary profile (D, b) *good* if b is strictly $\|\cdot\|_q$-Lipschitz. If $E = \mathbb{Z}$, then call an asymptotic boundary profile *good* if it is *non-taut*. An asymptotic boundary profile (D, b) is called *non-taut* if b has an extension $\bar{b}$ to $\bar{D}$ such that $\bar{b}|_D$ is locally strictly $\|\cdot\|_q$-Lipschitz. This is equivalent to asking that the largest and smallest $\|\cdot\|_q$-Lipschitz extensions $b^\pm$ of b to $\bar{D}$ satisfy $b^- < b^+$ on D.

Definition 3.4.7 (discrete approximations)**.** Let (D, b) denote an asymptotic boundary profile. Call a sequence of pairs $(D_n, b_n)_{n\in\mathbb{N}}$ of finite subsets of $\mathbb{Z}^d$ and height functions an *approximation* of (D, b) if

1. For all $n \in \mathbb{N}$, the function b_n is q-Lipschitz if $E = \mathbb{Z}$ or strictly q-Lipschitz if $E = \mathbb{R}$,
2. We have $\frac{1}{n}D_n \to D$ in the Hausdorff metric on $\mathbb{R}^d$,
3. We have $\frac{1}{n}\operatorname{Graph}(b_n|_{\partial D_n}) \to \operatorname{Graph}(b)$ in the Hausdorff metric on $\mathbb{R}^d \times \mathbb{R}$.

Moreover, if $E = \mathbb{R}$, then an approximation $(D_n, b_n)_{n\in\mathbb{N}}$ is called *good* if the constant $\varepsilon > 0$ which makes each function b_n a q_ε-Lipschitz function, is independent of n. If $E = \mathbb{Z}$, then any approximation is called *good*.

We have in mind a good approximation $(D_n, b_n)_{n\in\mathbb{N}}$ of some fixed good asymptotic boundary profile (D, b). The sequence of local Gibbs measures which are of interest in the LDP is the sequence $(\gamma_n)_{n\in\mathbb{N}}$ defined by $\gamma_n := \gamma_{D_n}(\cdot, b_n)$. All samples from the sequence of measures $(\gamma_n)_{n\in\mathbb{N}}$ must be brought to the same topological space, in order for us to formulate the LDP. We will now describe this topology, as well as the map from Ω to this topological space.

Definition 3.4.8 (topology for macroscopic profiles)**.** For any $U \subset \mathbb{R}^d$, write $\operatorname{Lip}(U)$ for the set of real-valued $K\|\cdot\|_1$-Lipschitz functions on U, where we recall that K is minimal subject to $Kd_1 \geq q$. Suppose given a sample ϕ from γ_n. Define the scaled interpolation $\mathfrak{G}_n(\phi) \in \operatorname{Lip}(\bar{D})$ of ϕ, which captures the global shape of ϕ, as follows. The sample ϕ is almost surely q-Lipschitz, and therefore also Kd_1-Lipschitz. First, write $\bar{\phi} : \mathbb{R}^d \to \mathbb{R}$ for the smallest $K\|\cdot\|_1$-Lipschitz extension of ϕ to $\mathbb{R}^d$. Next, we simply scale back each sample by n and restrict it to the set $\bar{D}$. Formally, this means that we define

$$\mathfrak{G}_n(\phi) : \bar{D} \to \mathbb{R},\, x \mapsto \frac{1}{n}\bar{\phi}(nx).$$

This function is $K\|\cdot\|_1$-Lipschitz, that is, $\mathfrak{G}_n(\phi) \in \operatorname{Lip}(\bar{D})$. Endow the space $\operatorname{Lip}(\bar{D})$ with the topology of uniform convergence, denoted by $\mathcal{X}^\infty$. The map $\mathfrak{G}_n : \Omega \to \operatorname{Lip}(\bar{D})$ captures the global profile of the height functions in the large deviations principle.

Definition 3.4.9 (rate function, pressure)**.** The *rate function* associated to the profile (D, b) is the function $I : \operatorname{Lip}(\bar{D}) \to [0, \infty]$ defined by

$$I(f) := -P_\Phi(D, b) + \int_D \sigma(\nabla f(x))dx$$

if $f|_{\partial D} = b$ and $I(f) := \infty$ otherwise. Here $P_\Phi(D, b)$ is the *pressure* associated to this profile, which is defined precisely such that the minimum of I is zero.

Theorem 3.4.10 (large deviations principle)**.** *Let $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and let $(D_n, b_n)_{n\in\mathbb{N}}$ denote a good approximation of some good asymptotic profile (D, b). Let γ_n^* denote the pushforward of $\gamma_n := \gamma_{D_n}(\cdot, b_n)$ along the map $\mathfrak{G}_n$, for any $n \in \mathbb{N}$. Then the sequence of probability measures $(\gamma_n^*)_{n\in\mathbb{N}}$ satisfies a large deviations principle with speed n^d and rate function I on the topological space $(\operatorname{Lip}(\bar{D}), \mathcal{X}^\infty)$. Moreover, the sequence of normalising constants $(Z_n)_{n\in\mathbb{N}} := (Z_{D_n}(b_n))_{n\in\mathbb{N}}$ satisfies $-n^{-d} \log Z_n \to P_\Phi(D, g)$ as $n \to \infty$.*

Corollary 3.4.11 (variational principle)**.** *Let $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and let $(D_n, b_n)_{n\in\mathbb{N}}$ denote a good approximation of some good asymptotic profile (D, b). Let γ_n^* denote the pushforward of $\gamma_n := \gamma_{D_n}(\cdot, b_n)$ along the map $\mathfrak{G}_n$, for any $n \in \mathbb{N}$. Write f_n for the random function in γ_n^*, which—as a random object—takes values in $\operatorname{Lip}(\bar{D})$. If σ is strictly convex on U_Φ, then the random function f_n converges to the unique minimiser f^* of the rate function I, in probability in the topology of uniform convergence as $n \to \infty$. In other words, f^* is the unique minimiser of the integral*

$$\int_D \sigma(\nabla f(x))dx$$

over all Lipschitz functions $f : \bar{D} \to \mathbb{R}$ which equal b on the boundary of D. If however σ fails to be strictly convex on U_Φ, then for any neighbourhood A of the set of minimisers of the integral in the topology of uniform convergence, we have $f_n \in A$ with high probability as $n \to \infty$.

3.4.3 The surface tension

Let us now state the motivating result on the surface tension.

Theorem 3.4.12 (strict convexity of the surface tension)**.** *Let Φ denote a potential in $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ which is monotone.*

1. *If $E = \mathbb{R}$, then σ is strictly convex on U_Φ,*
2. *If $E = \mathbb{Z}$, then σ is strictly convex on U_Φ if for any affine map $h : (\mathbb{R}^d)^* \to \mathbb{R}$ with $h \leq \sigma$, the set $\{h = \sigma\} \cap \partial U_\Phi$ is convex. In particular, σ is strictly convex on U_Φ if at least one of the following conditions is satisfied:*
   - *(a) σ is affine on ∂U_Φ, but not on $\bar{U}_\Phi$,*
   - *(b) σ is not affine on $[u_1, u_2]$ for any distinct $u_1, u_2 \in \partial U_\Phi$ such that $[u_1, u_2] \not\subset \partial U_\Phi$.*

Strict convexity of the surface tension is important because of Theorem 3.4.5, Theorem 3.4.10, and Corollary 3.4.11. Let us also mention some other properties of the surface tension which are useful to keep in mind.

Theorem 3.4.13 (general properties of the surface tension)**.** *If $\Phi \in \mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$, then*

1. *We have $U_\Phi = U_q$,*
2. *If $E = \mathbb{R}$, then $\sigma(u)$ tends to ∞ as u approaches the boundary of U_Φ,*
3. *If $E = \mathbb{Z}$, then σ is bounded and continuous on the closure of U_Φ.*

Theorem 3.4.12 is proven in Section 3.12, and Theorem 3.4.13 is proven in Section 3.7.

3.4.4 Note on the Lipschitz setting

Local Lipschitz constraints are designed to be as flexible as possible. Essential in the argument is that a height function $\phi : \mathbb{Z}^d \to E$ has finite energy *if and only if* it is Lipschitz with respect to the local Lipschitz constraint. This means that we can rely on the Kirszbraun theorem (Theorem 3.6.4) to join together Lipschitz functions defined on disjoint parts of the space. However, this formulation is sometimes inconvenient. There are, as we shall see, several natural models in which the admissible height functions are exactly the *graph homomorphisms* from $\mathbb{Z}^d$ to $\mathbb{Z}$: these are functions $\phi : \mathbb{Z}^d \to \mathbb{Z}$ which satisfy $\phi(0) \in 2\mathbb{Z}$ and $|\phi(y) - \phi(x)| = 1$ for each edge $\{x, y\}$ of the square lattice. For example, the canonical height functions corresponding to the six-vertex model are precisely the graph homomorphisms from $\mathbb{Z}^2$ to $\mathbb{Z}$. Since the zero transition is not allowed, it might appear that this model does not fit the Lipschitz framework: it is the first *if* in the *if and only if* that is violated. However, this problem is only cosmetic in nature: by a simple transformation one can move from graph homomorphisms to the Lipschitz framework. Write $h : \mathbb{Z}^d \to \mathbb{Z}$ for the function $h(x) := \sum_i x_i$, and consider the map

$$\phi \mapsto (\phi + h)/2.$$

This map is a bijection from the set of graph homomorphisms to the set of functions which are q-Lipschitz for q defined by

$$q(x, y) := \sum\nolimits_i 0 \vee (y - x)_i.$$

By applying this transformation, it is thus clear that models of graph homomorphisms *do* fit into the local Lipschitz setting of this chapter. In fact, the exact same trick applies to dimer models, and perhaps other models of discrete height functions.

3.4.5 Application to submodular potentials

A potential Φ is said to be *submodular* if for every $\Lambda \subset\subset \mathbb{Z}^d$, Φ_Λ has the property that

$$\Phi_\Lambda(\phi \wedge \psi) + \Phi_\Lambda(\phi \vee \psi) \leq \Phi_\Lambda(\phi) + \Phi_\Lambda(\psi).$$

Sheffield proposes this family of potentials as a natural generalisation of simply attractive potentials, and asks if similar results as the ones proved for simply attractive potentials in [54] could be proved for finite-range submodular potentials. We provide an answer to this question for the case that the model is also Lipschitz. (In fact, we do not even require the potential to be finite-range.) It is easy to see that submodular potentials generate monotone specifications. If $E=\mathbb{R}$ and Φ a submodular Lipschitz potential fitting the framework of this thesis (which is a very mild requirement), then we derive immediately from Theorem 3.4.12 that the surface tension is strictly convex. If $E=\mathbb{Z}$, then we must also fulfill the extra condition in Theorem 3.4.12. We show that we can fulfill the extra condition if all shift-invariant measures μ which are supported on q-Lipschitz functions and which have $S(\mu)\in\partial U_\Phi$, are *frozen*, in the sense that for any $\Lambda \subset\subset \mathbb{Z}^d$, the values of ϕ_Λ depend deterministically on $\phi_{\partial^R\Lambda}$ in μ. This is a property of the local Lipschitz constraint q, and such local Lipschitz constraints are called *freezing*.

Theorem 3.4.14 (strict convexity for submodular potentials)**.** *Suppose that the potential $\Phi\in\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$ is submodular. Then it is monotone. Moreover,*

1. *If $E=\mathbb{R}$, then σ is strictly convex on U_Φ,*
2. *If $E=\mathbb{Z}$, then σ is strictly convex on U_Φ if the local Lipschitz constraint q is freezing.*

Note that q is automatically freezing if it is $\mathbb{Z}^d$-invariant.

Of course, Theorem 3.4.5 applies, and if the potential is finite-range, then the specification is quasilocal ($\Omega=\Omega_\gamma$) so that all minimisers are Gibbs measures (Theorem 3.4.4).

3.4.6 Application to tree-valued graph homomorphisms

The flexibility of the main theorem in this chapter can also be used to prove statements about the behaviour of random functions taking values in target spaces other than $\mathbb{Z}$ and $\mathbb{R}$. A noteworthy example is the model of tree-valued graph homomorphisms described in [44]. In this context, tree-valued graph homomorphisms are functions from $\mathbb{Z}^d$ to a k-regular tree $\mathcal{T}_k$ which also map the edges of the square lattice to the edges of the tree. Regular trees are natural objects in several fields of mathematics: in group theory, for example, they arise as Cayley graphs of free groups on finitely many generators. As a significant result in [44], the authors characterise the surface tension for the model (there named *entropy*) and show that it is equivalent to the number of graph homomorphisms with nearly-linear boundary conditions. This entropy function describes the macroscopic behaviour of the model, as is extensively discussed in [44]. We confirm the conjecture in [44], which asserts that this entropy function is strictly convex. We can do so because the model of uniformly random $\mathcal{T}_k$-valued graph homomorphisms can be translated into a model of $\mathbb{Z}$-valued graph homomorphisms after introducing an infinite-range interaction.

Let us now rigorously describe the conjecture which we prove is correct. Write U for the set of slopes $u\in(\mathbb{R}^d)^*$ such that $|u(e_i)|<1$ for each element e_i in the natural basis of $\mathbb{R}^d$. For fixed $u\in\bar{U}$, write $\phi^u:\mathbb{Z}^d\to\mathbb{Z}$ for the graph homomorphism defined by

$$\phi^u(x):=\lfloor u(x)\rfloor+\begin{cases}0 & \text{if } d_1(0,x)\equiv\lfloor u(x)\rfloor \mod 2,\\ 1 & \text{if } d_1(0,x)\equiv\lfloor u(x)\rfloor+1 \mod 2.\end{cases}$$

Then ϕ^u approximates u and it thus nearly linear, in the sense that $\|\phi^u-u|_{\mathbb{Z}^d}\|_\infty\leq 1$. Let g denote a *bi-infinite geodesic* through $\mathcal{T}_k$, that is, a $\mathbb{Z}$-indexed sequence of vertices

$g = (g_n)_{n\in\mathbb{Z}} \subset \mathcal{T}_k$ such that $d_{\mathcal{T}_k}(g_n, g_m) = |m - n|$ for any $n, m \in \mathbb{Z}$. The geodesic g is thought of as a copy of $\mathbb{Z}$ in $\mathcal{T}_k$, and is used as reference frame. Write $\tilde{\phi}^u : \mathbb{Z}^d \to \mathcal{T}_k$ for the graph homomorphism defined by $\tilde{\phi}^u(x) := g_{\phi^u(x)}$ for every $x \in \mathbb{Z}^d$. It is shown in [44] that the macroscopic behaviour of uniformly random $\mathcal{T}_k$-valued graph homomorphisms is characterised by the function

$$\mathrm{Ent} : \bar{U} \to [-\log k, 0],\ u \mapsto \lim_{n\to\infty} -n^{-d} \log |\{\tilde{\phi} \in \tilde{\Omega} : \tilde{\phi}_{\mathbb{Z}^d \smallsetminus \Pi_n} = \tilde{\phi}^u_{\mathbb{Z}^d \smallsetminus \Pi_n}\}|,$$

where $\tilde{\Omega}$ denotes the set of all graph homomorphisms from $\mathbb{Z}^d$ to $\mathcal{T}_k$. It is conjectured in [44] that Ent is strictly convex on U, which we prove is correct. Figure 3.2 displays a sample from the model; the limit shape is clearly visible.

Theorem 3.4.15 (strict convexity of the entropy for tree-valued graph homomorphisms)**.** *For any $d, k \geq 2$, the entropy function* $\mathrm{Ent} : \bar{U} \to [-\log k, 0]$ *associated to uniformly random graph homomorphisms from $\mathbb{Z}^d$ to a k-regular tree, is strictly convex on U.*

3.5 Moats

The following section is at the heart of this work. Its purpose is to show that for a specification which is stochastically monotone, two configurations sampled independently with the same boundary conditions are, on the scale of the specific free energy, at least as likely to oscillate a large number of times than to deviate from each other macroscopically. *Moats* are introduced in Definition 3.5.2 to formalise this statement. Informally, moats are clusters surrounding a given connected set, and on which the height difference between two configurations is prescribed between two fixed bounds. The proof relies crucially on the reflection principle which is stated in Lemma 3.5.1.

In this section, the implicit graph structure on $\mathbb{Z}^d$ is always the square lattice. As per usual, $(\mathbb{A}, q)$ denotes the local Lipschitz constraint, and $K \in (0, \infty)$ is chosen minimal subject to $Kd_1 \geq q$. We have in mind a gradient specification γ which is q-Lipschitz and monotone over Ω_q. From this specification we draw two height functions $\phi_1, \phi_2 \in \Omega$, and f shall generally denote the difference function $\phi_1 - \phi_2$, which is thus $2K$-Lipschitz.

3.5.1 Reflection principle

We first state and prove the reflection principle, which does not rely on the Lipschitz property. Throughout this section only, we shall adopt the following notation. Suppose that f_1 and f_2 are random functions in Ω, in some probability measures μ_1 and μ_2 respectively. Then write $f_1 \preceq f_2$ if f_1 is *stochastically dominated* by f_2, that is, $\mu_1(f_1 \in A) \leq \mu_2(f_2 \in A)$ for any increasing set $A \in \mathcal{F}$. Note that this notation still makes sense if $\mu_1 = \mu_2$, even if μ_1 and μ_2 are finite measures rather than probability measures.

Lemma 3.5.1 (Reflection principle)**.** *Let $\gamma = (\gamma_\Lambda)_{\Lambda\subset\subset\mathbb{Z}^d}$ denote a monotone gradient specification. Fix $\Lambda \subset\subset \mathbb{Z}^d$, and consider a probability measure μ on the product space $(\Omega^2, \mathcal{F}^2)$, writing (ϕ_1, ϕ_2) for the random pair of height functions, and with $f := \phi_1 - \phi_2$. Suppose that*

$$\mu = \mu(\gamma_\Lambda \times \gamma_\Lambda).$$

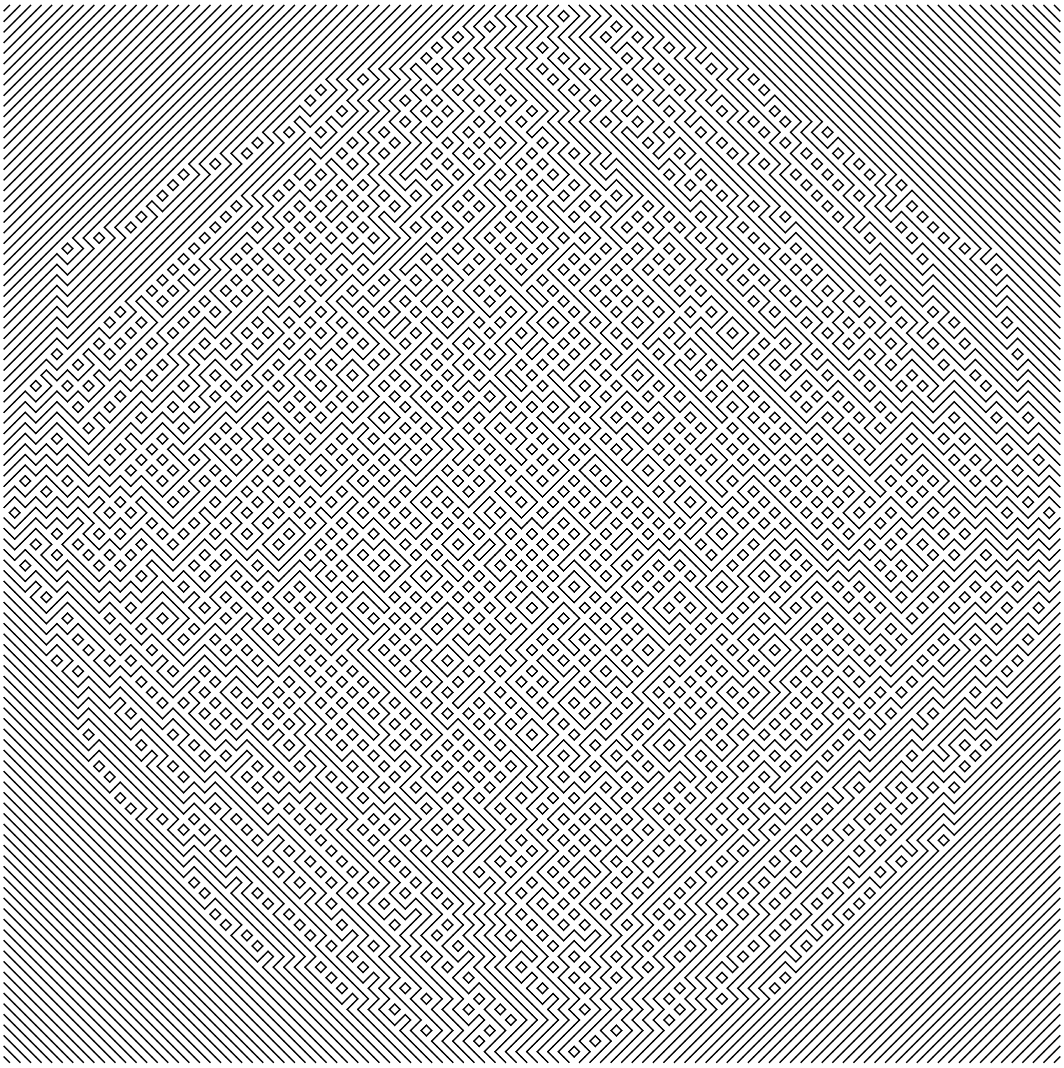


Figure 3.2: This figure shows the boundaries of the upper level sets of the horocylic height function (presented in Subsection 3.13.3) of a random $\mathcal{T}_3$-valued graph homomorphism. The boundary conditions resemble the Aztec diamond for domino tilings. The simulation hints at the presence of an arctic circle, alongside the limit shape which we prove appears inside.

If μ-almost surely $f_{\mathbb{Z}^d \smallsetminus \Lambda} \geq a$ for some $a \in \mathbb{R}$, then

$$-f \preceq f - 2a.$$

Similarly, if μ-almost surely $f_{\mathbb{Z}^d \smallsetminus \Lambda} \leq b$ for some $b \in \mathbb{R}$, then

$$f \preceq -f + 2b.$$

The same holds true if μ is a finite measure rather than a probability measure.

Proof. We focus on the first statement; the second statement then follows by symmetry. Fix $a \in \mathbb{R}$. Suppose first that μ restricted to $\mathbb{Z}^d \smallsetminus \Lambda$ is a Dirac measure, that is,

$$\mu = \gamma_\Lambda(\cdot, \psi_1) \times \gamma_\Lambda(\cdot, \psi_2)$$

for some $\psi_1, \psi_2 \in \Omega$ with $\psi_1 - \psi_2 \geq a$. As γ is a monotone gradient specification, we have

$$\phi_1 \succeq \phi_2 + a.$$

But ϕ_1 and ϕ_2 are independent, and therefore

$$-f = \phi_2 - \phi_1 \preceq (\phi_1 - a) - (\phi_2 + a) = f - 2a.$$

This inequality is generalised to the case that μ restricted to $\mathbb{Z}^d \smallsetminus \Lambda$ is not a Dirac measure, simply by averaging the inequality over all possible values of ϕ_1 and ϕ_2 on $\mathbb{Z}^d \smallsetminus \Lambda$ with respect to μ. □

3.5.2 Definition of moats

Definition 3.5.2 (Moats)**.** Let $f : \mathbb{Z}^d \to \mathbb{R}$ be a $2K$-Lipschitz function and $\Lambda \subset\subset \mathbb{Z}^d$ connected. Consider two real numbers a and b with $b - a \geq 4K$.

1. A set $M \subset \mathbb{Z}^d$ is a called an *a, b-moat of (f, Λ)* or simply a *moat* if M is a finite connected component of the set $\{a \leq f < b\} = \{x \in \mathbb{Z}^d : a \leq f(x) < b\} \subset \mathbb{Z}^d$ such that Λ is contained in a bounded connected component of $\mathbb{Z}^d \smallsetminus M$.

2. The boundary of M, that is, the set of vertices $x \in \mathbb{Z}^d \smallsetminus M$ adjacent to M, is denoted by ∂M. Write $\bar{M}$ for the closure of M, that is, $M \cup \partial M$.

3. The connected component of $\mathbb{Z}^d \smallsetminus M$ containing Λ is called the *inside* of M, and the *inside boundary* is the intersection of the inside with ∂M. Write M^Λ and $\partial_\Lambda M$ for the inside and the inside boundary respectively.

4. The unbounded connected component of $\mathbb{Z}^d \smallsetminus M$ is called the *outside* of M, and the *outside boundary* is the intersection of the outside with ∂M. Write M^∞ and $\partial_\infty M$ for the outside and the outside boundary respectively.

5. A moat M is said to *surround* another moat N, if $N \subset M^\Lambda$.

6. A moat M is called a *climbing moat* if $f_{\partial_\infty M} < a$ and $f_{\partial_\Lambda M} \geq b$ and it is called a *descending moat* if $f_{\partial_\infty M} \geq b$ and $f_{\partial_\Lambda M} < a$. From now on, we shall only consider moats which are either climbing or descending; when speaking of a moat, it is implicit that it belongs to one of these categories.

7. A finite sequence of moats $(M_k)_{1\leq k\leq n}$ is called *nested* if M_k surrounds M_{k+1} for all $1 \leq k < n$, and if the moats are alternatingly climbing and descending, with M_1 climbing.

We immediately collect a number of important properties.

Proposition 3.5.3. *Work in the context of the previous definition.*

1. *There exists at most one moat M with $x \in M$, for any fixed $x \in \mathbb{Z}^d$.*
2. *If M is a moat, then $a - 2K \leq f < b + 2K$ on $\bar{M}$.*
3. *Suppose that M is a moat, and that $p = (p_k)_{0\leq k\leq n} \subset \mathbb{Z}^d$ is a path through the square lattice from M^Λ to M^∞. Then $p_k \in M$ for at least $\lfloor (b-a)/2K \rfloor \geq 2$ consecutive integers k.*
4. *If $\Delta \subset\subset \mathbb{Z}^d$ contains Λ, then the number of moats M of (f, Λ) for which $M \cup M^\Lambda \subset \Delta$, is bounded by the d_1-distance from Λ to $\mathbb{Z}^d \smallsetminus \Delta$.*
5. *Suppose that M is a moat of (f, Λ), and that g is another $2K$-Lipschitz function with $g = f$ on $\bar{M}$. Then M is also a moat of (g, Λ). If M was climbing (resp. descending) w.r.t. (f, Λ) then it is climbing (resp. descending) w.r.t. (g, Λ). In other words, for $M \subset \mathbb{Z}^d$, the event*

   $$\{M \text{ is a (climbing or descending) moat of } (f, \Lambda)\}$$

   is $\mathcal{F}^2_{\bar{M}}$-measurable.
6. *Suppose that $A \subset \mathbb{Z}^d$ such that Λ is contained in a finite connected component of $\mathbb{Z}^d \smallsetminus A$, and write A^Λ for this connected component. If $f < a$ on A and $f \geq b$ on Λ, then A^Λ contains a climbing moat. If $f \geq b$ on A and $f < a$ on Λ, then A^Λ contains a descending moat.*
7. *If a' and b' are real numbers with $b' - a' \geq 4K$ and $[a', b'] \subset [a, b]$, then any a, b-moat contains an a', b'-moat.*

Proof. The first three statements follow from the definitions, where it is important that f is $2K$-Lipschitz and that any moat is either climbing or descending. For the fourth statement, observe that a path of minimal length from Λ to $\mathbb{Z}^d \smallsetminus \Delta$ through the square lattice must intersect any moat M for which $M \cup M^\Lambda \subset \Delta$. The fifth statement is immediate from the definition. The sixth statement follows from the connectivity properties of the square lattice, as well as the fact that f is $2K$-Lipschitz. The final statement is a corollary of the sixth. ☐

3.5.3 Moats and macroscopic deviations

Theorem 3.5.4. *Let $\gamma = (\gamma_\Lambda)_{\Lambda\subset\subset\mathbb{Z}^d}$ denote a q-Lipschitz gradient specification which is monotone over Ω_q. Fix $\Delta \subset\subset \mathbb{Z}^d$, and consider a q-Lipschitz probability measure μ on the product space $(\Omega^2, \mathcal{F}^2)$, writing (ϕ_1, ϕ_2) for the random pair of height functions, and with $f := \phi_1 - \phi_2$. Suppose that*

$$\mu = \mu(\gamma_\Delta \times \gamma_\Delta) \qquad \text{and} \qquad \mu\text{-almost surely } |f_{\mathbb{Z}^d \smallsetminus \Delta}| \leq 2K.$$

Fix a connected set $\Lambda \subset \Delta$, and write $E(n)$ for the event that there exists a sequence of n nested a,b-moats of (f,Λ), where $a = 4K$ and $b \geq 8K$. Then

$$m^{2n}\mu(E(2n)) \geq \mu(f_\Lambda \geq 3bn) \tag{3.5.5}$$

for all $n \in \mathbb{N}$, where $m = d_1(\Lambda, \mathbb{Z}^d \smallsetminus \Delta)$.

The idea of the proof is as follows. If $f \geq 3bn$ on Λ and $f \leq 2K$ on $\mathbb{Z}^d \smallsetminus \Delta$, then Δ must contain a climbing a,b-moat. Suppose now that we fix a subset M of Δ, and condition on the event

$$A := \{M \text{ is a climbing } a,b\text{-moat of } (f,\Lambda)\} \in \mathcal{F}^2_{\bar{M}}.$$

If we write Γ for the set $\Delta \smallsetminus \bar{M}$, then the conditioned measure $\mu(\cdot|A)$ satisfies

$$\mu(\cdot|A) = \mu(\cdot|A)(\gamma_\Gamma \times \gamma_\Gamma) \quad \text{ and } \quad \mu(\cdot|A)\text{-almost surely } -2K \leq f_{\mathbb{Z}^d \smallsetminus \Gamma} \leq b + 2K.$$

By the reflection principle, we thus have

$$\mu(f_\Lambda \leq -3bn + 2b + 4K|A) \geq \mu(f_\Lambda \geq 3bn|A).$$

In other words, this means that it is as least as likely to observe the set M as a climbing moat and a large negative deviation on Λ, than to see the set M as a climbing moat and a slightly larger positive deviation on Λ. But if f is negative on Λ then we can find a descending moat in the inside M^Λ of M. One repeats this reflection procedure to generate a full nested sequence of moats, while retaining a sufficiently large probability. The formalism is slightly more convoluted because one needs to choose the set M appropriately. This produces the extra factor m^{2n} in (3.5.5).

Proof of Theorem 3.5.4. We proceed along the same spirit. Write $\Delta_k := \{\Lambda \subset \Delta\}^k$, and define

$$A(M) := \{M \text{ is a nested sequence of } a,b\text{-moats of } (f,\Lambda)\} \in \mathcal{F}^2_{\cup_i \bar{M}_i}$$

for $M \in \Delta_k$. We also write $\Gamma(M) := \Delta \smallsetminus \cup_i \bar{M}_i$. Define $\mu_B := \mu(\cdot \cap B)$ for any $B \in \mathcal{F}^2$.

For any $k \in \mathbb{N}$ and $M \in \Delta_k$, we have

$$\mu_{A(M)} = \mu_{A(M)}(\gamma_{\Gamma(M)} \times \gamma_{\Gamma(M)}) \quad \text{ and } \quad \mu_{A(M)}\text{-a.e. } -2K \leq f_{\mathbb{Z}^d \smallsetminus \Gamma(M)} \leq b + 2K,$$

which means that the reflection principle applies to this measure. Claim that

$$\begin{aligned}
\mu(f_\Lambda \geq 3bn) &\leq \sum\nolimits_{M \in \Delta_1} \mu_{A(M)}(f_\Lambda \geq 3bn) && (3.5.6)\\
&\leq \sum\nolimits_{M \in \Delta_1} \mu_{A(M)}(f_\Lambda \leq -3bn + 2b + 4K) && (3.5.7)\\
&\leq \sum\nolimits_{M \in \Delta_2} \mu_{A(M)}(f_\Lambda \leq -3bn + 2b + 4K) && (3.5.8)\\
&\leq \sum\nolimits_{M \in \Delta_2} \mu_{A(M)}(f_\Lambda \geq 3bn - 2b - 8K) && (3.5.9)\\
&\leq \sum\nolimits_{M \in \Delta_2} \mu_{A(M)}(f_\Lambda \geq 3b(n-1)). && (3.5.10)
\end{aligned}$$

Here (3.5.6) follows from the fact that Δ contains a moat whenever $f \leq 2K$ on the complement of Δ and $f \geq 3bn$ on Λ, and (3.5.7) follows from the reflection principle applied to each measure in the finite sum. Now isolate one set $M \in \Delta_1$ and consider the measure $\mu_{A(M)}$. If $f_\Lambda \leq -3bn+2b+4K$, then there must be a descending moat in the inside of M_1—recall that M_1 is a climbing moat, by definition of a nested sequence of moats. In particular, this proves (3.5.8). Inequality (3.5.9) follows again from the reflection principle applied to each separate measure, and (3.5.10) follows from the fact that $3bn - 2b - 8K \geq 3b(n-1)$. A continuation of this series of inequalities leads to the equation

$$\mu(f_\Lambda \geq 3bn) \leq \sum_{M \in \Delta_{2n}} \mu_{A(M)}(f_\Lambda \geq 0).$$

The proof is nearly done. Note that $\mu_{A(M)}(\Omega^2) = \mu_{A(M)}(E(2n))$ for $M \in \Delta_{2n}$ by definition of $A(M)$ and $E(2n)$, and therefore

$$\mu(f_\Lambda \geq 3bn) \leq \sum_{M \in \Delta_{2n}} \mu_{A(M)}(f_\Lambda \geq 0) \leq \sum_{M \in \Delta_{2n}} \mu_{A(M)}(E(2n)).$$

To deduce (3.5.5), it suffices to demonstrate that, as measures,

$$\sum_{M \in \Delta_{2n}} \mu_{A(M)} \leq m^{2n}\mu.$$

The measure on the left equals $X\mu$, where X is the number of ways to choose a nested sequence of $2n$ moats contained in Δ. Since Δ contains at most m moats, we have $X \leq \binom{m}{2n} \leq m^{2n}$. □

We state an immediate corollary, which is an adaptation of the previous result to the case that Δ and Λ are not connected.

Proposition 3.5.11. *Assume the setting of the previous theorem, only suppose now that Δ and Λ each decompose into k connected components denoted by $(\Delta_i)_i$ and $(\Lambda_i)_i$ respectively with $\Lambda_i \subset \Delta_i$, and write $E(n)$ for the event that each Δ_i contains a sequence of n nested a, b-moats of (f, Λ_i). Then (3.5.5) holds true once we replace m by*

$$m = \prod_{i=1}^{k} d_1(\Lambda_i, \mathbb{Z}^d \smallsetminus \Delta_i).$$

3.6 Analysis of local Lipschitz constraints

This section contains several results on local Lipschitz constraints—most are deduced directly from Definition 3.3.1. Fix, throughout this section, a local Lipschitz constraint $(\mathbb{A}, q)$, and let $R \in \mathbb{N}$ denote a fixed constant such that $d_1(x, y) \leq R$ for all $\{x, y\} \in \mathbb{A}$. For example, one can take $(\mathbb{A}, q)$ to be the local Lipschitz constraint of Ψ, and R its range. These results are near-trivial for most commonly studied models; they require some work in the generality of Definition 3.3.1.

Throughout this section, we adopt the following notation. If $p = (p_k)_{0 \leq k \leq n}$ is a path through $(\mathbb{Z}^d, \mathbb{A})$, then we write $q(p)$ for $\sum_{k=1}^{n} q(p_{k-1}, p_k)$. If $q(p) = q(p_0, p_n)$, then p is called an *optimal path*.

3.6.1 Homogenisation of local Lipschitz constraints

The following lemma characterises U_q in terms of q. It also provides a relation between the local Lipschitz constraint q and the map $\|\cdot\|_q$ that it generates. The proof is similar to the proof in [54], although the formulation of the lemma is different.

Lemma 3.6.1. *The set U_q is nonempty. Its closure $\bar{U}_q$ can be written as the intersection of finitely many half-spaces. For each contributing half-space H, there exists a path $p = (p_k)_{0\leq k\leq n}$ through $(\mathbb{Z}^d, \mathbb{A})$ with $p_n - p_0 \in \mathcal{L}$ such that*

$$H = H(p) := \left\{u \in (\mathbb{R}^d)^* : u(p_n - p_0) \leq q(p)\right\}.$$

Moreover, there exists a constant $C < \infty$ such that

$$\|y - x\|_q - C \leq q(x, y) \leq \|y - x\|_q + C$$

for any $x, y \in \mathbb{Z}^d$.

Proof. Call some path $p = (p_k)_{0\leq k\leq n}$ through $(\mathbb{Z}^d, \mathbb{A})$ a *cycle lift* if the projection of p onto $\mathbb{Z}^d/\mathcal{L}$ is a cycle. Since $\mathbb{Z}^d/\mathcal{L}$ is finite and $(\mathbb{Z}^d, \mathbb{A})$ of bounded degree, there exist only finitely many cycle lifts once we identify paths which differ by a shift by a vector in $\mathcal{L}$.

Claim that

$$\{u \in (\mathbb{R}^d)^* : u|_{\mathcal{L}} \text{ is } q\text{-Lipschitz}\} = \cap_{p:\, p \text{ is a cycle lift}} H(p). \tag{3.6.2}$$

It is clear that the left set is contained in the right set. Focus now on the other containment. Fix a slope u in the set on the right. Suppose, for the sake of contradiction, that u is not in the set on the left, that is, that $u|_{\mathcal{L}}$ is not q-Lipschitz. Then there is some vertex $x \in \mathcal{L}$ and a path p from 0 to x through $(\mathbb{Z}^d, \mathbb{A})$ such that $u(x) > q(0, x) = q(p)$. But p decomposes into a finite collection of cycle lifts $(p^i)_i$. By choice of u, we have $u(x) \leq \sum_i q(p^i) = q(p)$, a contradiction. This proves the claim.

The set U_q equals the interior of the left and right in (3.6.2). Suppose that U_q is empty. Select a minimal family of cycle lifts $(p^k)_{1\leq k\leq m}$ such that the corresponding intersection of interiors of half-spaces $\cap_k \mathring{H}(p^k)$ is empty—by minimal we simply mean that m is as small as possible. For each $1 \leq k \leq m$, write $x^k \in \mathcal{L}$ for the endpoint of p^k minus p_0^k. Then each vector x^k is orthogonal to the affine hyperplane $\partial H(p^k)$. By Helly's theorem, we observe that $m \leq d + 1$. In fact, it is easy to see that, regardless of the value of m, the set $\{x^k : 1 \leq k \leq m\}$ is linearly dependent, with any strict subset linearly independent. It is a simple exercise in linear algebra to derive from the fact that the intersection of half-spaces $\cap_k \mathring{H}(p^k)$ is empty, that there is some slope u which is contained in the complement of $\mathring{H}(p^k)$ for any k, and that there exists a family of positive integers $(a^k)_{1\leq k\leq m} \subset \mathbb{N}$ such that $\sum_k a^k x^k = 0$. Since $u(x^k) \geq q(0, x^k)$ for each k by choice of u, we have

$$\sum\nolimits_k q(0, a^k x^k) \leq \sum\nolimits_k a^k q(0, x^k) \leq \sum\nolimits_k a^k u(x^k) = u(\textstyle\sum_k a^k x^k) = 0.$$

However, the triangle inequality and the inequality $q(x, y) + q(y, x) > 0$ for $x \neq y$ from the definition of an admissible quasimetric imply that

$$\begin{aligned}\sum\nolimits_{k=1}^m q(0, a^k x^k) &= q(0, a^1 x^1) + \sum\nolimits_{k=2}^m q(0, a^k x^k)\\ &\geq q(0, a^1 x^1) + q(0, -a^1 x^1) = q(0, a^1 x^1) + q(a^1 x^1, 0) > 0,\end{aligned}$$

a contradiction. This proves that U_q is nonempty.

Now let $x, y \in \mathbb{Z}^d$ arbitrary, and let p denote an optimal path from x to y. Then p decomposes into cycle lifts and at most $|\mathbb{Z}^d/\mathcal{L}| - 1$ remaining edges. It is straightforward to derive from this decomposition that the difference between $q(x, y)$ and $\|y - x\|_q$ is bounded uniformly over the choice of x and y. $\square$

Let us also state the following result, which follows immediately from the definition of $\|\cdot\|_q$ in terms of q.

Proposition 3.6.3. *If $f : D \to \mathbb{R}$ is $\|\cdot\|_q$-Lipschitz for $D \subset \mathbb{R}^d$, then $f|_{D\cap\mathcal{L}}$ is q-Lipschitz. If furthermore q is integral, then $\lfloor f \rfloor|_{D\cap\mathcal{L}}$ is also q-Lipschitz.*

3.6.2 General observations

First state the Kirszbraun theorem: this is an elementary result in the theory of Lipschitz functions. It asserts that a Lipschitz function defined on part of the space can be extended to a Lipschitz function on the entire space, with the same Lipschitz constant.

Proposition 3.6.4 (Kirszbraun theorem)**.** *If $\Lambda \subset \mathbb{Z}^d$ is nonempty and if $\phi : \Lambda \to \mathbb{R}$ is q-Lipschitz, then the function*

$$\phi^* : \mathbb{Z}^d \to \mathbb{R}, x \mapsto \sup_{y\in\Lambda} \phi(y) - q(x, y)$$

is the unique smallest q-Lipschitz extension of ϕ to $\mathbb{Z}^d$. If ϕ and q are integral, then so is ϕ^*. Suppose that $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ is any positive homogeneous function satisfying the triangle inequality. If $D \subset \mathbb{R}^d$ is nonempty and if $f : D \to \mathbb{R}$ is $\|\cdot\|$-Lipschitz, then the function

$$f^* : \mathbb{R}^d \to \mathbb{R}, x \mapsto \sup_{y\in D} f(y) - \|y - x\|$$

is the unique smallest $\|\cdot\|$-Lipschitz extension of f to $\mathbb{R}^d$.

Next, we discuss the derived local Lipschitz constraint q_ε for ε sufficiently small. For example, if $\mathbb{A}$ is the edge set of the square lattice and $q = Kd_1$ for $K \in [0, \infty)$, then q_ε is well-defined for $\varepsilon \in [0, K)$, and $q_\varepsilon = (K - \varepsilon)d_1$ for such ε. For the more general case, we use a technical construction to understand the derived local Lipschitz constraint q_ε.

Proposition 3.6.5. *There exist constants $\eta > 0$ and $C < \infty$ such that for any $0 \le \varepsilon \le \eta$,*

1. *We have $\varepsilon d_1/R \le q - q_\varepsilon \le C\varepsilon d_1$,*
2. *We have $q_{\varepsilon+\varepsilon'} = (q_\varepsilon)_{\varepsilon'} = (q_{\varepsilon'})_\varepsilon$ for any $\varepsilon' \ge 0$ with $\varepsilon + \varepsilon' \le \eta$,*
3. *For any $\varepsilon' \ge 0$ with $\varepsilon + 2\varepsilon' \le \eta$, if $\phi, \psi : \Lambda \to \mathbb{R}$ are functions for some $\Lambda \subset \mathbb{Z}^d$ where ϕ is $q_{\varepsilon+2\varepsilon'}$-Lipschitz and $\|\phi - \psi\|_\infty \le \varepsilon'$, then ψ is q_ε-Lipschitz.*

Proof outline. Claim that there exists a uniform constant $C < \infty$ such that $n(p) \le Cd_1(x, y)$ for any optimal path p from x to y, where $n(p)$ denotes the length of that path. To see that the claim is true, observe that U_q is nonempty and open, and therefore there exists a constant $\alpha > 0$ such that $\|x\|_q + \|-x\|_q \ge \alpha\|x\|_1$ for any $x \in \mathbb{R}^d$. Moreover, the difference between $q(x, y)$ and $\|y - x\|_q$ is bounded uniformly over $x, y \in \mathbb{Z}^d$ (Lemma 3.6.1). It is straightforward to deduce the claim from these two facts.

One now defines the map $X_q : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{Z}_{\geq 0}$ by

$$X_q(x, y) := \max\{n(p) : p \text{ is an optimal path from } x \text{ to } y\}.$$

Then $d_1/R \leq X_q \leq Cd_1$ by the previous discussion. It is straightforward, but slightly technical, to see that $q_\varepsilon = q - \varepsilon X_q$ for ε sufficiently small. This implies the three statements of the proposition. $\square$

Proposition 3.6.6. *We have $U_q = \cup_{\varepsilon>0} \bar{U}_{q_\varepsilon}$.*

3.6.3 Approximation of continuous profiles

Recall that $\Lambda^{-m}(D) := (\mathbb{Z}^d \cap D) \smallsetminus \partial^m(\mathbb{Z}^d \cap D)$ for any $m \in \mathbb{Z}_{\geq 0}$ and $D \subset \mathbb{R}^d$.

Theorem 3.6.7. *Consider $\varepsilon > 0$ sufficiently small so that q_ε is well-defined, and fix $C < \infty$. Then there is a constant $m \in \mathbb{Z}_{\geq 0}$ such that the following statement holds true. Suppose given a collection $(D_i)_i$ of disjoint subsets of $\mathbb{R}^d$, and write $D := \cup_i D_i$. Let $f : D \to \mathbb{R}$ denote a $\|\cdot\|_q$-Lipschitz function such that $f(y) - f(x) \leq \|y - x\|_{q_\varepsilon}$ for $x \in D_i$ and $y \in D_j$ with $i \neq j$. Define $\Lambda_i := \Lambda^{-m}(D_i)$ and $\Lambda := \cup_i \Lambda_i$. Let $\phi : \Lambda \to E$ denote a function such that ϕ_{Λ_i} is q-Lipschitz for all i and with $|\phi - f|_\Lambda| \leq C$. Then ϕ is q-Lipschitz, and has a q-Lipschitz extension to $\mathbb{Z}^d$.*

Proof. This follows from Lemma 3.6.1 and Proposition 3.6.5. $\square$

In the remainder of this section, we specialise to the case that $(\mathbb{A}, q)$ is the local Lipschitz constraint associated to the strong interaction Ψ as described in Subsection 3.3.4. The previous theorem is particularly useful in the case that the function f is affine on each set D_i, say with slope $u_i \in U_q$. In that case, we want the height function ϕ to approximate the slope u_i on each set Λ_i. To this end we will choose for each $u \in U_q$ a canonical Lipschitz height function ϕ^u to represent that slope u. This is the purpose of the following definition.

Definition 3.6.8. Consider some fixed slope $u \in U_q$. If $E = \mathbb{Z}$, then write $\phi^u \in \Omega$ for the unique smallest q-Lipschitz extension of the function $\lfloor u \rfloor|_{\mathcal{L}}$ to $\mathbb{Z}^d$. If $E = \mathbb{R}$, then write $\phi^u \in \Omega$ for the unique smallest q_ε-Lipschitz extension of $u|_{\mathcal{L}}$ to $\mathbb{Z}^d$, where ε is the largest positive real number such that $u|_{\mathcal{L}}$ is q_ε-Lipschitz (subject to $\varepsilon \leq \eta$, where η is as in Proposition 3.6.5).

If $E = \mathbb{Z}$, then q is integral, and therefore the smallest q-Lipschitz extension of $\lfloor u \rfloor|_{\mathcal{L}}$ to $\mathbb{Z}^d$ is also integer-valued. The rounding procedure in the discrete setting makes that the gradient of ϕ^u is not $\mathcal{L}$-invariant. In the continuous setting $E = \mathbb{R}$ there is no rounding, and therefore the gradient of ϕ^u is $\mathcal{L}$-invariant. Finally, we want to remark that, in both the discrete and the continuous setting, there exists a constant $C < \infty$ such that $|\phi^u - u|_{\mathbb{Z}^d}| \leq C$ for any $u \in U_q$. This is due to Lemma 3.6.1. This observation, combined with the previous theorem, implies the following result.

Theorem 3.6.9. *Let $C < \infty$ denote the smallest constant such that $|\phi^u - u|_{\mathbb{Z}^d}| + 1 \leq C$ for all $u \in U_q$. Consider $\varepsilon > 0$ so small that q_ε is well-defined. Then there exists a constant $m \in \mathbb{Z}_{\geq 0}$ such that the following holds true. Suppose given a collection $(D_i)_i$ of disjoint subsets of $\mathbb{R}^d$, and write $D := \cup_i D_i$, $\Lambda_i := \Lambda^{-m}(D_i)$, and $\Lambda := \cup_i \Lambda_i$. Let $f : D \to \mathbb{R}$ denote a $\|\cdot\|_{q_\varepsilon}$-Lipschitz function which is affine with slope $u_i \in \bar{U}_{q_\varepsilon}$*

whenever restricted to D_i. Then there exists a q-Lipschitz function $\phi : \Lambda \to E$ which satisfies $|\phi - f|_\Lambda| \leq C$ and $\nabla\phi|_{\Lambda_i} = \nabla\phi^{u_i}|_{\Lambda_i}$ for all i. If $E = \mathbb{R}$ then we may furthermore impose that ϕ is $q_{\varepsilon'}$-Lipschitz for fixed $0 < \varepsilon' < \varepsilon$ (that m is allowed to depend upon).

For this result, the notation $\nabla\phi = \nabla\psi$ means that the difference $\phi - \psi$ is constant.

3.7 The specific free energy

3.7.1 The attachment lemmas

The letter Φ denotes a potential in $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ throughout this section. For the thermodynamical formalism, it is crucial that we are able to attach height functions defined on disjoint subsets of $\mathbb{Z}^d$ without losing or gaining too much energy. More precisely, if $\Lambda_1, \Lambda_2 \subset\subset \mathbb{Z}^d$ are disjoint with $\Lambda := \Lambda_1 \cup \Lambda_2$, then we want to find bounds on the difference between $H^0_\Lambda(\phi)$ and $H^0_{\Lambda_1}(\phi) + H^0_{\Lambda_2}(\phi)$. Similarly, we will require bounds on the difference between $H^0_\Lambda(\phi)$ and $H_\Lambda(\phi)$. In this section, we present simple tools for doing this: the attachment lemmas. We first state and prove the lower attachment lemma, which is easier.

Lemma 3.7.1 (Lower attachment lemma)**.** *Let $\Lambda_1, \Lambda_2 \subset\subset \mathbb{Z}^d$ disjoint, and write $\Lambda := \Lambda_1 \cup \Lambda_2$. Then*

$$H^0_\Lambda \geq H^0_{\Lambda_1} + H^0_{\Lambda_2} - \min_{i \in \{1,2\}} e^-(\Lambda_i),$$

where e^- is the lower exterior bound of Ξ. We also have $H_\Lambda \geq H^0_\Lambda - e^-(\Lambda)$ for any $\Lambda \subset\subset \mathbb{Z}^d$.

Proof. The inequality $H^{0,\Psi}_\Lambda \geq H^{0,\Psi}_{\Lambda_1} + H^{0,\Psi}_{\Lambda_2}$ is obvious because Ψ is positive. The inequality $H^{0,\Xi}_\Lambda \geq H^{0,\Xi}_{\Lambda_1} + H^{0,\Xi}_{\Lambda_2} - \min_{i\in\{1,2\}} e^-(\Lambda_i)$ is immediate from the definition of e^- in terms of Ξ. This proves the inequality in the display. The other inequality follows from a similar decomposition. $\square$

More care is required for the upper bound. There is a difference between the discrete case $E = \mathbb{Z}$ and the continuous case $E = \mathbb{R}$. If $E = \mathbb{Z}$ then the strong interaction Ψ can be described by finite information. The effect of this is that there exists a uniform bound $C < \infty$ such that

$$H^\Psi_{\{x\}}(\phi) \leq C \tag{3.7.2}$$

for any $x \in \mathbb{Z}^d$ and any q-Lipschitz function $\phi \in \Omega$. If $E = \mathbb{R}$ then there exists no such *a priori* bound, and it is this specific reason reason that we introduce the *locally bounded* property in Subsection 3.3.2, so that at least

$$H^\Psi_{\{x\}}(\phi) \leq C_\varepsilon \tag{3.7.3}$$

whenever ϕ is q_ε-Lipschitz at x.

For the upper bound, one requires control especially over the potential Ψ which enforces the Lipschitz constraint. The height function ϕ must therefore be sufficiently well-behaved for the lemma to work, at least on the boundary where Λ_1 meets Λ_2.

Lemma 3.7.4 (Upper attachment lemma)**.** *Let $\phi \in \Omega$ and $\Lambda_1, \Lambda_2 \subset\subset \mathbb{Z}^d$ disjoint, and write $\Lambda := \Lambda_1 \cup \Lambda_2$. If $E = \mathbb{Z}$, then there exists an amenable function e^+, dependent only on Φ, such that*

$$H^0_\Lambda(\phi) \leq H^0_{\Lambda_1}(\phi) + H^0_{\Lambda_2}(\phi) + \min_{i\in\{1,2\}} e^+(\Lambda_i) \tag{3.7.5}$$

whenever $\phi_{\partial^R\Lambda_1 \cup \partial^R\Lambda_2}$ is q-Lipschitz, and such that

$$H_\Lambda(\phi) \leq H^0_\Lambda(\phi) + e^+(\Lambda) \tag{3.7.6}$$

whenever $\phi_{\partial^R\Lambda \cup \partial^R(\mathbb{Z}^d \smallsetminus \Lambda)}$ is q-Lipschitz. If $E = \mathbb{R}$ and $\varepsilon > 0$, then there exists an amenable function e^+_ε, dependent only on Φ and ε, such that (3.7.5) and (3.7.6) hold true whenever the restrictions of ϕ are q_ε-Lipschitz, and with e^+ replaced by e^+_ε.

Definition 3.7.7. The functions e^+ and e^+_ε are called *upper exterior bounds.*

Proof of Lemma 3.7.4. It suffices to consider the contributions of the potentials Ψ and Ξ to each Hamiltonian separately; one can simply sum the two upper exterior bounds $e^{+,\Psi}$ and $e^{+,\Xi}$ so obtained. In fact, the upper exterior bound $e^{+,\Xi} := e^-$ suffices for the long-range interaction Ξ. Let us therefore focus on the contribution from the potential Ψ.

We shall simultaneously consider the discrete case and the continuous case. In this proof we shall reserve the name *Lipschitz* for *q-Lipschitz* whenever $E = \mathbb{Z}$ and for *q_ε-Lipschitz* whenever $E = \mathbb{R}$. Write C for a fixed constant such that $H^\Psi_{\{x\}}(\psi) \leq C$ for any $x \in \mathbb{Z}^d$ and for any Lipschitz height function ψ. Because Ψ is positive and of range R and because the restriction of ϕ to $\partial^R\Lambda_1 \cup \partial^R\Lambda_2$ is Lipschitz, we have

$$\begin{aligned} H^{0,\Psi}_\Lambda(\phi) - H^{0,\Psi}_{\Lambda_1}(\phi) - H^{0,\Psi}_{\Lambda_2}(\phi) &= \sum_{\Delta\subset\Lambda,\, \Delta\not\subset\Lambda_1,\, \Delta\not\subset\Lambda_2} \Psi_\Delta(\phi) \\ &= \sum_{\Delta\subset\partial^R\Lambda_1\cup\partial^R\Lambda_2,\, \Delta\not\subset\Lambda_1,\, \Delta\not\subset\Lambda_2} \Psi_\Delta(\phi) \leq \min_{i\in\{1,2\}} H^\Psi_{\partial^R\Lambda_i, \partial^R\Lambda_1\cup\partial^R\Lambda_2}(\phi) \\ &\leq \min_{i\in\{1,2\}} C|\partial^R\Lambda_i| \leq \min_{i\in\{1,2\}} e^{+,\Psi}(\Lambda_i) \end{aligned}$$

if we define $e^{+,\Psi}(\Lambda) := C(2R+1)^d|\partial\Lambda|$; this function satisfies the desired constraints. It is clear that this choice for $e^{+,\Psi}$ also implies that

$$H^\Psi_\Lambda(\phi) \leq H^{0,\Psi}_\Lambda(\phi) + e^{+,\Psi}(\Lambda)$$

whenever the restriction of ϕ to $\partial^R\Lambda \cup \partial^R(\mathbb{Z}^d \smallsetminus \Lambda)$ is Lipschitz. □

3.7.2 Density limits of functions on finite subsets of $\mathbb{Z}^d$

Proposition 3.7.8. *Consider two $\mathcal{L}$-invariant real-valued functions f and b on the finite subsets of $\mathbb{Z}^d$, with b amenable and*

$$f(\Lambda_1 \cup \Lambda_2) \leq f(\Lambda_1) + f(\Lambda_2) + \min_{i\in\{1,2\}} b(\Lambda_i) \tag{3.7.9}$$

for disjoint $\Lambda_1, \Lambda_2 \subset\subset \mathbb{Z}^d$. Then $(n^{-d}f(\Pi_n))_{n\in\mathbb{N}}$ tends to a limit in $[-\infty, \infty)$ as $n \to \infty$, and

$$\lim_{n\to\infty} n^{-d} f(\Pi_n) = \inf_{n\in N\cdot\mathbb{N}} n^{-d}(f(\Pi_n) + b(\Pi_n))$$

where $N \in \mathbb{N}$ is minimal subject to $N \cdot \mathbb{Z}^d \subset \mathcal{L}$. Finally, if $(\Lambda_n)_{n\in\mathbb{N}} \uparrow \mathbb{Z}^d$, then

$$\limsup_{n\to\infty} |\Lambda_n|^{-1} f(\Lambda_n) \leq \lim_{n\to\infty} n^{-d} f(\Pi_n).$$

If we weaken the assumptions, and suppose only that (3.7.9) holds true whenever Λ_1 contains some vertex x adjacent to some vertex y in Λ_2, then each statement in this proposition remains valid, except that, for the final assertion, we also require that each set Λ_n is connected.

Definition 3.7.10. Write $\langle \cdot | \Phi \rangle : \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla) \to [-\|\Xi\|, \infty]$ for the unique functional which satisfies

$$\langle \mu | \Phi \rangle := \mu(\Phi) := \lim_{n\to\infty} n^{-d} \mu(H^0_{\Pi_n}).$$

The limit on the right converges due to the lower attachment lemma and the previous proposition. This quantity is called the *specific energy* of μ with respect to Φ.

3.7.3 Free energy attachment lemma

Definition 3.7.11. Define $e^* := e^- + \log(2K+1)$, where K is minimal subject to $Kd_1 \geq q$. Call the amenable function e^* the *free energy exterior bound*.

Lemma 3.7.12 (Free energy attachment lemma)**.** *Fix $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$, and consider some disjoint sets $\Lambda_1, \Lambda_2 \subset\subset \mathbb{Z}^d$ with some vertex x of Λ_1 adjacent to some vertex y of Λ_2 in the square lattice. Write $\Lambda := \Lambda_1 \cup \Lambda_2$. Then*

$$\mathcal{H}_\Lambda(\mu|\Phi) \geq \mathcal{H}_{\Lambda_1}(\mu|\Phi) + \mathcal{H}_{\Lambda_2}(\mu|\Phi) - \min_{i\in\{1,2\}} e^*(\Lambda_i).$$

Moreover, for $\Lambda \subset\subset \mathbb{Z}^d$ connected and nonempty, we have

$$\mathcal{H}_\Lambda(\mu|\Phi) \geq -(|\Lambda| - 1) \max_{x\in\mathbb{Z}^d/\mathcal{L}} e^*(\{x\}).$$

Proof. Fix K minimal subject to $Kd_1 \geq q$. Recall that $\mu\pi_\Lambda$ is the restriction of μ to Λ. We assume that $\mu\pi_\Lambda$ is supported on Kd_1-Lipschitz functions; if this is not the case, then $\mathcal{H}_\Lambda(\mu|\Phi)$ is infinite, and we are done. For any $\Delta \subset \Lambda$, we have

$$\mathcal{H}_\Delta(\mu|\Phi) = \mathcal{H}_{\mathcal{F}^\nabla_\Delta}(\mu|\lambda^{\Delta-1}) + \mu(H^0_\Delta).$$

Observe that $\mu(H^0_\Lambda) \geq \mu(H^0_{\Lambda_1}) + \mu(H^0_{\Lambda_2}) - \min_{i\in\{1,2\}} e^-(\Lambda_i)$ due to the lower attachment lemma. Therefore it suffices to show that

$$\mathcal{H}_{\mathcal{F}^\nabla_\Lambda}(\mu|\lambda^{\Lambda-1}) \geq \mathcal{H}_{\mathcal{F}^\nabla_{\Lambda_1}}(\mu|\lambda^{\Lambda_1-1}) + \mathcal{H}_{\mathcal{F}^\nabla_{\Lambda_2}}(\mu|\lambda^{\Lambda_2-1}) - \log(2K+1) \quad (3.7.13)$$

whenever $\mu\pi_\Lambda$ is supported on Kd_1-Lipschitz functions. This follows from the following two facts:

1. We have $\mathcal{H}_{\mathcal{F}^\nabla_{\{x,y\}}}(\mu|\lambda^{\{x,y\}-1}) \geq -\log(2K+1)$,

2. If $\Delta_1, \Delta_2 \subset \Lambda$ share a single vertex z and $\Delta := \Delta_1 \cup \Delta_2$, then

$$\mathcal{H}_{\mathcal{F}^\nabla_\Delta}(\mu|\lambda^{\Delta-1}) \geq \mathcal{H}_{\mathcal{F}^\nabla_{\Delta_1}}(\mu|\lambda^{\Delta_1-1}) + \mathcal{H}_{\mathcal{F}^\nabla_{\Delta_2}}(\mu|\lambda^{\Delta_2-1}).$$

Note that (3.7.13) then follows by applying the second fact twice, first to the sets Λ_1 and $\{x,y\}$, then to the sets $\Lambda_1 \cup \{y\}$ and Λ_2. Let us first prove the first fact. Since μ is supported on Kd_1-Lipschitz functions, we have

$$\mathcal{H}_{\mathcal{F}^\nabla_{\{x,y\}}}(\mu|\lambda^{\{x,y\}-1}) \geq -\log \lambda^{\{x,y\}-1}(\{|\phi(y)-\phi(x)| \leq K\}) \geq -\log(2K+1).$$

For the second fact, we can simply choose the point z as a reference point for all gradient measures, such that the measurable space $(\Omega, \mathcal{F}^\nabla_\Delta)$ becomes effectively a product space; the measure $\lambda^{\Delta-1}$ is then the product measure of λ^{Δ_1-1} and λ^{Δ_2-1}. The second fact now follows; the inequality in the display is well-known for product spaces.

The final assertion of the lemma is a direct consequence of the first assertion and the fact that $\mathcal{H}_\Lambda(\mu|\Phi) = 0$ whenever Λ is a singleton. □

3.7.4 Convergence and properties of the specific free energy

The two results in this subsection jointly imply Theorem 3.4.1.

Theorem 3.7.14. *If $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, then the functional $\mathcal{H}(\cdot|\Phi) : \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla) \to \mathbb{R} \cup \{\infty\}$ is well-defined and satisfies*

$$\mathcal{H}(\mu|\Phi) := \lim_{n\to\infty} n^{-d}\mathcal{H}_{\Pi_n}(\mu|\Phi)$$
$$= \sup_{n \in N\cdot\mathbb{N}} n^{-d}\left(\mathcal{H}_{\Pi_n}(\mu|\Phi) - e^*(\Pi_n)\right) \geq -\max_{x\in\mathbb{Z}^d/\mathcal{L}} e^*(\{x\}),$$

where N is minimal subject to $N \cdot \mathbb{Z}^d \subset \mathcal{L}$. Moreover, $\mathcal{H}(\cdot|\Phi)$ is lower-semicontinuous, and for each $C \in \mathbb{R}$ the lower level set

$$M_C := \{\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla) : \mathcal{H}(\mu|\Phi) \leq C\}$$

is a compact Polish space, with respect to the topology of (weak) local convergence. In fact, the two topologies coincide on each set M_C.

Proof. The statements in the first display follow from Lemma 3.7.12 and Proposition 3.7.8. For the remainder of the theorem, observe that

$$M_C = \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla) \cap \bigcap_{n\in N\cdot\mathbb{N}} \{\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla) : \mathcal{H}_{\Pi_n}(\mu|\Phi) \leq n^d C + e^*(\Pi_n)\}.$$

Each of these sets is closed (in the topology of weak local convergence), and therefore M_C is closed; the functional $\mathcal{H}(\cdot|\Phi)$ must be lower-semicontinuous (in either topology). Moreover, for each $n \in N \cdot \mathbb{N}$, the set

$$\{\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla_{\Pi_n}) : \mathcal{H}_{\Pi_n}(\mu|\Phi) \leq n^d C + e^*(\Pi_n)\}$$

is a compact Polish space with respect to both the weak and strong topologies, which coincide on this set. Write δ_n for the corresponding metric. Then M_C is a compact Polish space with metric $\delta(\mu,\nu) := \sum_{n\in N\cdot\mathbb{N}} e^{-n}(\delta_n(\mu,\nu) \wedge 1)$. □

Theorem 3.7.15. *If $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, then the functional $\mathcal{H}(\cdot|\Phi)$ is affine, in the sense that*

$$\mathcal{H}((1-t)\mu + t\nu|\Phi) = (1-t)\mathcal{H}(\mu|\Phi) + t\mathcal{H}(\nu|\Phi)$$

for $\mu, \nu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ and $0 \leq t \leq 1$.

Proof. It follows from a direct entropy calculation that for fixed $\Lambda \subset\subset \mathbb{Z}^d$,

$$0 \leq (1-t)\mathcal{H}_\Lambda(\mu|\Phi) + t\mathcal{H}_\Lambda(\nu|\Phi) - \mathcal{H}_\Lambda((1-t)\mu + t\nu|\Phi) \leq 2\log 2.$$

This error term vanishes in the normalisation of the specific free energy. □

3.7.5 The surface tension

Recall that the surface tension $\sigma : (\mathbb{R}^d)^* \to \mathbb{R} \cup \{\infty\}$ is defined by

$$\sigma(u) := \inf_{\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla}) \text{ with } S(\mu) = u} \mathcal{H}(\mu|\Phi).$$

The function σ must be convex because both $S(\cdot)$ and $\mathcal{H}(\cdot|\Phi)$ are affine. It is also bounded from below because $\mathcal{H}(\cdot|\Phi)$ is bounded from below by $-\max_{x\in\mathbb{Z}^d/\mathcal{L}} e^*(\{x\})$. Recall that U_Φ is defined to be the interior of the set $\{\sigma < \infty\} \subset (\mathbb{R}^d)^*$. The set U_Φ is convex, and σ is continuous on U_Φ. Moreover, σ must equal ∞ on the complement of the closure of U_Φ. Recall the statement of Theorem 3.4.13, for which we now provide a proof.

Proof of Theorem 3.4.13. Observe that σ is lower-semicontinuous, because $S(\cdot)$ is continuous and because $\mathcal{H}(\cdot|\Phi)$ is lower-semicontinuous with compact lower level sets.

Let us first prove that $U_\Phi \subset U_q$. Suppose that the slope $u := S(\mu)$ of $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ is not in $\bar{U}_q$. It suffices to demonstrate that $\mathcal{H}(\mu|\Phi) = \infty$. Since $u \notin \bar{U}_q$, we know that $u|_{\mathcal{L}}$ is not q-Lipschitz, and therefore with positive μ-probability, $\phi|_{\mathcal{L}}$ is not q-Lipschitz. In particular, this means that $\mu(H^0_{\Pi_n}) = \infty$ for n sufficiently large. This proves that $\mathcal{H}(\mu|\Phi) = \infty$.

For the remainder of the proof, we distinguish between the discrete and the continuous setting. Consider first the case that $E = \mathbb{Z}$. For the lemma, it suffices to demonstrate that σ is bounded on U_q. If $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ is supported on q-Lipschitz functions, then

$$\mathcal{H}_{\Pi_n}(\mu|\Phi) = \mu(H^0_{\Pi_n}) + \mathcal{H}_{\mathcal{F}^{\nabla}_{\Pi_n}}(\mu|\lambda^{\Pi_n - 1}) \leq Cn^d \qquad \text{where} \qquad C := \max_{x\in\mathbb{Z}^d/\mathcal{L}} e^+(\{x\});$$

the energy term is bounded by Cn^d because ϕ is q-Lipschitz μ-almost surely, and the entropy term is nonpositive because λ^{Π_n-1} is a counting measure. In particular, $\mathcal{H}(\mu|\Phi) \leq C$. Fix $u \in U_q$, and consider a subsequential limit μ of the sequence

$$\mu_n := \frac{1}{|\Pi_n \cap \mathcal{L}|} \sum_{x\in\Pi_n\cap\mathcal{L}} \delta_{\theta_x \phi^u}.$$

This limit μ is clearly supported on q-Lipschitz functions and is automatically shift-invariant and satisfies $S(\mu) = u$; in particular, $\sigma(u) \leq C < \infty$. This proves that σ is bounded by C on U_q.

Consider now the continuous case $E = \mathbb{R}$. For the lemma, we must show that σ is finite on U_q, and infinite on ∂U_q. Fix $u \in U_q$. Then ϕ^u is $q_{3\varepsilon}$-Lipschitz for $\varepsilon > 0$ sufficiently small. Let $X = (X_x)_{x\in\mathbb{Z}^d}$ denote an i.i.d. family of random variables which are uniformly random in the interval $[0, \varepsilon]$. Write $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ for the measure in which ϕ has the distribution of $\phi^u + X$. Then ϕ is q_ε-Lipschitz almost surely. It is straightforward to see that

$$\mathcal{H}_{\Pi_n}(\mu|\Phi) = \mu(H^0_{\Pi_n}) + \mathcal{H}_{\mathcal{F}^{\nabla}_{\Pi_n}}(\mu|\lambda^{\Pi_n-1}) \leq (C - \log\varepsilon)n^d \text{ where } C := \max_{x\in\mathbb{Z}^d/\mathcal{L}} e^+_\varepsilon(\{x\});$$

in particular, $\mathcal{H}(\mu|\Phi) \leq C - \log\varepsilon < \infty$. Clearly $S(\mu) = u$, and so $\sigma(u) < \infty$. Finally, consider $u \in \partial U_q$. Suppose that $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ has slope u. Then at least one of the following two must hold true:

1. ϕ is not q-Lipschitz, with positive μ-probability,

2. $\phi(y) - \phi(x)$ is deterministic in μ for some distinct vertices x and y.

This follows from Lemma 3.6.1 which gives a characterisation of U_q. In the former case we have $\mathcal{H}(\mu|\Phi) = \infty$ as was shown at the beginning of this proof. In the latter case, we observe that

$$\mathcal{H}_{\mathcal{F}^{\nabla}_{\Pi_n}}(\mu|\lambda^{\Pi_n - 1}) = \infty$$

for n sufficiently large, because $\mu\pi_{\Pi_n}$ is not absolutely continuous with respect to $\lambda^{\Pi_n - 1}$. This also implies that $\mathcal{H}(\mu|\Phi) = \infty$. We have now shown that $\sigma = \infty$ on ∂U_q. $\square$

3.8 Minimisers of the specific free energy

Recall that a minimiser is a shift-invariant measure $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ which satisfies

$$\mathcal{H}(\mu|\Phi) = \sigma(S(\mu)) < \infty,$$

and recall the discussion of minimisers in Subsection 3.4.1, in particular Definition 3.4.3. The purpose of this section is to prove the following theorem, which provides us with several properties of minimisers, and is equivalent to the conjunction of Theorem 3.4.2 and Theorem 3.4.4.

Theorem 3.8.1. *Let $\Phi \in \mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$, and consider a minimiser $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$. Fix $\Lambda \subset\subset \mathbb{Z}^d$, and write μ^{ϕ} for the regular conditional probability distribution of μ on $(\Omega, \mathcal{F})$ corresponding to the projection map $\Omega \to E^{\mathbb{Z}^d \smallsetminus \Lambda}$. Then for μ-almost every $\phi \in \Omega$, we have $\mu^{\phi}\pi_{\Lambda} \in \mathcal{A}_{\Lambda,\phi}$. In particular, if $\mu(\Omega_{\gamma}) = 1$, then μ is an almost Gibbs measure. In general, the former implies that μ has finite energy, in the sense that*

$$1_{\Omega_q}(\mu\pi_{\mathbb{Z}^d \smallsetminus \Lambda} \times \lambda^{\Lambda}) \ll \mu,$$

where Ω_q is the set of q-Lipschitz height functions.

We first introduce the definition of the max-entropy, which is due to Datta [8].

Definition 3.8.2. Let $(X, \mathcal{X})$ denote a measurable space, endowed with some finite measures μ and ν. Then the *max-entropy* of μ with respect to ν is defined by

$$\mathcal{H}^{\infty}(\mu|\nu) := \log\inf\{\lambda \geq 0 : \mu \leq \lambda\nu\} = \begin{cases} \operatorname{ess\,sup}\log f & \text{if } \mu \ll \nu \text{ where } f = d\mu/d\nu, \\ \infty & \text{otherwise.} \end{cases}$$

The *max-diameter* of a non-empty set $\mathcal{A}$ of finite measures on $(X, \mathcal{X})$ is defined by

$$\operatorname{Diam}^{\infty}\mathcal{A} := \sup_{\mu,\nu \in \mathcal{A}} \mathcal{H}^{\infty}(\mu|\nu).$$

If $\operatorname{Diam}^{\infty}\mathcal{A} < \infty$, then all measures in $\mathcal{A}$ are absolutely continuous with respect to one another, with uniform lower and upper bounds on the Radon-Nikodym derivatives.

Proposition 3.8.3. *Suppose that $\Lambda \subset \Delta \subset\subset \mathbb{Z}^d$ with $\Lambda \subset \Delta^{-R}$. Then we have* $\operatorname{Diam}^{\infty}\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi}) \leq 4e^{-}(\Lambda) < \infty$. *In particular,* $\operatorname{Diam}^{\infty}\mathcal{A}_{\Lambda,\phi} \leq 4e^{-}(\Lambda) < \infty$.

Proof. Claim first that $\operatorname{Diam}^\infty \mathcal{A}_{\Lambda,\Delta,\phi} \leq 4e^-(\Lambda)$. Consider two random fields $\nu_1, \nu_2 \in \mathcal{P}(\Omega, \mathcal{F})$ with $\nu_1\pi_\Delta = \nu_2\pi_\Delta = \delta_{\phi_\Delta}$. Then

$$\nu_i\gamma_\Lambda\pi_\Lambda = \int \frac{1}{Z_\Lambda(\psi)} e^{-H(\cdot,\psi_{\mathbb{Z}^d\smallsetminus\Lambda})}\lambda^\Lambda d\nu_i(\psi).$$

But since $\psi_\Delta = \phi_\Delta$ almost surely in both ν_1 and ν_2, the dependence of $H(\cdot, \psi_{\mathbb{Z}^d\smallsetminus\Delta})$ on ψ is bounded by $e^-(\Lambda)$. This error term appears twice in each measure $\nu_i\gamma_\Lambda\pi_\Lambda$; directly in the Hamiltonian, and indirectly in the normalisation constant. Thus, in calculating the Radon-Nikodym derivative between the two measures, the term appears four times. This proves the claim. By Lemma 2.5.1 in Chapter 2, this also implies that $\operatorname{Diam}^\infty \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi}) \leq 4e^-(\Lambda)$. □

We also need the following lemma, which is an adaptation of an intermediate result in Chapter 2 to the gradient setting.

Lemma 3.8.4. *Fix $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$, and define, for $\Lambda \subsetneq \Delta \subset\subset \mathbb{Z}^d$,*

$$K_\mu(\Lambda, \Delta) := \inf_{\nu\,\in\,\mathcal{P}(\Omega,\mathcal{F}^\nabla)\text{ with }\nu\pi_\Delta\,=\,\mu\pi_\Delta} \mathcal{H}_{\mathcal{F}^\nabla_\Delta}(\mu|\nu\gamma_\Lambda) \geq 0.$$

Then $K_\mu(\cdot,\cdot)$ is superadditive in the first argument, and increasing in the second argument.

Proof. It is straightforward to see that $K_\mu(\Lambda, \Delta)$ is increasing in Δ: increasing Δ restricts the set of measures ν for the infimum, while increasing the σ-algebra $\mathcal{F}^\nabla_\Delta$ for the entropy. Both operations increase the value of $K_\mu(\Lambda, \Delta)$. For superadditivity in Λ, it suffices to prove that

$$\begin{aligned}\inf_{\nu\,\in\,\mathcal{P}(\Omega,\mathcal{F}^\nabla)\text{ with }\nu\pi_\Delta\,=\,\mu\pi_\Delta} &\mathcal{H}_{\mathcal{F}^\nabla_\Delta}(\mu|\nu\gamma_\Lambda) \\ &\geq \sum_{i\in\{1,2\}} \inf_{\nu\,\in\,\mathcal{P}(\Omega,\mathcal{F}^\nabla)\text{ with }\nu\pi_\Delta\,=\,\mu\pi_\Delta} \mathcal{H}_{\mathcal{F}^\nabla_\Delta}(\mu|\nu\gamma_{\Lambda_i})\end{aligned}$$

for Λ_1 and Λ_2 disjoint with $\Lambda := \Lambda_1 \cup \Lambda_2 \subsetneq \Delta$. This follows from Lemma 2.4.1 in Chapter 2. Observe that that lemma does not concern the gradient setting, which provides us with a slight complication. However, since we choose Λ to be a strict subset of Δ, we can fix a vertex $x \in \Delta \smallsetminus \Lambda$ to serve as a reference vertex for the gradient setting for all three entropy calculations in the display, thus translating the inequality to the non-gradient setting. □

Lemma 3.8.5. *If μ is a minimiser, then $K_\mu \equiv 0$.*

Proof. Fix $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$. Then K_μ is shift-invariant, in the sense that $K_\mu(\Lambda, \Delta) = K_\mu(\theta\Lambda, \theta\Delta)$ for $\Lambda \subsetneq \Delta \subset\subset \mathbb{Z}^d$ and $\theta \in \Theta$. Using also the properties of K_μ in the previous defining lemma, it is immediate that $K_\mu \equiv 0$ if and only if $K_\mu(\Pi_n^{-R}, \Pi_n) = o(n^d)$ as $n \to \infty$. Moreover, by definition of K_μ, it is immediate that

$$K_\mu(\Pi_n^{-R}, \Pi_n) \leq \mathcal{H}_{\mathcal{F}^\nabla_{\Pi_n}}(\mu|\mu\gamma_{\Pi_n^{-R}}).$$

We must therefore prove that $\mathcal{H}(\mu|\Phi) = \sigma(S(\mu)) < \infty$ implies that the expression on the right in this display is of order $o(n^d)$ as $n \to \infty$. If this expression is not of order $o(n^d)$, then there is an $n \in \mathbb{N}$ and an $\varepsilon > 0$ such that

$$\mathcal{H}_{\mathcal{F}^\nabla_{\Pi_n}}(\mu|\mu\gamma_{\Pi_n^{-R}}) \geq 2e^-(\Pi_n) + \varepsilon. \tag{3.8.6}$$

We will use this inequality to construct another $\mathcal{L}$-invariant measure μ'' of the same slope as μ and with a strictly smaller specific free energy. This proves that $\mathcal{H}(\mu|\Phi) \neq \sigma(S(\mu))$.

For $\Lambda \subset\subset \mathbb{Z}^d$, we denote by γ_Λ^* the kernel $\gamma_{\Lambda^{-R}}$, only now with respect to the partial Hamiltonian $H_{\Lambda^{-R},\Lambda}$ rather than the full Hamiltonian $H_{\Lambda^{-R}}$. With a straightforward entropy calculation one can demonstrate that

$$\mathcal{H}_\Lambda(\nu\gamma_{\Pi_n}^*|\Phi) \leq \mathcal{H}_\Lambda(\nu|\Phi) - \varepsilon$$

for any $\Lambda \subset\subset \mathbb{Z}^d$ containing Π_n, and for any $\nu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ with $\nu\pi_{\Pi_n} = \mu\pi_{\Pi_n}$. This can be done by calculating each free energy term first over the σ-algebra generated by the vertices in $\Lambda \smallsetminus \Pi_n^{-R}$, then over the remaining vertices. The first term is the same for $\nu\gamma_{\Pi_n}^*$ and ν since the kernel modifies the values of ϕ in Π_n^{-R} only; the difference between the two measures for the second term is at least ε due to (3.8.6) and because

$$|H_{\Pi_n^{-R},\Lambda} - H_{\Pi_n^{-R},\Pi_n}| \leq e^-(\Pi_n).$$

If Λ and Δ are disjoint, then clearly γ_Λ^* and γ_Δ^* commute. Let M denote the smallest multiple of N which exceeds n, and write

$$\mu' := \mu \prod_{x \in M\cdot\mathbb{Z}^d} \gamma_{\Pi_n+x}^*;$$

this measure is $M\cdot\mathbb{Z}^d$-invariant, but not necessarily $\mathcal{L}$-invariant. By the inequality in the previous paragraph, we have $\mathcal{H}_{\Pi_{kM}}(\mu'|\Phi) \leq \mathcal{H}_{\Pi_{kM}}(\mu|\Phi) - k^d\varepsilon$ for any $k \in \mathbb{N}$. As $M\cdot\mathbb{Z}^d$-invariant measures, we have $S(\mu') = S(\mu)$ and $\mathcal{H}(\mu'|\Phi) \leq \mathcal{H}(\mu|\Phi) - \varepsilon/M^d < \mathcal{H}(\mu|\Phi)$. To make μ' also $\mathcal{L}$-invariant, simply define

$$\mu'' := \frac{1}{|\mathcal{L}/(M\cdot\mathbb{Z}^d)|} \sum_{x \in \mathcal{L}/(M\cdot\mathbb{Z}^d)} \theta_x\mu'.$$

The averaging procedure does not change the slope or the specific free energy. This is the desired measure. □

Proof of Theorem 3.8.1. The theorem contains three claims. The second claim follows directly from the first claim and the definition of an almost Gibbs measure. We shall quickly demonstrate that the third claim also follows from the first claim, before focusing on that first claim. Assume that the first claim holds true. Observe first that, by assumption, for μ-almost every ϕ,

$$1_{\Omega_q}(\delta_{\phi_{\mathbb{Z}^d\smallsetminus\Lambda}} \times \lambda^\Lambda)\pi_\Lambda \ll \gamma_\Lambda(\cdot,\phi)\pi_\Lambda \in \mathcal{A}_{\Lambda,\phi} \ni \mu^\phi\pi_\Lambda.$$

But all measures in $\mathcal{A}_{\Lambda,\phi}$ are absolutely continuous with respect to one another, by Proposition 3.8.3 and the comment preceding it. Therefore

$$1_{\Omega_q}(\delta_{\phi_{\mathbb{Z}^d\smallsetminus\Lambda}} \times \lambda^\Lambda) \ll \mu^\phi$$

for μ-almost every ϕ, which implies that $1_{\Omega_q}(\mu\pi_{\mathbb{Z}^d\smallsetminus\Lambda} \times \lambda^\Lambda) \ll \mu$.

Focus finally on the first claim. By the previous lemma, it suffices to prove that $K_\mu \equiv 0$ implies that $\mu^\phi\pi_\Lambda \in \mathcal{A}_{\Lambda,\phi}$ for μ-almost every ϕ. The proof is nearly identical to the proof of Lemma 2.5.7 in Chapter 2. Fix $\Delta \subset\subset \mathbb{Z}^d$ with $\Lambda \subset \Delta^{-R}$; it suffices to demonstrate that $K_\mu \equiv 0$ implies that $\mu^\phi\pi_\Lambda \in \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi})$ for μ-almost every ϕ.

By choice of Δ, we have $\operatorname{Diam}^\infty \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi}) < \infty$. Write Δ_n for $\{-n,\dots,n\}^d \subset\subset \mathbb{Z}^d$, and write μ_n^ϕ for the regular conditional probability distribution of μ on $(\Omega,\mathcal{F})$ corresponding to the projection map $\Omega \to E^{\Delta_n \smallsetminus \Lambda}$. We only consider $n \in \mathbb{N}$ so large that $\Lambda \subset \Delta \subset \Delta_n$. As in the proof of Lemma 2.5.7 in Chapter 2, we observe that $K_\mu(\Lambda,\Delta_n) = 0$ implies that for μ-almost every ϕ,

1. $\mu_n^\phi \pi_\Lambda \in \mathcal{C}(\mathcal{A}_{\Lambda,\Delta_n,\phi}) \subset \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi})$ for fixed n—this follows from Lemma 2.5.1 in Chapter 2,
2. $\mu_n^\phi(A) \to \mu^\phi(A)$ for fixed $A \in \mathcal{F}_\Lambda$, by the bounded martingale convergence theorem,
3. $\mu_n^\phi \pi_\Lambda \to \mu^\phi \pi_\Lambda \in \mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi})$ by compactness of $\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi})$ in the strong topology.

Compactness of $\mathcal{C}(\mathcal{A}_{\Lambda,\Delta,\phi})$ follows from Lemma 2.5.1 in Chapter 2 and the fact that $\mathcal{A}_{\Lambda,\Delta,\phi}$ has finite max-diameter. $\square$

3.9 Ergodic decomposition of shift-invariant measures

In this section we cite some standard results on ergodic decompositions of shift-invariant random fields from the work of Georgii [20]. Recall that $\mathcal{I}_\mathcal{L}^\nabla$ is the σ-algebra of shift-invariant gradient events, and that $\operatorname{ex}\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ is the set of ergodic gradient measures, endowed with the σ-algebra $e(\operatorname{ex}\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla))$.

The following result is a direct adaptation of Theorem 14.10 in [20] to the gradient setting of this chapter. Informally, the theorem asserts that if μ is a shift-invariant gradient random field, then the regular conditional probability distribution of μ given the information in $\mathcal{I}_\mathcal{L}^\nabla$ is well-defined.

Theorem 3.9.1. *There is a unique affine bijection*

$$w : \mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla) \to \mathcal{P}(\operatorname{ex}\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla), e(\operatorname{ex}\mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla))),\ \mu \mapsto w_\mu$$

such that

$$\mu = \int \nu dw_\mu(\nu)$$

for all $\mu \in \mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$. For any $A \in \mathcal{F}^\nabla$ and $c \in \mathbb{R}$, this bijection satisfies

$$w_\mu(\nu(A) \leq c) = \mu(\mu(A|\mathcal{I}_\mathcal{L}^\nabla) \leq c).$$

Definition 3.9.2. The measure w_μ is called the *ergodic decomposition* of μ.

Proof of Theorem 3.9.1. Let $(e_1,\dots,e_d)$ denote the standard basis of $\mathbb{R}^d$. The measure μ can be considered a non-gradient measure, by associating to each vertex $x \in \mathbb{Z}^d$ the tuple $(\phi(x+e_1)-\phi(x),\dots,\phi(x+e_d)-\phi(x)) \in E^d$. Theorem 14.10 in [20] applies to this non-gradient measure, which immediately implies the current theorem. $\square$

It was shown in previous sections that the slope and specific free energy are affine. In fact, these functionals are also strongly affine. This is the subject of the following two results.

Proposition 3.9.3. *The functional S is strongly affine, that is,*

$$S(\mu) = \int S(\nu) dw_\mu(\nu)$$

for any $\mu \in \mathcal{P}_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$ with finite slope.

This proposition is immediate from the definition of S.

Theorem 3.9.4. *If $\Phi \in \mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$, then the functional $\mathcal{H}(\cdot|\Phi)$ is strongly affine, that is,*

$$\mathcal{H}(\mu|\Phi) = \int \mathcal{H}(\nu|\Phi) dw_\mu(\nu) \tag{3.9.5}$$

for any $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$.

Proof. If μ is not supported on q-Lipschitz functions, then the left and right of (3.9.5) equal ∞; recall that $\mathcal{H}(\cdot|\Phi)$ is bounded below by Theorem 3.4.1 so that the integral on the right in (3.9.5) is always well-defined.

Consider now the case that μ is supported on q-Lipschitz functions, which means in particular that μ is K-Lipschitz for K minimal subject to $Kd_1 \geq q$. In that case we have

$$\mathcal{H}(\mu|\Phi) = \langle \mu|\Phi \rangle + \lim_{n\to\infty} n^{-d} \mathcal{H}_{\mathcal{F}^\nabla_{\Pi_n}}(\mu|\lambda^{\Pi_n - 1}), \tag{3.9.6}$$

once it is established that the sequence on the right tends to some limit in $(-\infty, \infty]$. The functional $\langle \cdot|\Phi \rangle$ is clearly strongly affine on $\mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$. Let us therefore focus on the limit on the right in the display. It suffices to demonstrate that the second limit in the display is well-defined, bounded below, and strongly affine in its dependence on μ, once restricted to K-Lipschitz measures. The idea is to use Theorem 15.20 in [20], which concerns the non-gradient setting. The measure μ can be made into a shift-invariant, non-gradient measure by considering the values of ϕ modulo $4K$. It is clear that the gradient of ϕ can be reconstructed from this reduced height function, if we use the extra information that ϕ is K-Lipschitz. This is formalised as follows. Write $\hat{E}$ for the set $E/4K\mathbb{Z}$, and endow it with the Borel σ-algebra $\hat{\mathcal{E}}$ and the Lebesgue measure $\hat{\lambda}$ which satisfies $\hat{\lambda}(\hat{E}) = 4K$. Write $\hat{\Omega}$ for the set of functions from $\mathbb{Z}^d$ to $\hat{E}$, and $\hat{\mathcal{F}}$ for the product σ-algebra on Ω. Define the measure $\hat{\mu}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ as follows: first sample a pair (ϕ, a) from $\mu \times (\hat{\lambda}/4K)$, the final sample $\hat{\phi}$ is then obtained by setting $\hat{\phi}(x) = \phi(x) - \phi(0) + a \in \hat{E}$. The measure $\hat{\mu}$ is clearly $\mathcal{L}$-invariant. Note that, for $\Lambda \subset\subset \mathbb{Z}^d$ nonempty,

$$\mathcal{H}_{\mathcal{F}^\nabla_\Lambda}(\mu|\lambda^{\Lambda-1}) = \mathcal{H}_{\hat{\mathcal{F}}_\Lambda}(\hat{\mu}|\hat{\lambda}^\Lambda) + \log 4K,$$

where $\hat{\mathcal{F}}_\Lambda := \sigma(\hat{\phi}(x) : x \in \Lambda)$. By Theorem 15.20 in [20], the limit

$$\lim_{n\to\infty} n^{-d} \mathcal{H}_{\mathcal{F}^\nabla_{\Pi_n}}(\mu|\lambda^{\Pi_n - 1}) = \lim_{n\to\infty} n^{-d} \mathcal{H}_{\hat{\mathcal{F}}_{\Pi_n}}(\hat{\mu}|\hat{\lambda}^{\Pi_n})$$

is well-defined, bounded below by $-\log 4K$, and strongly affine over μ. □

Definition 3.9.7. For $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ a K-Lipschitz measure, define

$$\mathcal{H}(\mu|\lambda) := \lim_{n\to\infty} n^{-d} \mathcal{H}_{\mathcal{F}^\nabla_{\Pi_n}}(\mu|\lambda^{\Pi_n - 1}) \in [-\log 4K, \infty],$$

the *specific entropy* of μ. This quantity is well-defined and strongly affine over μ due to the proof of the previous theorem. Remark that $\mathcal{H}(\mu|\lambda) \leq 0$ whenever $E = \mathbb{Z}$.

Lemma 3.9.8. *Consider a potential $\Phi \in \mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$ and a measure $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$. Fix K minimal subject to $Kd_1 \geq q$. If μ is not K-Lipschitz, then $\mathcal{H}(\mu|\Phi) = \langle \mu|\Phi \rangle = \infty$, and if μ is K-Lipschitz, then $\mathcal{H}(\mu|\Phi) = \langle \mu|\Phi \rangle + \mathcal{H}(\mu|\lambda)$.*

Proof. This also follows from the proof of the previous theorem. □

We are now able to prove Theorem 3.4.5.

Proof of Theorem 3.4.5. Suppose that $u \in \bar{U}_\Phi$ is an exposed point of σ. By compactness of the lower level sets of $\mathcal{H}(\cdot|\Phi)$ (Theorem 3.4.1) and by continuity of $S(\cdot)$, there exists a minimiser $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ of slope u. Write w_μ for the ergodic decomposition of μ. Since both $S(\cdot)$ and $\mathcal{H}(\cdot|\Phi)$ are strongly affine (due to Proposition 3.9.3 and Theorem 3.9.4) and because u is an exposed point, we observe that w_μ-almost every component ν is an ergodic minimiser of slope u. $\square$

3.10 Limit equalities

This section provides the fundamental building blocks for the large deviations principle in the next section. The motivating thesis for this section is that $\sigma(u)$ can be approximated by integrals of $\exp -H^0_{\Pi_n}$ after restricting to height functions which are close to the slope u on $\partial^R \Pi_n$. It is possible to be more subtle: if one considers a measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F})$ with $\mathcal{H}(\mu|\Phi) < \infty$ and $S(\mu) \in U_\Phi$, then one can approximate $\mathcal{H}(\mu|\Phi)$ by integrals of $\exp -H^0_{\Pi_n}$ after restricting to height functions which are close to the slope $S(\mu)$ on $\partial^R \Pi_n$, and after restricting further to height functions ϕ whose *empirical measure* in Π_n approximates μ. The empirical measure of ϕ in Π_n is obtained by randomly shifting ϕ by a vertex in $\mathcal{L} \cap \Pi_n$. Analogous results for finite-range non-Lipschitz potentials can be found in Chapter 6 in [54]. However, the proof presented here differs from the proof in [54] to account for the generality of our setting, and the specificity of the discrete Lipschitz case.

3.10.1 Formal statement

Let us first introduce some simple notation for fixing boundary conditions.

Definition 3.10.1. Write 0_Λ for the smallest element in Λ in the dictionary order on $\mathbb{Z}^d$ whenever $\Lambda \subset\subset \mathbb{Z}^d$. Let $u \in U_\Phi$. If $E = \mathbb{Z}$, then write C^u_Λ for the set of height functions

$$\{\phi \in \Omega : \phi_{\partial^R \Lambda} - \phi(0_\Lambda) = \phi^u_{\partial^R \Lambda} - \phi^u(0_\Lambda)\} \in \mathcal{F}^\nabla_{\partial^R \Lambda}.$$

Now consider $E = \mathbb{R}$, and fix $\varepsilon > 0$. Write $C^u_{\Lambda,\varepsilon}$ for the set

$$\{\phi \in \Omega : |(\phi_{\partial^R \Lambda} - \phi(0_\Lambda)) - (\phi^u_{\partial^R \Lambda} - \phi^u(0_\Lambda))| \leq \varepsilon\} \in \mathcal{F}^\nabla_{\partial^R \Lambda}.$$

Abbreviate $C^u_{\Pi_n}$ and $C^u_{\Pi_n,\varepsilon}$ to C^u_n and $C^u_{n,\varepsilon}$ respectively.

Next, we formally define the empirical measure of a height function ϕ in Λ. Recall the definition of the basis $\mathcal{B}$ of the topology of weak local convergence on $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$ from Subsection 3.2.1.

Definition 3.10.2. In this definition, we adopt the following notation: if ϕ is a height function and $\Lambda \subset\subset \mathbb{Z}^d$, then write $\bar{\phi}_\Lambda$ for unique extension of ϕ_Λ to $\mathbb{Z}^d$ which equals $\phi(0_\Lambda)$ on the complement of Λ. For $\Lambda \subset\subset \mathbb{Z}^d$ and $\phi \in \Omega$, we define the measure $L_\Lambda(\phi)$ by

$$L_\Lambda(\phi) := \frac{1}{|\mathcal{L} \cap \Lambda|} \sum_{x \in \mathcal{L} \cap \Lambda} \delta_{\theta_x \bar{\phi}_\Lambda}.$$

This is called the *empirical measure* of ϕ in Λ. The kernel L_Λ is thus a probability kernel from $(\Omega, \mathcal{F}_\Lambda)$ to $(\Omega, \mathcal{F})$ which restricts to a kernel from $(\Omega, \mathcal{F}^\nabla_\Lambda)$ to $(\Omega, \mathcal{F}^\nabla)$. Now consider $B \in \mathcal{B}$. Write B_Λ for the event $B_\Lambda := \{\phi \in \Omega : L_\Lambda(\phi) \in B\}$; this event is $\mathcal{F}^\nabla_\Lambda$-measurable. We shall also write L_n and B_n for L_{Π_n} and B_{Π_n} respectively.

We start with the introduction of free boundary limits, which is slightly easier than the definition of pinned boundary limits. For free boundary limits, we integrate over all height functions having the appropriate empirical measure, irrespective of boundary conditions. It will be useful to define free boundary limits also for measures $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ which are not shift-invariant.

Definition 3.10.3. Let $\Lambda \subset\subset \mathbb{Z}^d$ and $B \in \mathcal{B}$. The *free boundary* estimate of B over Λ is given by

$$\mathrm{FB}_\Lambda(B) := -\log \int_{B_\Lambda} e^{-H_\Lambda^0} d\lambda^{\Lambda-1}.$$

Let $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$. The *free boundary* limits of B and μ respectively are given by

$$\mathrm{FB}(B) := \liminf_{n\to\infty} n^{-d}\,\mathrm{FB}_{\Pi_n}(B) \qquad \text{and} \qquad \mathrm{FB}(\mu) := \sup_{A \in \mathcal{B} \text{ with } \mu \in A} \mathrm{FB}(A).$$

Free boundary limits should be thought of as an asymptotic upper bound on the integral in the display, and this is why we take the limit inferior in the definition of $\mathrm{FB}(B)$—taking into account the minus sign which appears in the definition of $\mathrm{FB}_\Lambda(B)$. Indeed, the free boundary estimates are useful in proving the upper bound on probabilities in the large deviations principle in the next section. Remark that it is immediate from the definition of $\mathrm{FB}(\mu)$ that $\mathrm{FB}(\cdot)$ is lower-semicontinuous on the set of gradient measures in the topology of weak local convergence for which $\mathcal{B}$ forms a basis.

Finally, we introduce pinned boundary limits, which take into consideration also the value of ϕ on the boundary of Π_n. In this case, it is the lower bound on the integral of interest that matters to us; pinned boundary limits play a crucial role in the proof of the lower bound on probabilities in the large deviations principle.

Definition 3.10.4. Fix $u \in U_\Phi$ and $\varepsilon > 0$, and let $\Lambda \subset\subset \mathbb{Z}^d$ and $B \in \mathcal{B}$. If $E = \mathbb{R}$, then define

$$\mathrm{PB}_{\Lambda,u,\varepsilon}(B) := -\log \int_{C^u_{\Lambda,\varepsilon} \cap B_\Lambda} e^{-H_\Lambda^0} d\lambda^{\Lambda-1}.$$

If $E = \mathbb{Z}$, then define

$$\mathrm{PB}_{\Lambda,u}(B) := -\log \int_{C^u_{\Lambda} \cap B_\Lambda} e^{-H_\Lambda^0} d\lambda^{\Lambda-1}.$$

These are called the *pinned boundary* estimates of B over Λ. In either case, we set $\mathrm{PB}_{\Lambda,u,\varepsilon}(B) := \infty$ and $\mathrm{PB}_{\Lambda,u}(B) := \infty$ whenever $u \notin U_\Phi$. Consider now also some random field $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$. The *pinned boundary* limits of B and μ are defined as follows:

$$\mathrm{PB}_{u,\varepsilon}(B) := \limsup_{n\to\infty} n^{-d}\,\mathrm{PB}_{\Pi_n,u,\varepsilon}(B), \;\; \mathrm{PB}(\mu) := \sup_{\varepsilon > 0 \text{ and } A \in \mathcal{B} \text{ with } \mu \in A} \mathrm{PB}_{S(\mu),\varepsilon}(A)$$

whenever $E = \mathbb{R}$, and if $E = \mathbb{Z}$, then

$$\mathrm{PB}_u(B) := \limsup_{n\to\infty} n^{-d}\,\mathrm{PB}_{\Pi_n,u}(B), \qquad \mathrm{PB}(\mu) := \sup_{A \in \mathcal{B} \text{ with } \mu \in A} \mathrm{PB}_{S(\mu)}(A).$$

It is again immediate from these definitions that for fixed $u \in U_\Phi$, the functional $\mathrm{PB}(\cdot)$ is lower-semicontinuous on the set $\{S(\cdot) = u\} \subset \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$.

For the proof of the large deviations principle in the next section, we require the following equalities and inequalities.

Theorem 3.10.5. *If $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ and $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$, then*

$$\mathcal{H}(\mu|\Phi) = \mathrm{FB}(\mu) = \mathrm{PB}(\mu),$$

unless $E = \mathbb{Z}$ and $S(\mu) \in \partial U_\Phi$. If however $E = \mathbb{Z}$ and $S(\mu) \in \partial U_\Phi$, then

$$\mathrm{FB}(\mu) \geq \mathcal{H}(\mu|\Phi).$$

Finally, if $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla) \smallsetminus \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$, then $\mathrm{FB}(\mu) = \infty$.

Free and pinned boundary limits are calculated along the sequence $(\Pi_n)_{n\in\mathbb{N}}$. This choice is convenient, but by no means necessary. In the following sections, we do not only prove the inequalities presented in the theorem: we also prove some generalisations thereof where these quantities are calculated over sequences of the form $(\Lambda_n)_{n\in\mathbb{N}}$ with $\Lambda_n := \Lambda^{-m}(nD)$, where D is a bounded convex subset of $\mathbb{R}^d$ of positive Lebesgue measure, and where $m \in \mathbb{Z}_{\geq 0}$. Observe that in this notation, $\Pi_n = \Lambda_n$ for $m = 0$ and $D = [0, 1)^d \subset \mathbb{R}^d$.

Definition 3.10.6. Write $\mathcal{C}$ for the set of bounded convex subsets of $\mathbb{R}^d$ of positive Lebesgue measure.

The definitions imply that $\mathrm{PB}(\mu) \geq \mathrm{FB}(\mu)$ for μ shift-invariant. In Subsection 3.10.2 we discuss free boundary limits. In particular, we show that $\mathrm{FB}(\mu) \geq \mathcal{H}(\mu|\Phi)$ whenever μ is shift-invariant, and that $\mathrm{FB}(\mu) = \infty$ whenever μ is not shift-invariant. In Subsection 3.10.3 we prove that $\mathrm{PB}(\mu) \leq \mathcal{H}(\mu|\Phi)$ whenever μ is ergodic with $S(\mu) \in U_\Phi$. In Subsection 3.10.4 we extend this inequality to shift-invariant measures μ which are not ergodic.

3.10.2 Free boundary limits: empirical measure argument

The idea in this subsection is always to use the set B, the empirical measures $L_n(\phi)$ for $\phi \in B_n$, as well as the subsequential limits thereof as $n \to \infty$, to derive the desired inequalities which were mentioned in the previous subsection. Let us first cover the case that μ is not shift-invariant.

Lemma 3.10.7. *If $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla) \smallsetminus \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$, then* $\mathrm{FB}(\mu) = \infty$.

Proof. If μ is not shift-invariant, then there is a shift $\theta \in \Theta$ and a continuous cylinder function $g : \Omega \to [0, 1]$ such that $\mu(g - \theta g) \neq 0$. Define $f := g - \theta g$; this is a bounded continuous cylinder function such that $\mu(f) \neq 0$. Define $\varepsilon := |\mu(f)|/2$ and $B := \{\nu : |\nu(f) - \mu(f)| < \varepsilon\} \in \mathcal{B}$. For $\Lambda \subset\subset \mathbb{Z}^d$ fixed and for n large, the measure $L_n(\phi) = L_{\Pi_n}(\phi)$ restricted to $\mathcal{F}_\Lambda^\nabla$ looks almost shift-invariant. More precisely, the sequence of functions

$$\Omega \to [-1, 1], \, \phi \mapsto L_n(\phi)(f)$$

converges to 0 uniformly over $\phi \in \Omega$ as $n \to \infty$. This proves that $B_n = B_{\Pi_n}$ is empty for n sufficiently large, that is, $\mathrm{FB}(\mu) \geq \mathrm{FB}(B) = \infty$. $\square$

Next, we consider shift-invariant gradient random fields.

Lemma 3.10.8. *For any $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$, we have* $\mathrm{FB}(\mu) \geq \mathcal{H}(\mu|\Phi)$.

We start with the following auxiliary lemma.

Lemma 3.10.9. *Suppose that $B \in \mathcal{B}$ satisfies* $\mathrm{FB}(B) < \infty$*. Then $\bar{B}$ contains a shift-invariant measure μ with $\mathcal{H}(\mu|\Phi) \leq \mathrm{FB}(B)$.*

Proof. Write ν_n^B for the normalised version of the measure $1_{B_n} e^{-H^0_{\Pi_n}} \lambda^{\Pi_n - 1}$ for each $n \in \mathbb{N}$, and observe that

$$\mathrm{FB}(B) = \liminf_{n\to\infty} n^{-d} \mathcal{H}_{\Pi_n}(\nu_n^B|\Phi).$$

We focus on *good* subsequences of n, that is, subsequences along which the limit inferior is reached.

Write $m : \mathbb{N} \to \mathbb{N}$ for a sequence of integers with $m(n) \to \infty$ and $m(n)/n \to 0$ as $n \to \infty$, and set $\Pi'_n := \Pi_n^{-m(n)} = \{m(n), \ldots, n - m(n) - 1\}^d \subset \Pi_n$. Fix $N \in \mathbb{N}$ minimal subject to $N \cdot \mathbb{Z}^d \subset \mathcal{L}$, and let k denote an integer multiple of N. Let n denote another integer, which is so large that $m(n) > k$. The idea is now to apply Lemma 3.7.12 to translates of Π_k. In particular, if we write $\Pi''_{n,k}$ for the set Π_n with the sets $\Pi_k + x$ removed for all x in $\Pi'_n \cap (k \cdot \mathbb{Z}^d)$, then that lemma asserts that

$$\mathcal{H}_{\Pi_n}(\nu_n^B|\Phi) \geq \mathcal{H}_{\Pi''_{n,k}}(\nu_n^B|\Phi) + \sum_{x \in \Pi'_n \cap (k\cdot\mathbb{Z}^d)} \mathcal{H}_{\Pi_k + x}(\nu_n^B|\Phi) - e^*(\Pi_k + x). \tag{3.10.10}$$

The set $\Pi''_{n,k}$ is always connected and, as $n \to \infty$, we have $|\Pi''_{n,k}| = o(n^d)$. Therefore the first term on the right in (3.10.10) has a lower bound of order $o(n^d)$. Moreover, the value of $e^*(\Pi_k + x)$ is independent of x as long as x lies in $\mathcal{L}$, and therefore we obtain the asymptotic bound

$$\frac{1}{|\Pi'_n \cap (k \cdot \mathbb{Z}^d)|} \sum_{x \in \Pi'_n \cap (k\cdot\mathbb{Z}^d)} \mathcal{H}_{\Pi_k + x}(\nu_n^B|\Phi) \leq k^d \, \mathrm{FB}(B) + e^*(\Pi_k) + o(1) \tag{3.10.11}$$

as $n \to \infty$ along a good subsequence. Moreover, if we write $\mu^{n,k}$ for the measure

$$\frac{1}{|\Pi'_n \cap (k \cdot \mathbb{Z}^d)|} \sum_{x \in \Pi'_n \cap (k\cdot\mathbb{Z}^d)} \theta_x \nu_n^B,$$

then the previous inequality and convexity of relative entropy imply that

$$\mathcal{H}_{\Pi_k}(\mu^{n,k}|\Phi) \leq k^d \, \mathrm{FB}(B) + e^*(\Pi_k) + o(1)$$

as $n \to \infty$ along a good subsequence. We may replace the sublattice $k \cdot \mathbb{Z}^d$ by another set $k \cdot \mathbb{Z}^d + y$ for $y \in \mathcal{L}/(k \cdot \mathbb{Z}^d)$ in the previous discussion, and by doing so and averaging further, it is immediate that the sequence of measures μ^n defined by

$$\frac{1}{|\Pi'_n \cap \mathcal{L}|} \sum_{x \in \Pi'_n \cap \mathcal{L}} \theta_x \nu_n^B,$$

also satisfies

$$\mathcal{H}_{\Pi_k}(\mu^n|\Phi) \leq k^d \, \mathrm{FB}(B) + e^*(\Pi_k) + o(1)$$

as $n \to \infty$ along a good subsequence. Compactness of the lower level sets of relative entropy implies that the sequence μ^n has a subsequential limit—at least when restricted to $\mathcal{F}^{\nabla}_{\Pi_k}$. Using a standard diagonal argument for convergence for all integers $k \in N \cdot \mathbb{N}$, one obtains a subsequential limit μ which is shift-invariant and satisfies

$$\mathcal{H}_{\Pi_k}(\mu|\Phi) \leq k^d \, \mathrm{FB}(B) + e^*(\Pi_k)$$

for all k, that is, $\mathcal{H}(\mu|\Phi) \leq \mathrm{FB}(B)$. This measure must clearly lie in $\bar{B}$ by construction. □

Proof of Lemma 3.10.8. Fix $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$, and suppose that $\mathrm{FB}(\mu) < \infty$. Then the lower level set of the specific free energy $M_{\mathrm{FB}(\mu)}$ endowed with the topology of weak local convergence is metrisable, and therefore we may choose for each n an open set $B^n \in \mathcal{B}$, containing μ and with $B^n \cap M_{\mathrm{FB}(\mu)}$ of diameter at most $1/n$ in this metric. Then each set $\bar{B}^n$ contains a measure μ^n with $\mathcal{H}(\mu^n|\Phi) \leq \mathrm{FB}(B^n) \leq \mathrm{FB}(\mu)$. By choice of B^n we must have $\mu^n \to \mu$, and lower-semicontinuity implies that

$$\mathcal{H}(\mu|\Phi) \leq \liminf_{n\to\infty} \mathcal{H}(\mu^n|\Phi) \leq \liminf_{n\to\infty} \mathrm{FB}(B^n) \leq \mathrm{FB}(\mu). \qquad \square$$

Finally, we discuss how to extend this result to other shapes.

Definition 3.10.12. For fixed $D \in \mathcal{C}$ and $\mu \in \mathcal{P}(\Omega, \mathcal{F}^{\nabla})$, we write

$$\mathrm{FB}(\mu : D) := \sup_{B \in \mathcal{B} \text{ with } \mu \in B} \liminf_{n\to\infty} n^{-d} \, \mathrm{FB}_{\Lambda_n}(B),$$

where we write Λ_n for $\Lambda(nD) = nD \cap \mathbb{Z}^d$.

The previous results extend as follows by analogous arguments—the Lebesgue measure $\mathrm{Leb}(D)$ would first appear as a factor on the left in (3.10.11) in the generalised argument.

Lemma 3.10.13. *Consider $D \in \mathcal{C}$ and $\mu \in \mathcal{P}(\Omega, \mathcal{F}^{\nabla})$. If μ is not shift-invariant, then $\mathrm{FB}(\mu : D) = \infty$, and if μ is shift-invariant, then $\mathrm{FB}(\mu : D) \geq \mathrm{Leb}(D) \cdot \mathcal{H}(\mu|\Phi)$.*

3.10.3 Pinned boundary limits for μ ergodic: truncation argument

The goal of this section is to derive the following lemma. The proof starts with a simple reduction, and is then intermitted to state an auxiliary result and to give an overview of the remainder of the proof. The proof extends the random truncation argument in [54] to the infinite-range Lipschitz setting.

Lemma 3.10.14. *If $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ is ergodic and $u := S(\mu) \in U_{\Phi}$, then $\mathrm{PB}(\mu) \leq \mathcal{H}(\mu|\Phi)$.*

Proof. It suffices to consider the case that $\mathcal{H}(\mu|\Phi) < \infty$, which implies in particular that μ is K-Lipschitz. We first focus on the discrete case $E = \mathbb{Z}$, then generalise to the continuous case $E = \mathbb{R}$; the latter comes with some additional technical complications.

The discrete case. Pick $B \in \mathcal{B}$ with $\mu \in B$. It suffices to show that

$$\mathcal{H}(\mu|\Phi) = \lim_{n\to\infty} n^{-d} \mathcal{H}_{\Pi_n}(\mu|\Phi) \geq \limsup_{n\to\infty} n^{-d} \, \mathrm{PB}_{\Pi_n,u}(B) = \limsup_{n\to\infty} n^{-d} \mathcal{H}_{\Pi_n}(\nu_n^B|\Phi),$$

where ν_n^B is the normalised measure

$$\nu_n^B := \frac{1}{Z} 1_{C_n^u \cap B_n} e^{-H^0_{\Pi_n}} \lambda^{\Pi_n - 1}.$$

Observe that ν_n^B minimises $\mathcal{H}_{\Pi_n}(\cdot|\Phi)$ over all measures which are supported on $C_n^u \cap B_n$. Therefore it suffices to construct a sequence of measures $(\mu_n)_{n\in\mathbb{N}}$, with each μ_n supported on $C_n^u \cap B_n$, and such that $\mathcal{H}_{\Pi_n}(\mu_n|\Phi) \leq \mathcal{H}_{\Pi_n}(\mu|\Phi) + o(n^d)$ as $n \to \infty$. Let us now intermit the proof to give an overview of the remainder of the proof, before continuing.

One continues roughly as follows. Always take 0 as a reference point for all gradient measures. This means that μ-almost surely $\phi(0) = 0$. Write $\phi_n^\pm$ for the largest and smallest q-Lipschitz extensions of $\phi^u_{\partial^R \Pi_n}$ to Π_n respectively, for each $n \in \mathbb{N}$. Define the random sets

$$A_n^- := \{x \in \Pi_n : \phi(x) < \phi_n^-(x)\} \quad \text{and} \quad A_n^+ := \{x \in \Pi_n : \phi(x) > \phi_n^+(x)\}.$$

Note that $\phi_n^\pm(0) = 0$ by definition; 0 is μ-almost surely not contained in $A_n^\pm$.

Since μ is ergodic and K-Lipschitz, almost every sample ϕ from μ is asymptotically close to u, in the sense of Theorem 3.10.15. As u belongs to U_q, the interior of the set of Lipschitz slopes, the function ϕ_n^+ is substantially larger than u on most vertices in Π_n. This means that $\mu(|A_n^+|) = o(n^d)$ as $n \to \infty$, and similarly $\mu(|A_n^-|) = o(n^d)$. For each $n \in \mathbb{N}$, define the measure $\mu_n^\pm$ as follows: to draw a sample from $\mu_n^\pm$, sample first a height function ϕ from μ, then replace this sample by $\psi := \phi_n^- \vee \phi_{\Pi_n} \wedge \phi_n^+$. Note that ϕ and ψ differ at at most $o(n^d)$ vertices in Π_n on average as $n \to \infty$. Moreover, the modified height function ψ is q-Lipschitz if the original height function ϕ was q-Lipschitz. In particular, we deduce that

$$\mathcal{H}_{\Pi_n}(\mu_n^\pm | \Phi) = \mathcal{H}_{\Pi_n}(\mu|\Phi) + o(n^d).$$

The measure $\mu_n^\pm$ is clearly supported on C_n^u, because ϕ_n^-, ϕ_n^+, and ϕ^u are equal on $\partial^R \Pi_n$. Using again the ergodicity of μ through Theorem 3.10.15, one can show that $\mu(B_n) \to 1$ as $n \to \infty$, and consequently $\mu_n^\pm(B_n) \to 1$ because ϕ and ψ agree on most vertices of Π_n. This proves that the sequence $(\mu_n)_{n\in\mathbb{N}}$ defined by $\mu_n := \mu_n^\pm(\cdot|B_n)$ is the desired sequence of measures. This concludes the proof overview for $E = \mathbb{Z}$. In the real case $E = \mathbb{R}$, the details are more involved, owing to the following two difficulties:

1. We cannot simply replace ϕ by $\phi_n^- \vee \phi_{\Pi_n} \wedge \phi_n^+$, because the measure so produced would not be absolutely continuous with respect to Lebesgue measure,

2. We only have a bound on $H_{\{x\}}(\phi)$ if ϕ is q_ε-Lipschitz at x; it is not sufficient to make modifications which are q-Lipschitz.

Let us now state Theorem 3.10.15 before continuing the proof of Lemma 3.10.14.

Theorem 3.10.15. *Consider $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ ergodic. Then $L_n(\phi)(f) \to \mu(f)$ as $n \to \infty$ for μ-almost every ϕ, for any bounded cylinder function f. Suppose now that μ is also K-Lipschitz with slope $u := S(\mu)$. Then μ-almost surely $\|\phi_{\Pi_n} - \phi(0) - u|_{\Pi_n}\|_\infty \leq \varepsilon n$ for n sufficiently large, for any fixed constant $\varepsilon > 0$.*

The first assertion is the ergodic theorem. The second assertion is straightforward: in the Lipschitz setting, the height difference $(\phi(x) - \phi(0))/\|x\|_1$ is approximately equal to the average of the gradient—which is bounded in magnitude—of ϕ over a large set $\Lambda \subset\subset \mathbb{Z}^d$.

Continuation of the proof of Lemma 3.10.14. Recall that $E = \mathbb{Z}$. By taking the set $B \in \mathcal{B}$ smaller if necessary, we suppose that B is of the form

$$B = \{\nu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla) : |\nu(f_i) - \mu(f_i)| < 2\eta \text{ for all } i\}$$

for a finite collection $(f_i)_i$ of continuous cylinder functions $f_i : \Omega \to [0, 1]$ and for some $\eta > 0$, and we write B^* for the same set with 2η replaced by η. The ergodic theorem asserts that $\mu(B_n^*) \to 1$ as $n \to \infty$. Consider μ a non-gradient measure on $(\Omega, \mathcal{F})$ by taking $0 \in \Pi_n$ as a reference point: this means that $\phi(0) = 0$ almost surely in μ.

Recall the definitions of $\phi_n^\pm$ and $A_n^\pm$ from the proof overview, and claim that $\mu(|A_n^\pm|) = o(n^d)$ as $n \to \infty$. The function ϕ_n^+ is pyramid-shaped, as in Figure 3.3—that figure concerns the more complicated continuous setting $E = \mathbb{R}$, but the shape of ϕ_n^+ is the same. Formally, this means that there exist constants $C' > 0$ and $\varepsilon' > 0$ such that for any $n \in \mathbb{N}$ and for any $x \in \Pi_n$,

$$\phi_n^+(x) \geq u(x) + \varepsilon' d_1(x, \partial^R \Pi_n) - C'. \tag{3.10.16}$$

This is a consequence of Lemma 3.6.1 and the fact that u is in U_Φ, the *interior* of the set of slopes u' for which $u'|_{\mathcal{L}}$ is q-Lipschitz. Now fix $\varepsilon'' > 0$. By (3.10.16), the number of points $x \in \Pi_n$ at which $\phi_n^+(x) \leq u(x) + \varepsilon'' n$ is bounded from above by $(2d\varepsilon''/\varepsilon')n^d + o(n^d)$ as $n \to \infty$. Theorem 3.10.15 tells us that

$$\mu(|\{x \in \Pi_n : \phi(x) > u(x) + \varepsilon'' n\}|) = o(n^d).$$

Combining the two bounds gives $\mu(|A_n^+|) \leq (2d\varepsilon''/\varepsilon')n^d + o(n^d)$. The constant ε'' may be chosen arbitrarily small, and therefore we obtain $\mu(|A_n^+|) = o(n^d)$. In the same spirit, one obtains $\mu(|A_n^-|) = o(n^d)$. This proves the claim.

Next, we construct for each $n \in \mathbb{N}$ a new measure μ_n^+, the *upper truncation* of μ. To sample from μ_n^+, first sample ϕ from μ, then replace $\phi(x)$ by $\phi_n^+(x)$ for any $x \in A_n^+$. This means that the distribution of ϕ_{Π_n} in μ_n^+ is the same as the distribution of $\phi_{\Pi_n} \wedge \phi_n^+$ in μ. Assert that

$$\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi) = \mathcal{H}_{\Pi_n}(\mu|\Phi) + o(n^d).$$

We present an alternative three-stage construction of μ_n^+, and demonstrate that the free energy changes by no more than $o(n^d)$ at every stage. Write $\mathcal{S}$ for the set of finite subsets of $\mathbb{Z}^d$, which is countable, and write α for the counting measure on $\mathcal{S}$. Write $\mathcal{G}_n$ for the smallest σ-algebra on $\Omega \times \mathcal{S}$ containing $A \times \{\Lambda\}$ for any $A \in \mathcal{F}_{\Pi_n}^\nabla$ and any $\Lambda \subset \Pi_n$.

For the first stage, write $\tilde{\mu}_n$ for the measure μ with the set A_n^+ attached to every sample $\phi \in \Omega$. The measure $\tilde{\mu}_n$ is thus a probability measure on the measurable space $(\Omega \times \mathcal{S}, \mathcal{G}_n)$. Moreover, the distribution of ϕ_{Π_n} is the same in μ as it is in $\tilde{\mu}_n$, and the set A_n^+ depends deterministically on ϕ_{Π_n}. Therefore

$$\mathcal{H}_{\Pi_n}(\mu|\Phi) = \mathcal{H}_{\mathcal{F}_{\Pi_n}^\nabla}\left(\mu \middle| e^{-H_{\Pi_n}^0}\lambda^{\Pi_n - 1}\right) = \mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n \middle| \left(e^{-H_{\Pi_n}^0}\lambda^{\Pi_n - 1}\right) \times \alpha\right). \tag{3.10.17}$$

For the second stage, introduce a new measure $\tilde{\mu}_n^+$ on $(\Omega \times \mathcal{S}, \mathcal{G}_n)$. To sample from $\tilde{\mu}_n^+$, sample first a pair (ϕ, A) from $\tilde{\mu}_n$, then replace $\phi(x)$ by $\phi_n^+(x)$ for every $x \in A$. Remark that the distribution of ϕ_{Π_n} is the same in $\tilde{\mu}_n^+$ as it is in μ_n^+. Write A' for the set $\Pi_n \smallsetminus A$. The entropies of $\tilde{\mu}_n$ and $\tilde{\mu}_n^+$ (relative to the reference measure in the final term of (3.10.17)) can be calculated in three steps. First, calculate the entropy of the choice of the set A. Second, calculate the entropy of the choice of the values of ϕ on A'. Third, calculate the entropy of the choice of the values of ϕ on A. In the construction of $\tilde{\mu}_n^+$ we only change the values of ϕ on A, and therefore the third step is the only step that produces a different entropy term. We have

$$\begin{aligned}
&\mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n^+\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right)-\mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right)\\
&\quad=\int\left(\mathcal{H}\left(\delta_{\phi_n^+|_A}\middle|e^{-H_{A,\Pi_n}(\cdot,\phi_{A'})}\lambda^A\right)-\mathcal{H}\left(\mu^{(A,\phi_{A'})}\pi_A\middle|e^{-H_{A,\Pi_n}(\cdot,\phi_{A'})}\lambda^A\right)\right)d\tilde{\mu}_n(\phi,A)\\
&\quad=\tilde{\mu}_n\left(H_{A,\Pi_n}(\phi\wedge\phi_n^+)-H_{A,\Pi_n}(\phi)\right)-\int\mathcal{H}\left(\mu^{(A,\phi_{A'})}\pi_A\middle|\lambda^A\right)d\tilde{\mu}_n(\phi,A).
\end{aligned}\tag{3.10.18}$$

In these equations, δ denotes the Dirac measure, π_A is the projection kernel onto A, and $\mu^{(A,\phi_{A'})}$ denotes the original measure μ conditioned on seeing $A_n^+=A$ and on the values of ϕ on the set A'. For the first term in (3.10.18) we observe that

$$\left|\tilde{\mu}_n\left(H_{A,\Pi_n}(\phi\wedge\phi_n^+)-H_{A,\Pi_n}(\phi)\right)\right|=O(\mu(|A_n^+|))=o(n^d);$$

this follows from the claim and (3.7.2)—noting that ϕ and $\phi\wedge\phi_n^+$ are q-Lipschitz. For $\tilde{\mu}_n$-a.e. (ϕ,A), we observe that the measure $\mu^{(A,\phi_{A'})}$ produces K-Lipschitz height functions almost surely, and consequently the same measure—restricted to A—is supported on a set of cardinality at most $(2K+1)^{|A|}$. Conclude that the second term in (3.10.18) is bounded absolutely by

$$\mu(|A_n^+|)\log(2K+1)=o(n^d).$$

To sample from μ_n^+, sample a pair (ϕ,A) from $\tilde{\mu}_n^+$, then simply forget about the set A. This is the third stage. Write ν_n for marginal of $\tilde{\mu}_n^+$ on $\mathcal{S}$. Then

$$\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi)+\mathcal{H}(\nu_n|\alpha)\leq\mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n^+\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right)\leq\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi).$$

Evidently $\mathcal{H}(\nu_n|\alpha)\leq 0$; the goal is to find a lower bound on $\mathcal{H}(\nu_n|\alpha)$. The measure ν_n is a probability measure on the set of subsets of Π_n and we also know that $\nu_n(|A|)=\mu(|A_n^+|)$. The entropy of ν_n is minimised (among all probability measures with these two properties) if one samples from ν_n by flipping a coin independently for every vertex $x\in\Pi_n$ to determine if $x\in A$. The Bernoulli parameter of the coin is $\mu(|A_n^+|)/n^d$ so that $\nu_n(|A|)=\mu(|A_n^+|)$. Write $f(p)=p\log p+(1-p)\log(1-p)$, the entropy of a Bernoulli trial with parameter p. Then the entropy of the entropy-minimising measure is $n^d f(\mu(|A_n^+|)/n^d)$. Now $\lim_{p\to 0}f(p)=0$ and therefore $\mathcal{H}(\nu_n|\alpha)=o(n^d)$. Conclude that the assertion holds true, that is,

$$\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi)=\mathcal{H}_{\Pi_n}(\mu|\Phi)+o(n^d).$$

The measure $\mu_n^\pm$ is now obtained from μ_n^+ by applying a *lower truncation*. To sample from $\mu_n^\pm$, sample first a height function ϕ from μ_n^+, then replace $\phi(x)$ by $\phi_n^-(x)$ for any $x\in A_n^-$. Alternatively, sample ϕ from μ, then replace ϕ_{Π_n} by $\phi_n^-\vee\phi_{\Pi_n}\wedge\phi_n^+$. By similar arguments as before we have

$$\mathcal{H}_{\Pi_n}(\mu_n^\pm|\Phi)=\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi)+o(n^d)=\mathcal{H}_{\Pi_n}(\mu|\Phi)+o(n^d)$$

as $n\to\infty$. The measure $\mu_n^\pm$ is supported on C_n^u. Moreover, because $\mu(|A_n^\pm|)=o(n^d)$ and because $\mu(B_n^*)\to 1$ as $n\to\infty$, we have $\mu_n^\pm(B_n)\to 1$ as $n\to\infty$. In particular, this means that the measures $\mu_n:=\mu_n^\pm(\cdot|B_n)$ are supported on $C_n^u\cap B_n$ and satisfy $\mathcal{H}_{\Pi_n}(\mu_n|\Phi)\leq\mathcal{H}_{\Pi_n}(\mu|\Phi)+o(n^d)$ as $n\to\infty$. This concludes the proof for $E=\mathbb{Z}$.

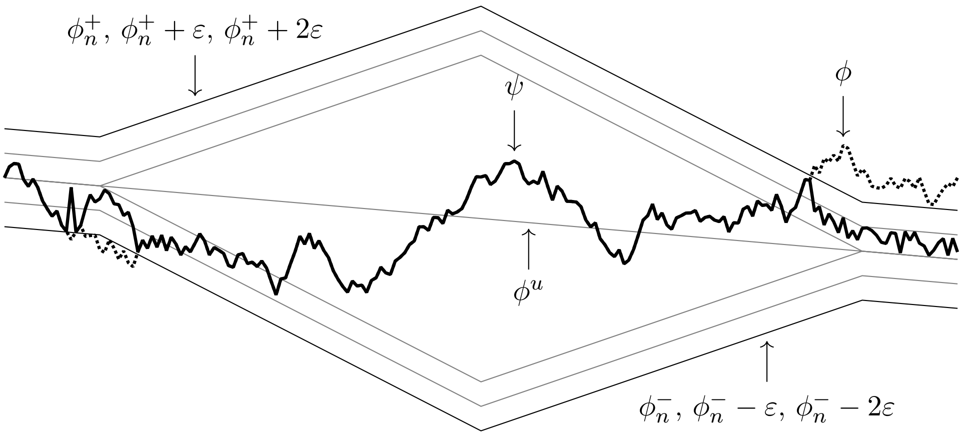


Figure 3.3: A random truncation for $E = \mathbb{R}$. The randomly truncated sample ψ remains between $\phi_n^- - 2\varepsilon$ and $\phi_n^+ + 2\varepsilon$.

The continuous case. Fix $\varepsilon > 0$ so small that ϕ^u is $q_{4\varepsilon}$-Lipschitz, and pick $B \in \mathcal{B}$ with $\mu \in B$. Assume a choice of B and B^* as for the discrete case. It suffices to find a sequence of measures $(\mu_n)_{n\in\mathbb{N}}$ with μ_n supported on $C^u_{n,2\varepsilon} \cap B_n$ and with $\mathcal{H}_{\Pi_n}(\mu_n|\Phi) \leq \mathcal{H}_{\Pi_n}(\mu|\Phi) + o(n^d)$. Write $\phi_n^\pm$ for the largest and smallest $q_{3\varepsilon}$-Lipschitz extensions of $\phi^u_{\partial^R \Pi_n}$ to Π_n respectively, for each $n \in \mathbb{N}$. Take again 0 as reference vertex for the gradient setting (as for the discrete case), and define the random sets

$$A_n^- := \{x \in \Pi_n : \phi(x) < \phi_n^-(x) - 2\varepsilon\} \quad \text{and} \quad A_n^+ := \{x \in \Pi_n : \phi(x) > \phi_n^+(x) + 2\varepsilon\}.$$

Note that $\phi_n^\pm(0) = 0$ by definition, and therefore μ-almost surely $0 \notin A_n^\pm$. Observe that $\mu(|A_n^\pm|) = o(n^d)$ by arguments identical to the case $E = \mathbb{Z}$; one can show that ϕ_n^+ is pyramid-shaped in the sense of (3.10.16) because ϕ^u is $q_{4\varepsilon}$-Lipschitz and because we chose the extension ϕ_n^+ to be the largest $q_{3\varepsilon}$-Lipschitz extension.

For each $n \in \mathbb{N}$ we construct a new measure μ_n^+ on $(\Omega, \mathcal{F}^\nabla_{\Pi_n})$, the *upper truncation* of μ. Let $(X(x))_{x\in\mathbb{Z}^d}$ be a process of i.i.d. random variables, uniformly random in the interval $[0, \varepsilon]$, in some new measure ν. To sample from μ_n^+, first sample (ϕ, X) from $\mu \times \nu$. Then, for each $x \in A_n^+$, replace $\phi(x)$ by $\phi_n^+(x) + X(x)$. Figure 3.3 displays the original function ϕ and the randomly truncated function ψ; the upper truncation is located on the right hand side, a lower truncation (which is defined at a later stage) occurs on the left. The new measure μ_n^+ is absolutely continuous with respect to $\lambda^{\Pi_n - 1}$ because we replaced each value $\phi(x)$ by a continuously distributed random variable. Assert that

$$\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi) \leq \mathcal{H}_{\Pi_n}(\mu|\Phi) + o(n^d). \tag{3.10.19}$$

Again, we present an alternative three-stage construction of μ_n^+, and we demonstrate that the entropy does not increase by more than $o(n^d)$ at every stage.

For the first stage, write $\tilde{\mu}_n$ for the measure μ with the set A_n^+ attached to every sample $\phi \in \Omega$. Define α and $\mathcal{G}_n$ as before. The measure $\tilde{\mu}_n$ is a probability measure on $(\Omega \times \mathcal{S}, \mathcal{G}_n)$. Note that (3.10.17) holds for this measure as A_n^+ depends deterministically on ϕ_{Π_n}.

For the second stage, introduce a new measure $\tilde{\mu}_n^+$ on $(\Omega \times \mathcal{S}, \mathcal{G}_n)$. To sample from $\tilde{\mu}_n^+$, sample first a triple (ϕ, A, X) from $\tilde{\mu}_n \times \nu$, then replace $\phi(x)$ by $\phi_n^+(x) + X(x)$ for every $x \in A$. Write A' for $\Pi_n \smallsetminus A$. Write ψ for the function on Π_n defined by

$\psi_A = \phi_n^+|_A + X_A$ and $\psi_{A'} = \phi_{A'}$. One calculates the entropies of $\tilde{\mu}_n$ and $\tilde{\mu}_n^+$ as in the discrete case to deduce that

$$\mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n^+\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right) - \mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right)$$
$$= \int\left(\mathcal{H}\left((\nu+\phi_n^+)\pi_A\middle|e^{-H_{A,\Pi_n}(\cdot,\phi_{A'})}\lambda^A\right) - \mathcal{H}\left(\mu^{(A,\phi_{A'})}\pi_A\middle|e^{-H_{A,\Pi_n}(\cdot,\phi_{A'})}\lambda^A\right)\right) d\tilde{\mu}_n(\phi,A)$$
$$= \tilde{\mu}_n\times\nu\left(H_{A,\Pi_n}(\psi) - H_{A,\Pi_n}(\phi)\right) - \tilde{\mu}_n(|A|)\log\varepsilon - \int\mathcal{H}\left(\mu^{(A,\phi_{A'})}\pi_A\middle|\lambda^A\right) d\tilde{\mu}_n(\phi,A).$$

In these equations, $\mu^{(A,\phi_{A'})}$ denotes the original measure μ conditioned on seeing $A_n^+ = A$ and on the values of ϕ on the set A'. By $\nu+\phi_n^+$ we simply mean the measure obtained by shifting each sample X from ν by ϕ_n^+. As in the discrete setting, the last two terms have an upper bound of order $o(n^d)$ as $n\to\infty$. It suffices to find an appropriate upper bound for the first term in the final expression.

Let $(\mathbb{A},q)$ denote the local Lipschitz constraint. By Proposition 3.6.5, it is possible to find a constant $0<\varepsilon'\leq\varepsilon$, such that for any $\{x,y\}\in\mathbb{A}$, we have

$$q_{\varepsilon'}(x,y)\geq q(x,y)-\varepsilon.$$

Claim that $\tilde{\mu}_n\times\nu$-almost surely, ψ is $q_{\varepsilon'}$-Lipschitz at every $x\in A$. In other words, we claim that

$$-q_{\varepsilon'}(y,x)\leq\psi(y)-\psi(x)\leq q_{\varepsilon'}(x,y) \tag{3.10.20}$$

whenever $x\in A$, $y\in\Pi_n$, and $\{x,y\}\in\mathbb{A}$. Suppose first that $y\in A$. The function ϕ_n^+ is $q_{3\varepsilon}$-Lipschitz and $0\leq(\psi-\phi_n^+)_{\{x,y\}}=X_{\{x,y\}}\leq\varepsilon$ for $x,y\in A$, and therefore (3.10.20) holds true with ε' replaced by ε. But $q_\varepsilon\leq q_{\varepsilon'}$, which implies (3.10.20) without said replacement. Now suppose that $y\not\in A$, so that $\psi(y)-\psi(x)=\phi(y)-\phi_n^+(x)-X(x)$. For the righthand inequality of (3.10.20) we have (almost surely)

$$\phi(y)-\phi_n^+(x)-X(x)\leq(\phi_n^+(y)+2\varepsilon)-\phi_n^+(x)\leq q_{3\varepsilon}(x,y)+2\varepsilon\leq q_\varepsilon(x,y)\leq q_{\varepsilon'}(x,y).$$

For the inequality on the left we see that (using $\phi_n^+(x)+X(x)\leq\phi(x)-\varepsilon$ for the first inequality)

$$\phi(y)-\phi_n^+(x)-X(x)\geq\phi(y)-\phi(x)+\varepsilon\geq -q(y,x)+\varepsilon\geq -q_{\varepsilon'}(y,x).$$

The middle inequality in this equation is due to the fact that ϕ is μ-almost surely q-Lipschitz. This proves the claim.

By the claim and (3.7.3), we have

$$\tilde{\mu}_n\times\nu\left(H_{A,\Pi_n}(\psi)\right)\leq O(\mu(|A_n^+|))=o(n^d).$$

For the other Hamiltonian we simply observe that

$$\tilde{\mu}_n\times\nu\left(H_{A,\Pi_n}(\phi)\right)=\tilde{\mu}_n\left(H_{A,\Pi_n}(\phi)\right)\geq -\|\Xi\|\mu(|A_n^+|)=o(n^d).$$

Putting all estimates together, we see that

$$\mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n^+\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right)\leq\mathcal{H}_{\Pi_n}(\mu|\Phi)+o(n^d).$$

To prove the original assertion, simply observe that, as in the discrete case, forgetting about the information encoded in the set A changes the entropy of $\tilde{\mu}_n^+$ by no more than $o(n^d)$:

$$\mathcal{H}_{\Pi_n}(\mu_n^+|\Phi)=\mathcal{H}_{\mathcal{G}_n}\left(\tilde{\mu}_n^+\middle|\left(e^{-H^0_{\Pi_n}}\lambda^{\Pi_n-1}\right)\times\alpha\right)+o(n^d).$$

This proves the assertion (3.10.19).

Finally one constructs a *lower truncation* $\mu_n^{\pm}$ from μ_n^+. To sample from μ_n, one first samples ϕ from μ_n^+. Then, for every $x \in A_n^-$, one resamples $\phi(x)$ independently and uniformly at random from the interval $[\phi_n^-(x) - \varepsilon, \phi_n^-(x)]$. As before, we have

$$\mathcal{H}_{\Pi_n}(\mu_n^{\pm}|\Phi) \leq \mathcal{H}_{\Pi_n}(\mu_n^+|\Phi) + o(n^d) \leq \mathcal{H}_{\Pi_n}(\mu|\Phi) + o(n^d).$$

Now $\phi_n^- - 2\varepsilon \leq \phi_{\Pi_n} \leq \phi_n^+ + 2\varepsilon$ almost surely in the measure $\mu_n^{\pm}$ and this implies in particular that

$$\phi^u_{\partial^R \Pi_n} - 2\varepsilon \leq \phi_{\partial^R \Pi_n} \leq \phi^u_{\partial^R \Pi_n} + 2\varepsilon,$$

that is, $\mu_n^{\pm}$ is supported on $C^u_{n,2\varepsilon}$. Since $\mu(B_n^*) \to 1$ and $\mu(|A_n^{\pm}|) = o(n^d)$ as $n \to \infty$, we have $\mu_n^{\pm}(B_n) \to 1$ as $n \to \infty$. This proves that the sequence $\mu_n := \mu_n^{\pm}(\cdot|B_n)$ has the desired properties. □

We now proceed as for free boundary limits, and define pinned boundary limits over other Van Hove sequences.

Definition 3.10.21. Fix $D \in \mathcal{C}$ and $m \in \mathbb{Z}_{\geq 0}$, and write Λ_n for $\Lambda^{-m}(nD) = (nD \cap \mathbb{Z}^d)^{-m}$. Consider $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$. If $E = \mathbb{Z}$, then define

$$\mathrm{PB}(\mu : D, m) := \sup_{B \in \mathcal{B} \text{ with } \mu \in B} \limsup_{n \to \infty} n^{-d} \, \mathrm{PB}_{\Lambda_n, S(\mu)}(B),$$

and if $E = \mathbb{R}$, then define

$$\mathrm{PB}(\mu : D, m) := \sup_{\varepsilon > 0 \text{ and } B \in \mathcal{B} \text{ with } \mu \in B} \limsup_{n \to \infty} n^{-d} \, \mathrm{PB}_{\Lambda_n, S(\mu), \varepsilon}(B).$$

Finally, write $\mathrm{PB}^*(\mu) := \sup_{(D,m) \in \mathcal{C} \times \mathbb{Z}_{\geq 0}} \mathrm{PB}(\mu : D, m) / \operatorname{Leb}(D)$.

It is immediate that $\mathrm{PB}^*(\mu) \geq \mathrm{PB}(\mu)$ because one can take $D = [0,1)^d$ and $m = 0$ in the supremum in this new definition. By reordering the suprema in the definitions, it is also clear that PB^* is lower-semicontinuous on the set $\{S(\cdot) = u\} \subset \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ for any u.

Lemma 3.10.22. *Consider $D \in \mathcal{C}$, $m \in \mathbb{Z}_{\geq 0}$, and $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ ergodic with $S(\mu) \in U_\Phi$. Then*

$$\mathrm{PB}(\mu : D, m) \leq \operatorname{Leb}(D) \cdot \mathcal{H}(\mu|\Phi).$$

In other words, $\mathrm{PB}^*(\mu) \leq \mathcal{H}(\mu|\Phi)$.

Proof. Write $u := S(\mu)$ and $\Lambda_n := \Lambda^{-m}(nD)$, and fix $B \in \mathcal{B}$ with $\mu \in B$. The truncation argument in the proof of Lemma 3.10.14 implies that $\mathrm{PB}_{\Lambda_n,u}(B) \leq \mathcal{H}_{\Lambda_n}(\mu|\Phi) + o(n^d)$ as $n \to \infty$ if $E = \mathbb{Z}$, and $\mathrm{PB}_{\Lambda_n,u,\varepsilon}(B) \leq \mathcal{H}_{\Lambda_n}(\mu|\Phi) + o(n^d)$ as $n \to \infty$ for any $\varepsilon > 0$ if $E = \mathbb{R}$. Therefore it suffices to demonstrate that

$$\limsup_{n \to \infty} n^{-d} \mathcal{H}_{\Lambda_n}(\mu|\Phi) \leq \operatorname{Leb}(D) \cdot \mathcal{H}(\mu|\Phi). \tag{3.10.23}$$

Without loss of generality, we suppose that $D \subset [\varepsilon, 1-\varepsilon]^d \subset [0,1)^d \subset \mathbb{R}^d$ for some $\varepsilon > 0$. Define $\Delta_n := \Pi_n \smallsetminus \Lambda_n$. Then $n^{-d}|\Delta_n| \to 1 - \operatorname{Leb}(D)$ as $n \to \infty$, and therefore Proposition 3.7.8 implies that

$$\liminf_{n \to \infty} n^{-d} \mathcal{H}_{\Delta_n}(\mu|\Phi) \geq (1 - \operatorname{Leb}(D)) \cdot \mathcal{H}(\mu|\Phi).$$

Note that (3.10.23) now follows from the fact that

$$n^{-d}(\mathcal{H}_{\Lambda_n}(\mu|\Phi) + \mathcal{H}_{\Delta_n}(\mu|\Phi)) \leq \mathcal{H}(\mu|\Phi) + o(1)$$

as $n \to \infty$. □

3.10.4 Pinned boundary limits for all μ: washboard argument

The purpose of this subsection is to demonstrate that $\mathrm{PB}^*(\mu) \leq \mathcal{H}(\mu|\Phi)$ for any shift-invariant random field μ with $S(\mu) \in U_\Phi$. The previous subsection proved this for μ ergodic. First, we demonstrate that PB^* is convex (Lemma 3.10.24)—recall that $\mathcal{H}(\cdot|\Phi)$ is affine. The idea is then to use lower-semicontinuity of PB^* in the topology of weak local convergence to derive the inequality for all non-ergodic measures (Lemma 3.10.33). Extra care must be taken whenever $E = \mathbb{Z}$, because in that case there exist ergodic measures with finite specific free energy which have their slope in ∂U_Φ rather than U_Φ. This pathology is dealt with in Lemma 3.10.32.

Lemma 3.10.24. *The functional* PB^* *is convex.*

Consider $\nu_1, \nu_2 \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with $S(\nu_1), S(\nu_2) \in U_\Phi$, and define $\mu := (1-t)\nu_1 + t\nu_2$ for some $t \in (0,1)$. If we take the value of $\mathrm{PB}^*(\nu_1)$ and $\mathrm{PB}^*(\nu_2)$ for granted, then we look for an upper bound on $\mathrm{PB}^*(\mu)$. This means that we look for asymptotic lower bounds on the integrals defining the pinned boundary estimates of μ.

The proof of the lemma uses a general strategy which produces an asymptotic lower bound on this particular integral, and which is used again twice in this chapter: in the lower bound on probabilities in the proof of the large deviations principle in Subsection 3.11.4, and when constructing the contradiction which leads to a proof of strict convexity of the surface tension in Subsection 3.12.2. The general idea is as follows: Lemma 3.10.14, and later (once it is proven) Lemma 3.10.33, provide the fundamental building blocks for the lower bounds. One then shows that these building blocks can be put together without gaining too much energy, that is, without decreasing the value of the integral of interest by too much. For this, one appeals to Theorem 3.6.9, which allows one to find suitable discrete approximations of continuous Lipschitz profiles, and the upper attachment lemma (Lemma 3.7.4), which allows one to bound the energy increase due to combining height functions defined on different parts of $\mathbb{Z}^d$. This is already sufficient to understand the macroscopic shape of the height functions. In the context of boundary limits, this is expressed through the pinning of the height functions on the boundary $\partial^R\Lambda$ of the set Λ of interest—essentially by restricting to the set C_Λ^u or $C_{\Lambda,\varepsilon}^u$. It is, however, also necessary to understand the behaviour of the local statistics of the height functions—expressed in the boundary limits through the sets B_Λ—under the operation of putting together the fundamental building blocks. For this, one appeals to the following result.

Proposition 3.10.25. *Consider some set $D \in \mathcal{C}$ and a nonnegative integer $m \in \mathbb{Z}_{\geq 0}$. Consider also some finite family $(D_i, m_i)_i \subset \mathcal{C} \times \mathbb{Z}_{\geq m}$ with the sets D_i disjoint and contained in D. Write $\Lambda_n := \Lambda^{-m}(nD)$ and $\Lambda_n^i := \Lambda^{-m_i}(nD_i)$. Then for any cylinder function $f : \Omega \to [0,1]$, we have*

$$\lim_{n\to\infty} \sup_{\phi\in\Omega} \left| L_{\Lambda_n}(\phi)(f) - \textstyle\sum_i \frac{\mathrm{Leb}(D_i)}{\mathrm{Leb}(D)} L_{\Lambda_n^i}(\phi)(f) \right| \leq \frac{\mathrm{Leb}(D \smallsetminus \cup_i D_i)}{\mathrm{Leb}(D)}.$$

Proof. Note that

$$\frac{|\mathcal{L} \cap \Lambda_n^i|}{|\mathcal{L} \cap \Lambda_n|} \to \frac{\mathrm{Leb}(D_i)}{\mathrm{Leb}(D)}$$

as $n \to \infty$, and therefore it suffices to prove the proposition for the latter fraction replaced by the former. Suppose that f is $\mathcal{F}_\Delta$-measurable for some $\Delta \subset\subset \mathbb{Z}^d$ which contains 0. Write $\mathbb{P}_n$ for the uniform probability measure on $\{\theta_x : x \in \mathcal{L} \cap \Lambda_n\}$. By coupling the measures in the obvious way, we observe that

$$L_{\Lambda_n}(\phi)(f) - \sum_i \frac{|\mathcal{L} \cap \Lambda_n^i|}{|\mathcal{L} \cap \Lambda_n|} L_{\Lambda_n^i}(\phi)(f) = \mathbb{E}_n(g_n)$$

where g_n is defined by

$$g_n(\theta) = \begin{cases} 0 & \text{if } \theta\Delta \subset \Lambda_n^i \text{ for some } i, \\ f(\theta\bar{\phi}_{\Lambda_n}) - f(\theta\bar{\phi}_{\Lambda_n^i}) & \text{if } \theta 0 \in \Lambda_n^i \text{ for some } i \text{ but } \theta\Delta \not\subset \Lambda_n^i, \\ f(\theta\bar{\phi}_{\Lambda_n}) & \text{otherwise.} \end{cases}$$

Now $|g_n| \leq 1$ and $\mathbb{P}_n(\theta\Delta \not\subset \Lambda_n^i \text{ for any } i) = \operatorname{Leb}(D \smallsetminus \cup_i D_i)/\operatorname{Leb}(D) + o(1)$ as $n \to \infty$, which implies the proposition. □

The particular proof of Lemma 3.10.24 utilises the so-called washboard construction (see Figure 3.4), which appears in the work of Sheffield [54], and is adapted here to the particular Lipschitz setting.

Proof of Lemma 3.10.24. Consider $\mu := s\nu_1 + t\nu_2$ for $s, t \in (0,1)$ with $s + t = 1$ and for some measures $\nu_1, \nu_2 \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ which have their slope in U_Φ. The goal is to prove that $\operatorname{PB}^*(\mu) \leq s\operatorname{PB}^*(\nu_1) + t\operatorname{PB}^*(\nu_2)$.

Write $u := S(\mu)$, $u_1 := S(\nu_1)$, and $u_2 := S(\nu_2)$. Consider $D \in \mathcal{C}$, $m \in \mathbb{Z}_{\geq 0}$, and $B \in \mathcal{B}$ with $\mu \in B$. Write $\Lambda_n := \Lambda^{-m}(nD)$. Fix also some $\varepsilon > 0$. If $E = \mathbb{Z}$, then we must show that

$$\limsup_{n\to\infty} n^{-d} \operatorname{PB}_{\Lambda_n,u}(B) \leq \operatorname{Leb}(D)(s\operatorname{PB}^*(\nu_1) + t\operatorname{PB}^*(\nu_2)),$$

and if $E = \mathbb{R}$, then we must show that

$$\limsup_{n\to\infty} n^{-d} \operatorname{PB}_{\Lambda_n,u,\varepsilon}(B) \leq \operatorname{Leb}(D)(s\operatorname{PB}^*(\nu_1) + t\operatorname{PB}^*(\nu_2)).$$

By choosing $\varepsilon > 0$ smaller if necessary, we suppose that $u, u_1, u_2 \in U_{q_{7\varepsilon}}$. By choosing B smaller if necessary, we suppose that B is of the form

$$B = \{\pi : |\mu(f_i) - \pi(f_i)| < 2\eta \text{ for all } i\} \in \mathcal{B}$$

for some finite family $(f_i)_i$ of continuous cylinder functions $f_i : \Omega \to [0,1]$ and for some $\eta > 0$, and we write

$$B^j := \{\pi : |\nu_j(f_i) - \pi(f_i)| < \eta \text{ for all } i\} \in \mathcal{B}$$

for $j \in \{1,2\}$.

The idea of the proof is roughly as follows. First, we partition a large subset of D into finitely many convex shapes. Second, we find a continuous Lipschitz function f which equals u on ∂D, and which is affine on each convex shape in this partition, with slope either u_1 or u_2. This function is chosen such that the Lebesgue measure of the convex shapes with slope u_j is roughly $s\operatorname{Leb}(D)$ for $j = 1$ and roughly $t\operatorname{Leb}(D)$ for $j = 2$. Informally, the function f looks like a "washboard". Next, we define $f_n := nf(\cdot/n)$, and use the existence of the function f_n and Theorem 3.6.9 to find for each $n \in \mathbb{N}$ a corresponding height function ϕ_n. The existence of the height function

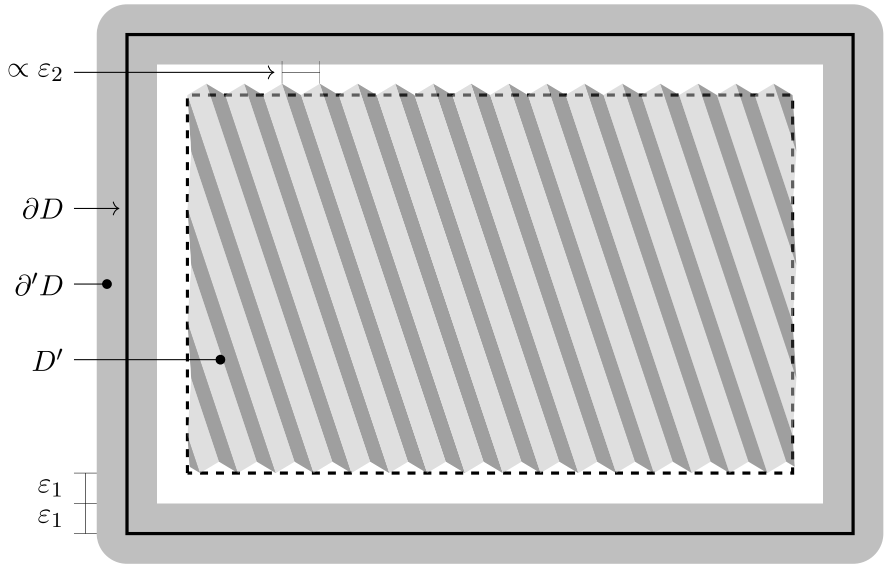


Figure 3.4: The washboard and the function $f : \partial' D \cup D' \to \mathbb{R}$

ϕ_n and the previously described general strategy allow us to build a direct comparison between $\mathrm{PB}_{\Lambda_n,u}(B)$ or $\mathrm{PB}_{\Lambda_n,u,\varepsilon}(B)$ and the numbers $\mathrm{PB}^*(\nu_1)$ and $\mathrm{PB}^*(\nu_2)$.

We start by constructing the continuous "washboard"—see Figure 3.4. Set $v := u_1 - u_2$ if $u_1 \neq u_2$, and choose $v \in (\mathbb{R}^d)^* \smallsetminus \{0\}$ arbitrary otherwise. Define

$$w : \mathbb{R}^d \to \mathbb{Z},\ x \mapsto \begin{cases} 2\lfloor v(x) \rfloor & \text{if } v(x) - \lfloor v(x) \rfloor \in [0, t), \\ 2\lfloor v(x) \rfloor + 1 & \text{if } v(x) - \lfloor v(x) \rfloor \in [t, 1). \end{cases}$$

Write $p : \mathbb{R}^d \to \mathbb{R}$ for the unique continuous function that maps 0 to 0, and which has gradient u_j on the interior of $\{w \in 2\mathbb{Z} + j\} \subset \mathbb{R}^d$ for $j \in \{1, 2\}$. For $\alpha > 0$, write w_α for the map $w_\alpha(\cdot) := w(\cdot/\alpha)$, and write p_α for the map $p_\alpha(\cdot) := \alpha p(\cdot/\alpha)$. It is straightforward to see that p_α has gradient u_j on $\{w_\alpha \in 2\mathbb{Z} + j\}$ for $j \in \{1, 2\}$, and that $\|p_\alpha - u\|_\infty \propto \alpha$. Observe also that p and p_α are $\|\cdot\|_{q_{7\varepsilon}}$-Lipschitz.

In the remainder of the proof, we shall work with three limits. First we take $n \to \infty$, then $\varepsilon_2 \to 0$, then $\varepsilon_1 \to 0$. Reference to these variables is sometimes omitted for brevity. Define

$$\begin{aligned} \partial' D &:= \{x \in \mathbb{R}^d : d_2(x, \partial D) < \varepsilon_1\}, \\ D' &:= \{x \in D : d_2(x, \partial D) > 2\varepsilon_1\}, \\ D'_k &:= D' \cap \{w_{\varepsilon_2} = k\}, \end{aligned}$$

where d_2 denotes Euclidean distance. Write $f : \partial' D \cup D' \to \mathbb{R}$ for the function defined by

$$f(x) := \begin{cases} u(x) & \text{if } x \in \partial' D, \\ p_{\varepsilon_2}(x) & \text{if } x \in D'. \end{cases}$$

This function is $\|\cdot\|_{q_{6\varepsilon}}$-Lipschitz for ε_2 sufficiently small, depending on ε_1. Note that f is affine with gradient u_1 on D'_k for k odd and with gradient u_2 on D'_k for k even. Moreover, the family $(D'_k)_{k \in \mathbb{Z}}$ is a partition of D'. Only finitely many members are nonempty, and the nonempty members are convex, bounded, and have positive

Lebesgue measure. The merit of this construction is that

$$\operatorname{Leb}(\cup_{k\in 2\mathbb{Z}+1}D'_k) \to_{\varepsilon_2\to 0} s\operatorname{Leb}(D') \to_{\varepsilon_1\to 0} s\operatorname{Leb}(D), \tag{3.10.26}$$
$$\operatorname{Leb}(\cup_{k\in 2\mathbb{Z}}D'_k) \quad \to_{\varepsilon_2\to 0} t\operatorname{Leb}(D') \to_{\varepsilon_1\to 0} t\operatorname{Leb}(D). \tag{3.10.27}$$

For $n \in \mathbb{N}$, define $f_n : n(\partial' D \cup D') \to \mathbb{R}$ by $f_n(\cdot) := nf(\cdot/n)$—this function is also $\|\cdot\|_{q_{6\varepsilon}}$-Lipschitz. In particular, Theorem 3.6.9 implies that for some $M \in \mathbb{Z}_{\geq m}$ depending only on ε, there exists a q-Lipschitz height function $\phi_n \in \Omega$ such that

1. $\nabla\phi_n|_{\Lambda^{-M}(n\partial' D)} = \nabla\phi^u|_{\Lambda^{-M}(n\partial' D)}$,
2. $\nabla\phi_n|_{\Lambda^{-M}(nD'_k)} = \nabla\phi^{u_1}|_{\Lambda^{-M}(nD'_k)}$ for all k odd,
3. $\nabla\phi_n|_{\Lambda^{-M}(nD'_k)} = \nabla\phi^{u_2}|_{\Lambda^{-M}(nD'_k)}$ for all k even,
4. ϕ_n is $q_{5\varepsilon}$-Lipschitz if $E = \mathbb{R}$.

Recall the definition of Λ_n, and define

$$\Lambda_{n,k} := \Lambda^{-M}(nD'_k), \qquad \Lambda_n^0 := \Lambda_n \smallsetminus \cup_k(\Lambda_{n,k} \smallsetminus \{0_{\Lambda_{n,k}}\}), \qquad \Lambda_n^* := \Lambda_n \smallsetminus \cup_k \Lambda_{n,k}^{-R}.$$

Note that $\partial^R\Lambda_n \subset \Lambda^{-M}(n\partial' D)$ for n sufficiently large, and consequently $\nabla\phi_n|_{\partial^R\Lambda_n} = \nabla\phi^u|_{\partial^R\Lambda_n}$. This also implies that the sets $\partial^R\Lambda_n$ and $\Lambda_{n,k}$ are all disjoint for fixed n as k ranges over $\mathbb{Z}$. Finally, $\Lambda_{n,k} \subset \Lambda_n$ for all k.

The idea is now to use the existence of the function ϕ_n to derive the inequalities. We distinguish two cases, depending on whether $E = \mathbb{Z}$ or $E = \mathbb{R}$. Start with the former, which is easier. Write A_n for the set of height functions ϕ such that

1. $\nabla\phi$ equals $\nabla\phi_n$ on Λ_n^*,
2. $\phi \in B^1_{\Lambda_{n,k}}$ for all k odd,
3. $\phi \in B^2_{\Lambda_{n,k}}$ for all k even.

Note that $A_n \subset C^u_{\Lambda_n}$ because $\partial^R\Lambda_n \subset \Lambda_n^*$ and because $\nabla\phi_n = \nabla\phi^u$ on $\partial^R\Lambda_n$. It is straightforward to work out that $A_n \subset B_{\Lambda_n}$ for n sufficiently large and $\varepsilon_1, \varepsilon_2$ sufficiently small, by application of Proposition 3.10.25 combined with (3.10.26) and (3.10.27).

Therefore it suffices to demonstrate that

$$\liminf n^{-d} \log \int_{A_n} e^{-H^0_{\Lambda_n}} d\lambda^{\Lambda_n - 1} \geq -\operatorname{Leb}(D)(s\operatorname{PB}^*(\nu_1) + t\operatorname{PB}^*(\nu_2)) \tag{3.10.28}$$

where the limit is in the variables n, ε_2, and ε_1. Moreover, since $\nabla\phi$ equals $\nabla\phi_n$ on Λ_n^* for any $\phi \in A_n$, this restriction to Λ_n^* is q-Lipschitz, and the upper attachment lemma (Lemma 3.7.4) implies that

$$H^0_{\Lambda_n} \leq \sum_k H^0_{\Lambda_{n,k}} + \sum_k e^+(\Lambda_{n,k}) + |\Lambda_n \smallsetminus \cup_k \Lambda_{n,k}| \max_{x\in\mathbb{Z}^d/\mathcal{L}} e^+(\{x\}) \tag{3.10.29}$$

on A_n. For the third term we have $n^{-d}|\Lambda_n \smallsetminus \cup_k\Lambda_{n,k}| \to_{n\to\infty} \operatorname{Leb}(D \smallsetminus D') \to_{\varepsilon_1\to 0} 0$, and the second term is of order $o(n^d)$ as $n \to \infty$. This implies that

$$\liminf n^{-d} \log \int_{A_n} e^{-H^0_{\Lambda_n}} d\lambda^{\Lambda_n - 1} \geq \liminf n^{-d} \log \int_{A_n} e^{-\sum_k H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_n - 1}. \tag{3.10.30}$$

Recall the definition of Λ_n^0, and consider λ^{Λ_n-1} a product measure, by writing

$$\lambda^{\Lambda_n-1} := \lambda^{\Lambda_n^0-1} \times \prod_k \lambda^{\Lambda_{n,k}-1}.$$

Note that A_n contains exactly all height functions ϕ such that

1. $\nabla\phi$ equals $\nabla\phi_n$ on Λ_n^0,
2. $\phi \in C^{u_1}_{\Lambda_{n,k}} \cap B^1_{\Lambda_{n,k}}$ for all k odd,
3. $\phi \in C^{u_2}_{\Lambda_{n,k}} \cap B^2_{\Lambda_{n,k}}$ for all k even,

and therefore

$$\int_{A_n} e^{-\sum_k H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_n-1} = \int_{\{\nabla\phi \text{ equals } \nabla\phi_n \text{ on } \Lambda_n^0\}} d\lambda^{\Lambda_n^0-1}(\phi)$$
$$\cdot \prod_{k\in 2\mathbb{Z}+1} \int_{C^{u_1}_{\Lambda_{n,k}} \cap B^1_{\Lambda_{n,k}}} e^{-H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_{n,k}-1} \cdot \prod_{k\in 2\mathbb{Z}} \int_{C^{u_2}_{\Lambda_{n,k}} \cap B^2_{\Lambda_{n,k}}} e^{-H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_{n,k}-1}.$$

The first factor equals one since we are dealing with the counting measure, and therefore

$$\log \int_{A_n} e^{-\sum_k H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_n-1} = -\sum_{k\in 2\mathbb{Z}+1} \mathrm{PB}_{\Lambda_{n,k},u_1}(B^1) - \sum_{k\in 2\mathbb{Z}} \mathrm{PB}_{\Lambda_{n,k},u_2}(B^2).$$

For fixed $\varepsilon_1, \varepsilon_2$ only finitely many terms are possibly nonzero—those corresponding to nonempty sets D_k'—and for each term we have (for $j \in \{1,2\}$)

$$\limsup_{n\to\infty} n^{-d}\, \mathrm{PB}_{\Lambda_{n,k},u_j}(B^j) \leq \mathrm{PB}(\nu_j : D_k', M) \leq \mathrm{Leb}(D_k')\, \mathrm{PB}^*(\nu_j).$$

Therefore (3.10.26) and (3.10.27) imply

$$\liminf n^{-d} \log \int_{A_n} e^{-H^0_{\Lambda_n}} d\lambda^{\Lambda_n-1} \geq -\mathrm{Leb}(D)(s\, \mathrm{PB}^*(\nu_1) + t\, \mathrm{PB}^*(\nu_2)),$$

the desired inequality.

Let us now discuss what changes for $E = \mathbb{R}$. Write A_n for the set of samples ϕ such that

1. $|(\phi_{\Lambda_n^0} - \phi(0_{\Lambda_n})) - (\phi_n|_{\Lambda_n^0} - \phi_n(0_{\Lambda_n}))| \leq \varepsilon$,
2. $\phi \in C^{u_1}_{\Lambda_{n,k},\varepsilon}$ and $\phi \in B^1_{\Lambda_{n,k}}$ for all k odd,
3. $\phi \in C^{u_2}_{\Lambda_{n,k},\varepsilon}$ and $\phi \in B^2_{\Lambda_{n,k}}$ for all k even.

Note that $A_n \subset C^u_{\Lambda_n,\varepsilon}$. The proof that $A_n \subset B_{\Lambda_n}$ is the same as before. We must again prove (3.10.28). The definition of A_n implies that $|(\phi_{\Lambda_n^*} - \phi(0_{\Lambda_n})) - (\phi_n|_{\Lambda_n^*} - \phi_n(0_{\Lambda_n}))| \leq 2\varepsilon$ for any $\phi \in A_n$, which in turn implies that $\phi_{\Lambda_n^*}$ is q_ε-Lipschitz as ϕ_n was $q_{5\varepsilon}$-Lipschitz—see Proposition 3.6.5. Therefore (3.10.29) holds true with $e^+(\cdot)$ replaced by $e^+_\varepsilon(\cdot)$, which implies (3.10.30). We now have

$$\int_{A_n} e^{-\sum_k H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_n-1} = \int_{\{|(\phi_{\Lambda_n^0} - \phi(0_{\Lambda_n})) - (\phi_n|_{\Lambda_n^0} - \phi_n(0_{\Lambda_n}))| \leq \varepsilon\}} d\lambda^{\Lambda_n^0-1}(\phi)$$
$$\cdot \prod_{k\in 2\mathbb{Z}+1} \int_{C^{u_1}_{\Lambda_{n,k},\varepsilon} \cap B^1_{\Lambda_{n,k}}} e^{-H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_{n,k}-1} \cdot \prod_{k\in 2\mathbb{Z}} \int_{C^{u_2}_{\Lambda_{n,k},\varepsilon} \cap B^2_{\Lambda_{n,k}}} e^{-H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_{n,k}-1}.$$

The first integral equals $(2\varepsilon)^{|\Lambda_n^0|-1}$, and therefore

$$\log \int_{A_n} e^{-\sum_k H^0_{\Lambda_{n,k}}} d\lambda^{\Lambda_n - 1} \\ = (|\Lambda_n^0| - 1)\log(2\varepsilon) - \sum_{k\in 2\mathbb{Z}+1} \mathrm{PB}_{\Lambda_{n,k},u_1,\varepsilon}(B^1) - \sum_{k\in 2\mathbb{Z}} \mathrm{PB}_{\Lambda_{n,k},u_2,\varepsilon}(B^2).$$

The first term vanishes in the limit in the three variables after normalising by n^{-d}. The remainder of the proof is the same as before. $\square$

Let us now discuss briefly how to deal with ergodic measures with finite specific free energy which have their slope in ∂U_Φ, before proving that $\mathrm{PB}^*(\mu) \leq \mathcal{H}(\mu|\Phi)$ for any shift-invariant random field μ with $S(\mu) \in U_\Phi$.

Definition 3.10.31. Consider a measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with finite specific free energy. Classify μ as *taut* if w_μ-almost surely $S(\nu) \in \partial U_\Phi$, and as *non-taut* if w_μ-almost surely $S(\nu) \in U_\Phi$. A *non-taut approximation* of μ is a sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ of non-taut measures such that $\mathcal{H}(\mu_n|\Phi) \to \mathcal{H}(\mu|\Phi)$ and $\mu_n \to \mu$ in the topology of weak local convergence as $n \to \infty$.

If $E = \mathbb{R}$ and μ a shift-invariant random field with finite specific free energy, then w_μ-almost surely $S(\nu) \in U_\Phi$, due to Theorem 3.4.13 and because $\mathcal{H}(\cdot|\Phi)$ is strongly affine. In other words, μ is automatically non-taut. The following lemma is therefore meaningful for $E = \mathbb{Z}$ only.

Lemma 3.10.32. *Any ergodic gradient random field with finite specific free energy has a non-taut approximation.*

Proof. Let $E = \mathbb{Z}$, and let μ denote an ergodic random field with $\mathcal{H}(\mu|\Phi) < \infty$ and $S(\mu) \in \partial U_\Phi$. In this pathological case, we must modify μ slightly, so that the modified measure is non-taut, and without changing the specific free energy too much. Let ξ denote another ergodic measure in $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with $\mathcal{H}(\xi|\Phi) < \infty$ and with $S(\xi) \in U_\Phi$—such measures exist, due to the proof of Theorem 3.4.13 on Page 95. Write ρ_n for the uniform probability measure on the set $\{0, \dots, n-1\}$. Fix $n \in \mathbb{N}$; we are going to define a new measure μ_n. To sample a height function ϕ from μ_n, sample first a triple (ϕ^μ, ϕ^ξ, a) from the measure $\mu \times \xi \times \rho_n$. The final sample ϕ is then given by the equation

$$\phi := \phi^\mu - \left\lfloor \frac{(\phi^\mu - \phi^\mu(0)) - (\phi^\xi - \phi^\xi(0)) + a}{n} \right\rfloor.$$

The random choice of a makes the rounding operation shift-invariant. Note that the numerator in this fraction is $2K$-Lipschitz almost surely for K minimal subject to $Kd_1 \geq q$, and therefore the rounded function is 1-Lipschitz for n sufficiently large. In fact, the density of edges on which the rounded function is not constant, has a bound of order $O(1/n)$ as $n \to \infty$. In particular, this implies that $\mu_n \to \mu$ in the topology of (weak) local convergence. Recall (3.9.6) from the proof of Theorem 3.9.4, and observe that the specific free energy of μ and μ_n can be calculated as in this equation because either measure is K-Lipschitz. If $f(p)$ denotes the entropy function of a Bernoulli trial with parameter p as in the proof of Lemma 3.10.14, then by arguments similar to those used in that proof, we can bound the difference in the specific entropy between μ and μ_n:

$$|\mathcal{H}(\mu|\lambda) - \mathcal{H}(\mu_n|\lambda)| = O(f(O(1/n))) = o(1)$$

as $n \to \infty$. For $E = \mathbb{Z}$, we have a lower and upper bound on $H_{\{x\}}(\phi)$ for q-Lipschitz ϕ, and this and amenability of the weak interaction Ξ imply that the specific energy functional

$$\mu \mapsto \mu(\Phi)$$

is continuous with respect to the topology of local convergence whenever restricted to shift-invariant random fields which are supported on q-Lipschitz functions. Jointly these two observations imply that $\mathcal{H}(\mu_n|\Phi) \to \mathcal{H}(\mu|\Phi)$. It suffices to demonstrate that each measure μ_n is non-taut. Claim that w_{μ_n}-almost every ergodic component ν satisfies $S(\nu) = (1 - \frac{1}{n})S(\mu) + \frac{1}{n}S(\xi) \in U_\Phi$. Recall Theorem 3.10.15. The final assertion of that theorem tells us that the slope $S(\nu)$ of each ergodic component can be read off from almost every sample ϕ from ν, since the slope $u := S(\nu)$ is almost surely the unique slope such that for any fixed $\varepsilon > 0$,

$$\|\phi_{\Pi_m} - \phi(0) - u|_{\Pi_m}\|_\infty \leq \varepsilon m$$

for m sufficiently large. The slope $(1 - \frac{1}{n})S(\mu) + \frac{1}{n}S(\xi)$ makes this inequality work for samples ϕ from the original measure μ_n, because μ and ξ are ergodic, and because ϕ equals $(1 - \frac{1}{n})\phi^\mu + \frac{1}{n}\phi^\xi$ up to bounded differences. ∎

Lemma 3.10.33. *For any $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with $S(\mu) \in U_\Phi$, we have $\mathrm{PB}^*(\mu) \leq \mathcal{H}(\mu|\Phi)$.*

Proof. Let μ denote an arbitrary shift-invariant random field with $H := \mathcal{H}(\mu|\Phi) + 1 < \infty$ and $u := S(\mu) \in U_\Phi$. If μ is non-taut and a convex combination of finitely many ergodic random fields, then the lemma follows immediately from Theorem 3.4.1 and Lemmas 3.10.22 and 3.10.24. Let us now consider the case that μ is non-taut, but not a convex combination of finitely many ergodic random fields. The lower level set of the specific free energy M_H is a compact Polish space, and therefore there exists a sequence of continuous cylinder functions $(f_k)_{k\in\mathbb{N}}$ with $f_k : \Omega \to [0,1]$ such that some sequence $(\mu_n)_{n\in\mathbb{N}} \subset M_H$ satisfies $\mu_n \to \mu$ in the topology of weak local convergence if and only if $\mu_n(f_k) \to \mu(f_k)$ as $n \to \infty$ for every $k \in \mathbb{N}$. Write w_μ for the ergodic decomposition of μ. Let $(\nu_i)_{i\in\mathbb{N}}$ denote an i.i.d. sequence of samples from w_μ. Define

$$\mu_n := \sum\nolimits_{i=1}^n \tfrac{1}{n}\nu_i.$$

Then w_μ-almost surely, $\mathcal{H}(\mu_n|\Phi) \to \mathcal{H}(\mu|\Phi)$ and $\mu_n(f_k) \to \mu(f_k)$ as $n \to \infty$ for all $k \in \mathbb{N}$. This implies that $\mu_n \to \mu$ in the topology of weak local convergence. Finally, we have $S(\mu_n) \to u$. By altering the coefficients in the definition of each measure μ_n slightly, we can make sure that $S(\mu_n) = u$ for n sufficiently large, while retaining the other properties of this sequence. For each measure μ_n we have $\mathrm{PB}^*(\mu_n) \leq \mathcal{H}(\mu_n|\Phi)$ by the first part of this proof, and $\mathrm{PB}^*(\mu) \leq \mathcal{H}(\mu|\Phi)$ because $\mathrm{PB}^*(\cdot)$ is lower-semicontinuous when restricted to $\{S(\cdot) = u\}$, while $\mathcal{H}(\mu_n|\Phi) \to \mathcal{H}(\mu|\Phi)$ as $n \to \infty$.

We now prove the lemma for the case that μ is a convex combination of finitely many ergodic measures, but without imposing that μ is non-taut. Write

$$\mu = \sum\nolimits_{i=1}^n a_i \nu^i$$

for the decomposition of μ into ergodic components. Since each ν^i is ergodic, it has a non-taut approximation $(\nu_k^i)_{k\in\mathbb{N}}$. Define $\mu_k := \sum_{i=1}^n a_i \nu_k^i$, so that $\mu_k \to \mu$ in the topology of weak local convergence with $\mathcal{H}(\mu_k|\Phi) \to \mathcal{H}(\mu|\Phi)$ as $k \to \infty$. This implies also that $S(\mu_k) \to S(\mu)$, and by altering the coefficients in the definition of each measure μ_k slightly, we may ensure that $S(\mu_k) = u$ for k sufficiently large, while retaining the previously mentioned properties. By arguing as before, we have $\mathrm{PB}^*(\mu_k) \leq \mathcal{H}(\mu_k|\Phi)$ and therefore $\mathrm{PB}^*(\mu) \leq \mathcal{H}(\mu|\Phi)$. The generalisation to those measures μ which are not a convex combination of finitely many ergodic measures and not non-taut is the same as before. □

3.11 Large deviations principle

Large deviations are the subject of a vast literature within statistical physics [12, 11, 49]. In the context of gradient models, the pioneering result was derived by Sheffield in [54]. In this section we prove a large deviations principle (LDP) of similar strength to the one contained in Chapter 7 of [54], with the noteworthy difference that we express it directly in terms of the Gibbs specification. The large deviations principle applies to all models described in the introduction, including for example perturbed dimer models [21, 22] which are not monotone, even if the perturbation has infinite range. This LDP captures both the macroscopic profile of each sample, as well as its local statistics. We will be using some notations and ideas from [54] and [36]. Recall Subsection 3.4.2 for a description of good asymptotic boundary profiles and good approximations. That subsection also contains a description of the topology for the macroscopic profile of each function. The letter Φ denotes a fixed potential belonging to the class $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ throughout this section.

3.11.1 Formal description of the LDP

Recall Subsection 3.4.2, which gave an overview of the large deviations principle without local statistics. Throughout this section, the sequence $(D_n, b_n)_{n\in\mathbb{N}}$ denotes a good approximation of some fixed good asymptotic boundary profile (D, b). The sequence of local Gibbs measures which are of interest in the LDP is the sequence $(\gamma_n)_{n\in\mathbb{N}}$ defined by $\gamma_n := \gamma_{D_n}(\cdot, b_n)$. We shall also write Z_n for $Z_{D_n}(b_n)$; the normalising constant in the definition of the measure $\gamma_{D_n}(\cdot, b_n)$. Finally, $\tilde{\gamma}_n$ shall denote the non-normalised version of γ_n, that is, $\tilde{\gamma}_n := Z_n\gamma_n$.

The topological space

All samples from the sequence of measures $(\gamma_n)_{n\in\mathbb{N}}$ must be brought to the same topological space, in order to formulate the large deviations principle. We want our large deviations principle to describe both the global profile of each sample as well as its local statistics, and this is reflected in the choice of topological space. More concretely, the topological space that we have in mind decomposes as the product of two topological spaces, each describing one of the two aspects of each sample. Recall from Definition 3.4.8 that $(\mathrm{Lip}(\bar{D}), \mathcal{X}^\infty)$ is space of $K\|\cdot\|_1$-Lipschitz functions functions on $\bar{D}$ endowed with the topology of uniform convergence. Recall also the definition of $\mathfrak{G}_n$; each map $\mathfrak{G}_n$ is used to map samples from γ_n to $\mathrm{Lip}(\bar{D})$. This map characterises the macroscopic profile of each sample.

Next, we define the empirical measure profile $\mathfrak{L}_n(\phi)$ of the sample ϕ from γ_n. The empirical measure profile captures the local statistics of the height function ϕ in the large deviations principle.

Definition 3.11.1 (topology for local statistics)**.** Write $\mathcal{D}$ for the Borel σ-algebra on D, and recall that $\mathcal{M}(X, \mathcal{X})$ denotes the set of σ-finite measures on the measurable space $(X, \mathcal{X})$. Throughout this chapter, we shall write $\mathcal{M}^D$ for the set of measures $\mu \in \mathcal{M}(D \times \Omega, \mathcal{D} \times \mathcal{F}^\nabla)$ which have the property that the first marginal $\mu_D = \mu(\cdot, \Omega)$ equals the Lebesgue measure on D. The empirical profile $\mathfrak{L}_n(\phi) \in \mathcal{M}^D$ of ϕ is now defined by the equation

$$\mathfrak{L}_n(\phi) := \int_D \delta_{(x, \theta_{[nx]_\mathcal{L}} \phi)} dx,$$

where δ denotes the Dirac measure and $[nx]_\mathcal{L}$ is the vertex in $\mathcal{L}$ closest to nx in the Euclidean metric—this is well-defined for almost every x with respect to the Lebesgue measure. Thus, to "sample" from $\mathfrak{L}_n(\phi)$—this language is abusive because the size of the measure $\mathfrak{L}_n(\phi)$ is $\operatorname{Leb}(D)$ and therefore not generally a probability measure—one first samples x from D uniformly at random; then one shifts the sample ϕ by $[nx]_\mathcal{L}$. The map $\mathfrak{L}_n : \Omega \to \mathcal{M}^D$ thus captures the local statistics of the height functions in the large deviations principle. For the statement of the large deviations principle, we endow the space $\mathcal{M}^D$ with the topology $\mathcal{X}^\mathfrak{L}$. This is defined to be the weakest topology which makes the map $\mu \mapsto \mu(R, f)$ continuous for any rectangular subset R of D, and for any continuous cylinder function $f : \Omega \mapsto [0, 1]$.

Remark. If ϕ is a height function, $R \subset D$ a bounded convex set of positive Lebesgue measure, and n large, then

$$\operatorname{Leb}(R)^{-1} \mathfrak{L}_n(\phi)(R, \cdot) \approx L_{\Lambda(nR)}(\phi).$$

More precisely, the total variation distance between the two measures goes to zero as $n \to \infty$, uniformly over the choice of ϕ.

Definition 3.11.2 (Product topology for the large deviation principle)**.** The large deviations principle is formulated on the space $X^\mathfrak{P} := \operatorname{Lip}(\bar{D}) \times \mathcal{M}^D$ endowed with the topology $\mathcal{X}^\mathfrak{P} := \mathcal{X}^\infty \times \mathcal{X}^\mathfrak{L}$, and we map each sample ϕ from γ_n to this space by applying the map $\mathfrak{P}_n := \mathfrak{G}_n \times \mathfrak{L}_n$.

The rate function

Before proceeding, a few definitions for measures $\mu \in \mathcal{M}^D$ are introduced. The measure μ is called *$\mathcal{L}$-invariant* if $\operatorname{Leb}(U)^{-1} \mu(U, \cdot) \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ for any $U \in \mathcal{D}$ of positive Lebesgue measure. Write $\mathcal{M}^D_\mathcal{L}$ for the set of all such shift-invariant measures. If μ is $\mathcal{L}$-invariant and $U \in \mathcal{D}$ has positive Lebesgue measure, then we write $S(\mu(U, \cdot))$ for the slope of $\operatorname{Leb}(U)^{-1} \mu(U, \cdot)$. Call a pair $(g, \mu) \in X^\mathfrak{P}$ *compatible*, and write $g \sim \mu$, if μ is $\mathcal{L}$-invariant with $\nabla g(x) = S(\mu(x, \cdot))$ as a distribution on D. Finally, write w_μ for the ergodic decomposition of the shift-invariant non-normalised measure $\mu(D, \cdot)$, and define

$$\mathcal{H}(\mu|\Phi) := \mathcal{H}(\mu(D, \cdot)|\Phi) := \int \mathcal{H}(\nu|\Phi) dw_\mu(\nu) = \operatorname{Leb}(D) \mathcal{H}(\operatorname{Leb}(D)^{-1} \mu(D, \cdot)|\Phi).$$

Definition 3.11.3. Consider a good asymptotic boundary profile (D, b). The *rate function* associated to this profile is the function $I : X^\mathfrak{P} \to \mathbb{R} \cup \{\infty\}$ defined by

$$I(g, \mu) := \tilde{I}(g, \mu) - P_\Phi(D, b) \quad \text{where} \quad \tilde{I}(g, \mu) := \begin{cases} \mathcal{H}(\mu|\Phi) & \text{if } g|_{\partial D} = b \text{ and } g \sim \mu, \\ \infty & \text{otherwise.} \end{cases}$$

Here $P_\Phi(D, b)$ denotes the *pressure* of (D, b), which is given by

$$P_\Phi(D, b) := \min_{g \in \operatorname{Lip}(\bar{D}) \text{ with } g|_{\partial D} = b} \int_D \sigma(\nabla g(x)) dx.$$

The function $\tilde{I}$ is useful because its definition does not appeal to the pressure. It will later appear as the rate function of the LDP corresponding to the sequence of measures $(\tilde{\gamma}_n)_{n\in\mathbb{N}}$ defined by $\tilde{\gamma}_n := Z_n\gamma_n$, the non-normalised versions of the local Gibbs measures $\gamma_{D_n}(\cdot, b_n)$.

Lemma 3.11.4. *The following hold true:*

1. *The rate functions I and $\tilde{I}$ are convex,*
2. *The rate functions I and $\tilde{I}$ are lower-semicontinuous,*
3. *The lower level sets $\{I \leq C\}$ and $\{\tilde{I} \leq C\}$ are compact Polish spaces for $C < \infty$,*
4. *There is a probability kernel $u \mapsto \mu_u$ such that for any $u \in \{\sigma < \infty\}$, we have $\mu_u \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ with $S(\mu_u) = u$ and $\mathcal{H}(\mu_u|\Phi) = \sigma(u)$,*
5. *For fixed $g \in \mathrm{Lip}(\bar{D})$ with $g|_{\partial D} = b$, we have*
$$\min_{\mu\in\mathcal{M}^D} \tilde{I}(g, \mu) = \int_D \sigma(\nabla g(x))dx,$$
6. *The minimum of I is 0, and the minimum of $\tilde{I}$ is $P_\Phi(D, b)$.*

We provide a proof in the next subsection.

Statement of the LDP

Theorem 3.11.5 (Large deviations principle)**.** *Let $\Phi \in \mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$, and let $(D_n, b_n)_{n\in\mathbb{N}}$ denote a good approximation of some good asymptotic profile (D, b). Let γ_n^* denote the pushforward of $\gamma_n := \gamma_{D_n}(\cdot, b_n)$ along the map $\mathfrak{P}_n$, for any $n \in \mathbb{N}$. Then the sequence of probability measures $(\gamma_n^*)_{n\in\mathbb{N}}$ satisfies a large deviations principle with speed n^d and rate function I on the topological space $(X^{\mathfrak{P}}, \mathcal{X}^{\mathfrak{P}})$. Moreover, the sequence of normalising constants $(Z_n)_{n\in\mathbb{N}} := (Z_{D_n}(b_n))_{n\in\mathbb{N}}$ satisfies $-n^{-d}\log Z_n \to P_\Phi(D, g)$ as $n \to \infty$.*

Remark that Theorem 3.4.10 follows immediately from this theorem in combination with Lemma 3.11.4, Statement 5.

3.11.2 Proof overview

Let us start with a proof of some key properties of the rate functions I and $\tilde{I}$.

Proof of Lemma 3.11.4. Note that $\mathcal{M}_{\mathcal{L}}^D$ is closed in $(\mathcal{M}^D, \mathcal{X}^{\mathfrak{L}})$. It follows immediately from the properties of the original specific free energy functional (Theorem 3.4.1) that the map
$$\mathcal{H}(\cdot|\Phi) : \mathcal{M}_{\mathcal{L}}^D \to \mathbb{R} \cup \{\infty\}$$
is affine and lower-semicontinuous, and that its lower level sets are compact Polish spaces with respect to the $\mathcal{X}^{\mathfrak{L}}$-topology.

Observe that the set $\{g \sim \mu\} \subset X^{\mathfrak{P}}$ is convex. This implies that I and $\tilde{I}$ are convex, since the map $(g, \mu) \mapsto \mathcal{H}(\mu|\Phi)$ is affine on $\{g \sim \mu\}$. Observe that the set

$\{g \sim \mu\}$ is also closed in $\mathcal{X}^{\mathfrak{P}}$. The lower level sets of I and $\tilde{I}$ are compact Polish spaces because

$$\{\tilde{I} \leq C\} = (\{g \in \mathrm{Lip}(\bar{D}) : g|_{\partial D} = b\} \times \{\mu \in \mathcal{M}_{\mathcal{L}}^D : \mathcal{H}(\mu|\Phi) \leq C\}) \cap \{g \sim \mu\},$$

that is, $\{\tilde{I} \leq C\}$ is as a closed subset of a product of two compact Polish spaces. This also implies that I and $\tilde{I}$ are lower-semicontinuous.

The fourth statement is a simple exercise in measure theory; it follows from the topological properties of the specific free energy stated in Theorem 3.4.1. If $g \sim \mu$, then it is clear that

$$\tilde{I}(g, \mu) = \int_D \mathcal{H}(\mu(x, \cdot)|\Phi) dx \geq \int_D \sigma(S(\mu(x, \cdot))) dx = \int_D \sigma(\nabla g(x)) dx.$$

For fixed g, this inequality can be turned into an equality, by constructing μ in terms of ∇g and the kernel from the fourth statement. This proves the fifth statement. The final statement is now obvious. $\square$

Theorem 3.11.5 states the LDP for the sequence of normalised measures $(\gamma_n)_{n \in \mathbb{N}}$. For the proof, however, it will be beneficial to consider also the sequence of non-normalised measures $(\tilde{\gamma}_n)_{n \in \mathbb{N}}$. Write $\tilde{\gamma}_n^*$ for the pushforward of $\tilde{\gamma}_n$ along $\mathfrak{P}_n$. Theorem 3.11.5 is equivalent to the conjunction of the following two statements:

1. The minimum of $\tilde{I}$ is $P_\Phi(D, b)$,
2. The sequence $(\tilde{\gamma}_n^*)_{n \in \mathbb{N}}$ satisfies an LDP with speed n^d and rate function $\tilde{I}$ in $(X^{\mathfrak{P}}, \mathcal{X}^{\mathfrak{P}})$.

The first statement was proven in Lemma 3.11.4. The second statement is somewhat easier to prove than the original LDP, because it appeals to non-normalised measures only.

Let us first describe a particular basis for the topological space

$$(X^{\mathfrak{P}}, \mathcal{X}^{\mathfrak{P}}) = (\mathrm{Lip}(\bar{D}), \mathcal{X}^\infty) \times (\mathcal{M}^D, \mathcal{X}^{\mathfrak{L}}).$$

As a basis $\mathcal{B}^\infty$ for $\mathcal{X}^\infty$, we take the sets of the form

$$B_\varepsilon^\infty(g) := \{h \in \mathrm{Lip}(\bar{D}) : \|h - g\|_\infty < \varepsilon\}$$

where $g \in \mathrm{Lip}(\bar{D})$ and $\varepsilon > 0$. Write

$$B_\varepsilon^{\mathfrak{L}}(\mu, (R_i)_i, (f_j)_j) := \{\nu : |\mu(R_i, f_j) - \nu(R_i, f_j)| < \mathrm{Leb}(R_i)\varepsilon \text{ for all } i, j\} \subset \mathcal{M}^D,$$

where $\varepsilon > 0$, μ is a measure in $\mathcal{M}^D$, $(R_i)_i$ is a finite collection of closed rectangular subsets of D, and $(f_j)_j$ is a finite collection of continuous cylinder functions $f_j : \Omega \to [0, 1]$. The collection $\mathcal{B}^{\mathfrak{L}}$ of such sets forms a basis of $\mathcal{X}^{\mathfrak{L}}$. As a basis $\mathcal{B}^{\mathfrak{P}}$ for $\mathcal{X}^{\mathfrak{P}}$, we choose the collection of open sets of the form $B_\varepsilon^{\mathfrak{P}}(\cdot, \cdot, \cdot, \cdot) := B_\varepsilon^\infty(\cdot) \times B_\varepsilon^{\mathfrak{L}}(\cdot, \cdot, \cdot)$.

To prove a large deviations principle, it must first be checked that the rate function is lower-semicontinuous. For this refer again to Lemma 3.11.4. The large deviations principle (with non-normalised measures) is now a corollary of the following three claims:

1. *Lower bound on probabilities.* For any $(g, \mu) \in A \in \mathcal{B}^{\mathfrak{P}}$, we have

$$\liminf_{n \to \infty} n^{-d} \log \tilde{\gamma}_n^*(A) \geq -\tilde{I}(g, \mu).$$

2. *Upper bound on probabilities.* For any $(g, \mu) \in X^{\mathfrak{P}}$, we have

$$\inf_{A \in \mathcal{B}^{\mathfrak{P}} \text{ with } (g,\mu) \in A} \limsup_{n \to \infty} n^{-d} \log \tilde{\gamma}_n^*(A) \leq -\tilde{I}(g, \mu).$$

3. *Exponential tightness.* For all $\alpha > -\infty$, there is a compact set $K_\alpha \subset X^{\mathfrak{P}}$ such that

$$\limsup_{n \to \infty} n^{-d} \log \tilde{\gamma}_n^*(X^{\mathfrak{P}} \smallsetminus K_\alpha) \leq \alpha$$

The next subsection contains an auxiliary result on approximations of Lipschitz functions which is useful for proving the lower bound. Each of the three subsequent sections addresses one of the three claims formulated above.

3.11.3 Simplicial approximations of Lipschitz function

This subsection is dedicated to providing some results on affine approximations of Lipschitz functions necessary to prove the lower bound on probabilities. For $x \in \mathbb{R}^d$, the point $\lfloor x \rfloor \in \mathbb{Z}^d$ is obtained by rounding down each coordinate.

Definition 3.11.6. Let S_d denote the group of permutations on $\{1, \ldots, d\}$. For $x \in \mathbb{R}$, we write $s(x) \in S_d$ for the permutation which rank-orders the coordinate indices of $x - \lfloor x \rfloor$. For $x \in \mathbb{Z}^d$ and $s \in S_d$, we define the *simplex* $C(x, s)$ to be the closure of the set

$$\{y \in \mathbb{R}^d : \lfloor y \rfloor = x, \, s(y) = s\}.$$

By a *simplex of scale* ε, we simply mean a scaled simplex of the form $\varepsilon C(x, s)$. A *simplex domain of scale* ε is a union of finitely many simplices of scale ε. If D is a domain, then write D_ε for the largest simplex domain of scale ε contained in D.

Definition 3.11.7. Let D denote a domain, and g a real-valued function on D. Consider $\varepsilon > 0$. Write $F_\varepsilon = F_\varepsilon(g)$ for the unique real-valued function on D_ε which equals g on $D_\varepsilon \cap \varepsilon\mathbb{Z}^d$, interpolated linearly on each simplex.

We will make use of the simplicial Rademacher theorem proven in [36] for which we recall a statement here.

Lemma 3.11.8 (Lemma 6.1 from [36])**.** *Consider a positive homogeneous function $\| \cdot \| : \mathbb{R}^d \to \mathbb{R}$ satisfying the triangle inequality. Let $D \subset \mathbb{R}^d$ be a domain and $g : D \to \mathbb{R}$ a $\| \cdot \|$-Lipschitz function. For any $\delta > 0$ and any $\varepsilon > 0$ sufficiently small (depending on δ), we have*

1. $\mathrm{Leb}(D \smallsetminus D_\varepsilon) \leq \delta$,

2. $\|F_\varepsilon - g|_{D_\varepsilon}\|_\infty \leq \delta\varepsilon$,

3. $\mathrm{Leb}(\{x \in D_\varepsilon : \|\nabla F_\varepsilon(x) - \nabla g(x)\|_2 \geq \delta\}) \leq \delta$.

Moreover, F_ε is $\| \cdot \|$-Lipschitz for any $\varepsilon > 0$.

The first property is obvious, and the proof of the second and third property is identical to the proof in [36].

3.11.4 The lower bound on probabilities

Fix $(g, \mu) \in A \in \mathcal{B}^{\mathfrak{P}}$ and $\beta > 0$; the goal of this subsection is to prove that

$$\liminf_{n\to\infty} n^{-d} \log \tilde{\gamma}_n^*(A) \geq -\tilde{I}(g, \mu) - \beta.$$

We suppose of course that $\tilde{I}(g, \mu)$ is finite. For the proof, we require the following result.

Lemma 3.11.9. *Consider some fixed $\varepsilon > 0$. Then there exists a sufficiently small constant $\alpha > 0$ such that the following statement holds true. Suppose that $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F})$ satisfies $u := S(\mu) \in \bar{U}_{q_\varepsilon} \subset U_q$, and that $v \in U_q$ is another slope with $\|u - v\|_2 \leq \alpha$. Then there is another measure $\nu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F})$ such that $S(\nu) = v$ and $\mathcal{H}(\nu|\Phi) \leq \mathcal{H}(\mu|\Phi) + \varepsilon$ and $\|\mu - \nu\|_{\mathrm{TV}} := \|\mu - \nu\|_\infty < \varepsilon$. In particular, if $f : \Omega \to [0, 1]$ is measurable, then $|\mu(f) - \nu(f)| < \varepsilon$.*

Proof. Note that σ is bounded uniformly on a neighbourhood A of $\bar{U}_{q_\varepsilon}$. The proof of the lemma is straightforward: one simply defines $\nu := (1 - t)\mu + t\mu'$ for t small and μ' some minimiser with $S(\mu') \in A$ in order to adjust the slope of the measure of interest. □

Proof of the lower bound on probabilities. It suffices to consider the case that $\tilde{I}(g, \mu)$ is finite. We claim that it is sufficient to consider the case that g is strictly $\|\cdot\|_q$-Lipschitz (if $E = \mathbb{R}$) or that $g|_D$ is locally strictly $\|\cdot\|_q$-Lipschitz (if $E = \mathbb{Z}$). If g were not $\|\cdot\|_q$-Lipschitz and $g \sim \mu$, then μ cannot be supported on q-Lipschitz functions, and consequently $\tilde{I}(g, \mu) = \infty$. Therefore g must be $\|\cdot\|_q$-Lipschitz. There is some pair $(h, \nu) \in X^{\mathfrak{P}}$ such that h is strictly $\|\cdot\|_q$-Lipschitz (if $E = \mathbb{R}$) or such that $h|_D$ is locally strictly $\|\cdot\|_q$-Lipschitz (if $E = \mathbb{Z}$), and such that $\tilde{I}(h, \nu) < \infty$—this follows from the definition of a good asymptotic boundary profile and from Lemma 3.11.4. Define $g_t := (1 - t)g + th$ and $\mu_t := (1 - t)\mu + t\nu$. Then $(g_t, \mu_t) \in A$ for t sufficiently small and $\limsup_{t\to 0} \tilde{I}(g_t, \mu_t) \leq \tilde{I}(g, \mu)$ as $\tilde{I}$ is convex. Moreover, for any $t > 0$, the function g_t has the desired properties. Thus, we may replace (g, μ) by (g_t, μ_t) for small t, by choosing β smaller if necessary. This proves the claim.

The proof follows the general strategy that was outlined after the statement of Lemma 3.10.24. Let us first consider the case that $E = \mathbb{R}$. We find an appropriate approximation of g using the simplicial Rademacher theorem, and then apply Lemma 3.11.9 and the limit equalities to obtain the desired lower bound on probabilities. For the approximations, it is necessary to take limits in three variables: first we take $n \to \infty$, then $\varepsilon_2 \to 0$, and finally $\varepsilon_1 \to 0$. There is also another variable ε; it is not necessary to take a limit in this variable, but it must be small for the arguments to work.

First fix $\varepsilon > 0$ so small that b and g are $\|\cdot\|_{q_{8\varepsilon}}$-Lipschitz, and such that all functions b_n are $q_{8\varepsilon}$-Lipschitz. We also suppose that $A = B_{8\varepsilon}^{\mathfrak{P}}(g, \mu, (R_i)_i, (f_j)_j)$, by choosing ε and A smaller if necessary, where $(R_i)_i$ is a finite family of rectangular subsets of D, and $(f_j)_j$ a finite family of continuous cylinder functions $f_j : \Omega \to [0, 1]$.

Consider some $\varepsilon_1 > 0$, and write D' for the points in D at distance more than ε_1 from the complement of D. Consider additionally some $\varepsilon_2 > 0$, and write D'' for D'_{ε_2}: the largest simplex domain of scale ε_2 contained in D'. See Figure 3.5 for a drawing of this construction. Write $F = F(g)$ for the unique $\|\cdot\|_{q_{8\varepsilon}}$-Lipschitz function on D'' which equals g on $\varepsilon_2\mathbb{Z}^d \cap D''$, and which is affine on each simplex of D''. For ε_2 sufficiently small, this function has a $\|\cdot\|_{q_{7\varepsilon}}$-Lipschitz extension $\bar{F}$ to $\bar{D}$ which

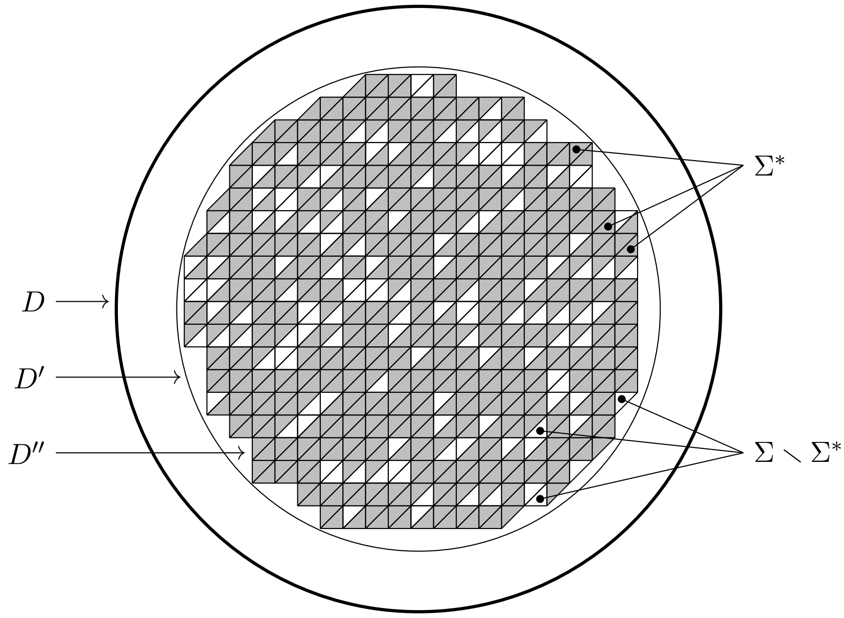


Figure 3.5: The sets $D'' \subset D' \subset D \subset \mathbb{R}^d$, and the sets $\Sigma^* \subset \Sigma$ of simplices of scale ε_2

equals b on ∂D. It is clear that any such extension $\bar{F}$ is contained in $B_\varepsilon^\infty(g)$, that is, $\|\bar{F} - g\|_\infty < \varepsilon$, for ε_1 and ε_2 sufficiently small.

Write Σ for the set of simplices of scale ε_2 in D''—this is a finite set. The slope ∇F of F is constant on any $\Delta \in \Sigma$; write $S(\Delta) \in \bar{U}_{q_{8\varepsilon}}$ for this slope. Write Σ^* for the set of simplices $\Delta \in \Sigma$ for which $\|S(\Delta) - S(\mu(\Delta, \cdot))\|_2 \leq \varepsilon_1$; Lemma 3.11.8 asserts that $|\Sigma^*|/|\Sigma| \geq 1 - \varepsilon_1$ for ε_2 sufficiently small. See again Figure 3.5 for an example of the sets Σ and Σ^*.

Choose C minimal subject to $\|\phi^u - u|_{\mathbb{Z}^d}\|_\infty + 1 \leq C$ for all $u \in U_\Phi$. Let M denote a constant which makes Theorem 3.6.9 work for the local Lipschitz constraint $q_{6\varepsilon}$, and for the constants ε and C—this constant M depends on ε only. We shall also suppose that $M \geq R$, by choosing M larger if necessary. For $\Delta \in \Sigma$ and $n \in \mathbb{N}$, define $\Delta_n := \Lambda^{-M}(n\Delta)$. Write also $D_n'' := \cup_{\Delta \in \Sigma} \Delta_n$ and $D_n^* := \cup_{\Delta \in \Sigma^*} \Delta_n$. It follows from the definition of an approximation that $D_n'' \subset D_n$ for n sufficiently large. By Theorem 3.6.9 there exists, for any $n \in \mathbb{N}$, a $q_{6\varepsilon}$-Lipschitz function $F_n : D_n'' \to E$ such that:

1. $|F_n(x) - nF(x/n)| \leq C$ for all $x \in D_n''$,
2. $\nabla F_n|_{\Delta_n} = \nabla \phi^{S(\Delta)}|_{\Delta_n}$ for all $\Delta \in \Sigma$.

It is straightforward to see that for n sufficiently large, the function F_n extends to a $q_{5\varepsilon}$-Lipschitz height function $\bar{F}_n$ which equals b_n on the complement of D_n.

We now use the existence of the function $\bar{F}_n$ to demonstrate that there exists a set $A_n \in \mathcal{F}$ such that $\mathfrak{P}_n(A_n) \subset A$, and for which we show that $\tilde{\gamma}_n(A_n)$ is sufficiently large as $n \to \infty$. Define A_n to be the set of height functions ϕ which are q-Lipschitz, and which satisfy the following criteria:

1. If $x \in \mathbb{Z}^d \smallsetminus D_n$, then $\phi(x) = \bar{F}_n(x) = b_n(x)$,
2. If $x \in D_n \smallsetminus D_n^*$, then $|\phi(x) - \bar{F}_n(x)| \leq \varepsilon$,
3. If $x = 0_{\Delta_n}$ for some $\Delta \in \Sigma^*$, then $|\phi(x) - \bar{F}_n(x)| \leq \varepsilon$,

4. For each $\Delta \in \Sigma^*$, we have $\phi \in C^{S(\Delta)}_{\Delta_n,\varepsilon}$,

5. For each $\Delta \in \Sigma^*$, we have $\phi \in B^{\Delta}_{\Delta_n}$, that is, $L_{\Delta_n}(\phi) \in B^{\Delta}$, where

$$B^{\Delta} := \{\nu \in \mathcal{P}(\Omega, \mathcal{F}^{\nabla}) : |\operatorname{Leb}(\Delta)\nu(f_j) - \mu(\Delta, f_j)| < \operatorname{Leb}(\Delta)\varepsilon \text{ for all } j\} \in \mathcal{B}.$$

It suffices to demonstrate that for ε, ε_1, and ε_2 sufficiently small, and for n sufficiently large, we have $\mathfrak{P}_n(A_n) \subset A$ and

$$\liminf_{n\to\infty} n^{-d} \log \tilde{\gamma}_n(A_n) \geq -I(g,\mu) - \beta.$$

Claim first that, in the limit, $\mathfrak{P}_n(A_n) \subset A$. This is equivalent to asking that $\mathfrak{G}_n(A_n) \subset B^{\infty}_{8\varepsilon}(g)$ and $\mathfrak{L}_n(A_n) \subset B^{\mathfrak{L}}_{8\varepsilon}(\mu, (R_i)_i, (f_j)_j)$. The former of the two holds true because $\|\bar{F} - g\|_\infty < \varepsilon$ and because $\|\bar{F} - \mathfrak{G}_n(\phi)\|_\infty$ is small in the described limit, uniformly over the choice of $\phi \in A_n$. The proof that $\mathfrak{L}_n(A_n) \subset B^{\mathfrak{L}}_{8\varepsilon}(\mu, (R_i)_i, (f_j)_j)$ in the limit relies again on Proposition 3.10.25; observe in particular that in the limit most of the volume of each fixed rectangle R_i is covered by simplices in Σ^* which are entirely contained in R_i.

In the sequel, we shall pretend that $A_n \in \mathcal{E}^{D_n}$ by restricting each height function in A_n to D_n. If $\phi \in E^{D_n}$, then we write ψ for the height function which restricts to ϕ on D_n and to b_n on the complement of D_n. We aim to find an asymptotic lower bound on

$$n^{-d} \log \tilde{\gamma}_n(A_n) = n^{-d} \log \int_{A_n} e^{-H_{D_n}(\psi)} d\lambda^{D_n}(\phi).$$

If $\phi \in A$, then ψ is q_ε-Lipschitz whenever restricted to $\mathbb{Z}^d \smallsetminus \cup_{\Delta\in\Sigma^*} \Delta_n^{-R}$, because $\bar{F}_n$ is $q_{5\varepsilon}$-Lipschitz and because ψ and $\bar{F}_n$ differ by at most 2ε at each vertex in this set. Therefore the upper attachment lemma (Lemma 3.7.4) implies

$$H_{D_n}(\psi) \leq H^0_{D_n \smallsetminus D_n^*}(\psi) + e^+_\varepsilon(D_n) + \sum_{\Delta\in\Sigma^*} H^0_{\Delta_n}(\psi) + e^+_\varepsilon(\Delta_n)$$

for any $\phi \in A$. For fixed ε, ε_1, and ε_2, the terms of the form $e^+_\varepsilon(\cdot)$ in this expression are of order $o(n^d)$ as $n \to \infty$, and therefore we may omit them in calculating the limit inferior. Moreover, since ψ is q_ε-Lipschitz on $D_n \smallsetminus D_n^*$, the term $H^0_{D_n\smallsetminus D_n^*}(\psi)$ has an upper bound $C'|D_n \smallsetminus D_n^*|$, where C' depends on ε only. In particular,

$$\begin{aligned}\liminf_{n\to\infty} & n^{-d} \log \tilde{\gamma}_n(A_n) \\ &\geq \liminf_{n\to\infty} n^{-d} \left[-C'|D_n \smallsetminus D_n^*| + \log \int_{A_n} e^{-\sum_{\Delta\in\Sigma^*} H^0_{\Delta_n}(\psi)} d\lambda^{D_n}(\phi) \right].\end{aligned}$$

It follows from the definition of A_n, that the integral decomposes as follows:

$$\begin{aligned}\int_{A_n} & e^{-\sum_{\Delta\in\Sigma^*} H^0_{\Delta_n}(\psi)} d\lambda^{D_n}(\phi) \\ &= \left[\prod_{x\in D_n\smallsetminus D_n^*} \int_{\bar{F}_n(x)-\varepsilon}^{\bar{F}_n(x)+\varepsilon} d\lambda \right] \left[\prod_{\Delta\in\Sigma^*} \int_{\bar{F}_n(0_{\Delta_n})-\varepsilon}^{\bar{F}_n(0_{\Delta_n})+\varepsilon} d\lambda \right] \cdot \\ &\qquad \cdot \left[\prod_{\Delta\in\Sigma^*} \int_{C^{S(\Delta)}_{\Delta_n,\varepsilon}\cap B^{\Delta}_{\Delta_n}} e^{-H^0_{\Delta_n}} d\lambda^{\Delta_n - 1} \right],\end{aligned}$$

and therefore the logarithm of this integral equals

$$(|D_n \smallsetminus D_n^*| + |\Sigma^*|) \log 2\varepsilon - \sum_{\Delta \in \Sigma^*} \mathrm{PB}_{\Delta_n, S(\Delta), \varepsilon}(B^\Delta).$$

But $|\Sigma^*|$ does not depend on n, and by choosing C' larger, we obtain

$$\liminf_{n\to\infty} n^{-d} \log \tilde{\gamma}_n(A_n) \geq \liminf_{n\to\infty} n^{-d} \left[-C'|D_n \smallsetminus D_n^*| - \sum_{\Delta\in\Sigma^*} \mathrm{PB}_{\Delta_n, S(\Delta),\varepsilon}(B^\Delta) \right].$$

It is easy to see that $n^{-d}|D_n \smallsetminus D_n^*| \to \operatorname{Leb}(D \smallsetminus \cup\Sigma^*)$ as $n \to \infty$. Fix $\Delta \in \Sigma^*$. By definition of Σ^*, we have $\|S(\Delta) - S(\mu(\Delta, \cdot))\|_2 \leq \varepsilon_1$. Then Lemma 3.11.9 tells us that for ε_1 sufficiently small, the set B^Δ contains another shift-invariant measure ν of slope $S(\Delta)$ such that $\mathcal{H}(\nu|\Phi) \leq \mathcal{H}(\operatorname{Leb}(\Delta)^{-1}\mu(\Delta,\cdot)|\Phi) + \varepsilon$. In particular, this means that

$$\begin{aligned}\limsup_{n\to\infty} n^{-d}\, \mathrm{PB}_{\Delta_n, S(\Delta),\varepsilon}(B^\Delta) &\leq \operatorname{Leb}(\Delta)\left(\mathcal{H}(\operatorname{Leb}(\Delta)^{-1}\mu(\Delta,\cdot)|\Phi) + \varepsilon\right) \\ &= \mathcal{H}(\mu(\Delta,\cdot)|\Phi) + \operatorname{Leb}(\Delta)\varepsilon.\end{aligned}$$

Conclude that

$$\liminf_{n\to\infty} n^{-d} \log \tilde{\gamma}_n(A_n) \geq -C' \operatorname{Leb}(D \smallsetminus \cup\Sigma^*) - \mathcal{H}(\mu(\cup\Sigma^*, \cdot)|\Phi) - \varepsilon \operatorname{Leb}(\cup\Sigma^*).$$

As $\varepsilon_2 \to 0$ and then $\varepsilon_1 \to 0$, we have

$$\operatorname{Leb}(\cup\Sigma^*) \to \operatorname{Leb}(D), \qquad \operatorname{Leb}(D \smallsetminus \cup\Sigma^*) \to 0, \qquad \mathcal{H}(\mu(\cup\Sigma^*,\cdot)|\Phi) \to \mathcal{H}(\mu|\Phi).$$

The desired lower bound is thus obtained by setting ε so small that $\varepsilon \operatorname{Leb}(D) < \beta$.

Let us finally describe what changes for $E = \mathbb{Z}$. The first part of the proof is the same, except that the functions F_n and $\bar{F}_n$ are q-Lipschitz, and not $q_{6\varepsilon}$-Lipschitz or $q_{5\varepsilon}$-Lipschitz. The q-Lipschitz extension $\bar{F}_n$ exists for n sufficiently large, because $g|_D$ is locally strictly $\|\cdot\|_q$-Lipschitz. The only thing that changes in the remainder of the proof is that λ is now the counting measure rather than the Lebesgue measure. This makes the remainder of the proof easier, exactly as in the proof of Lemma 3.10.24. □

3.11.5 The upper bound on probabilities

Proof of the upper bound on probabilities. Let us first consider the case $\tilde{I}(g,\mu) < \infty$, in which case $\tilde{I}(g,\mu) = \mathcal{H}(\mu|\Phi)$. Let Σ denote a finite set of closed disjoint rectangles, contained in D. Define $R_n := \Lambda(nR)$ for $R \in \Sigma$ and $\Sigma_n := \cup_{R\in\Sigma} R_n$, and note that $\Sigma_n \subset D_n$ for n sufficiently large. Now choose for each $R \in \Sigma$ an open set $B^R \in \mathcal{B}$ with $\mu(R,\cdot)/\operatorname{Leb}(R) \in B^R$, and define

$$A_n := \cap_{R\in\Sigma} B^R_{R_n}.$$

It is straightforward to show that (g,μ) has a fixed neighbourhood which is contained in all sets $\mathfrak{P}_n(A_n)$ for n sufficiently large. Fix $\beta > 0$. It suffices to find an appropriate choice for the set of rectangles Σ and the collection of balls $(B^R)_{R\in\Sigma}$, such that

$$\limsup_{n\to\infty} n^{-d} \log \tilde{\gamma}_n(A_n) \leq -\mathcal{H}(\mu|\Phi) + \beta.$$

Remove all height functions from A_n which do not equal b_n on $\mathbb{Z}^d \smallsetminus D_n$ or which are not Kd_1-Lipschitz; this obviously does not change the value of $\tilde{\gamma}_n(A_n)$. As in the proof of the lower bound, we shall sometimes pretend that $A_n \in \mathcal{E}^{D_n}$ by restricting each height function in A_n to D_n. If $\phi \in E^{D_n}$, then we write ψ for the height function which restricts to ϕ on D_n and to b_n on the complement of D_n. We are thus interested in the asymptotic behaviour of

$$\tilde{\gamma}_n(A_n) = \int_{A_n} e^{-H_{D_n}(\psi)} d\lambda^{D_n}(\phi).$$

The lower attachment lemma (Lemma 3.7.1) asserts that

$$\begin{aligned} H_{D_n} &\geq H^0_{D_n \smallsetminus \Sigma_n} - e^-(D_n) + \sum_{R\in\Sigma} H^0_{R_n} - e^-(R_n) \\ &\geq -\|\Xi\| \cdot |D_n \smallsetminus \Sigma_n| - e^-(D_n) + \sum_{R\in\Sigma} H^0_{R_n} - e^-(R_n). \end{aligned}$$

The terms of the form $e^-(\cdot)$ are of order $o(n^d)$ as $n \to \infty$. Moreover, $n^{-d}|D_n \smallsetminus \Sigma_n| \to \operatorname{Leb}(D \smallsetminus \cup\Sigma)$ as $n \to \infty$, and therefore

$$\begin{aligned} &\limsup_{n\to\infty} n^{-d} \log \tilde{\gamma}_n(A_n) \\ &\qquad \leq \|\Xi\| \operatorname{Leb}(D \smallsetminus \cup\Sigma) + \limsup_{n\to\infty} n^{-d} \log \int_{A_n} e^{-\sum_{R\in\Sigma} H^0_{R_n}(\psi)} d\lambda^{D_n}(\phi). \end{aligned}$$

Write $D_n^0 := D_n \smallsetminus \cup_{R\in\Sigma}(R_n \smallsetminus \{0_{R_n}\})$, so that $\lambda^{D_n} = \lambda^{D_n^0} \times \prod_{R\in\Sigma} \lambda^{R_n - 1}$. Then

$$\int_{A_n} e^{-\sum_{R\in\Sigma} H^0_{R_n}(\psi)} d\lambda^{D_n}(\phi) \leq \left[\int_{W_n} d\lambda^{D_n^0}\right] \left[\prod_{R\in\Sigma} \int_{B^R_{R_n}} e^{-H^0_{R_n}} d\lambda^{R_n - 1}\right],$$

where W_n is the set of Kd_1-Lipschitz functions $\phi : D_n^0 \to E$ such that $\phi b_n|_{\mathbb{Z}^d \smallsetminus D_n}$ is also Kd_1-Lipschitz. Remark that

$$\log \int_{W_n} d\lambda^{D_n^0} \leq |D_n^0| \log(2K+1)$$

and that

$$\log \int_{B^R_{R_n}} e^{-H^0_{R_n}} d\lambda^{R_n - 1} = -\operatorname{FB}_{R_n}(B^R).$$

If we write $m := \|\Xi\| + \log(2K+1)$, then we have now shown that

$$\limsup_{n\to\infty} n^{-d} \log \tilde{\gamma}_n(A_n) \leq m \operatorname{Leb}(D \smallsetminus \cup\Sigma) - \sum_{R\in\Sigma} \operatorname{Leb}(R) \operatorname{FB}(B^R).$$

It now suffices to show that the expression on the right is at most $-\mathcal{H}(\mu|\Phi) + \beta$ for an appropriate choice of the set of rectangles in Σ and for the collection $(B^R)_{R\in\Sigma} \subset \mathcal{B}$. By choosing the rectangles in Σ such that they exhaust most of the space, we can ensure that $\operatorname{Leb}(D \smallsetminus \cup\Sigma) \leq \beta/2m$, and by also taking each ball B^R sufficiently small, we can ensure that $\sum_{R\in\Sigma} \operatorname{Leb}(R) \operatorname{FB}(B^R)$ is at least $\mathcal{H}(\mu|\Phi) - \beta/2$. This proves the upper bound on probabilities.

Consider now the case that $\tilde{I}(g, \mu) = \infty$. We distinguish several reasons which may cause $\tilde{I}(g, \mu)$ to be infinite. If μ is shift-invariant but $\mathcal{H}(\mu|\Phi) = \infty$, then the proof is the same as before. If μ is not shift-invariant, then there is a closed rectangle $R \subset D$ such that $\mu(R, \cdot)$ is not shift-invariant, and by including R in Σ and using the free boundary limits for non-shift-invariant measures, we obtain the same result. In fact, in that case it is readily seen that A_n is empty for n sufficiently large. (See also the proof of Lemma 3.10.7).

The remaining cases are: either $g|_{\partial D}$ does not equal b, or it is not true that $\nabla g(x) = S(\mu(x, \cdot))$ as a distribution on D. Consider first the case that $g|_{\partial D}$ does not equal b. Choose $\varepsilon := \|g|_{\partial D} - b\|_\infty/2$. In that case, it is readily seen that $\tilde{\gamma}_n(\mathfrak{G}_n^{-1}(B_\varepsilon^\infty(g))) = 0$ for n sufficiently large. Finally consider the case that it is not true that $\nabla g(x) = S(\mu(x, \cdot))$ as a distribution on D. In that case, there is a closed rectangle $R \subset D$ such that the average of ∇g over R does not equal $S(\mu(R, \cdot))$. But if $\mathfrak{G}_n(\phi)$ is close to g, then $\mathfrak{L}_n(\phi)(R, \cdot)$ must have its approximate slope close to the average of ∇g over R. Note that we use the words *approximate slope* here rather than the word *slope*, because $\mathfrak{L}_n(\phi)(R, \cdot)$ is not shift-invariant, but it is almost shift-invariant in the sense that $\mathfrak{L}_n(\phi)(R, f - \theta f)$ goes to zero uniformly over ϕ as $n \to \infty$ for f a bounded continuous cylinder function and $\theta \in \Theta(\mathcal{L})$; see the proof of Lemma 3.10.7. In particular, by including R in Σ in the previous discussion and choosing B^R sufficiently small, it can again be seen that A_n is empty for n sufficiently large, which leads to the desired bound. $\square$

3.11.6 Exponential tightness

Proof of exponential tightness. The proof is easy. Fix a positive constant $\varepsilon > 0$, and let K denote the smallest constant such that $Kd_1 \geq q$. Define

$$K_\varepsilon^\infty := \{g \in \operatorname{Lip}(\bar{D}) : \|g|_{\partial D} - b\|_\infty \leq \varepsilon\}, \; K^{\mathfrak{L}} := \{\mu \in \mathcal{M}^D : \mu(D, \cdot) \text{ is } K\text{-Lipschitz}\}.$$

It is clear that K_ε^∞ is compact in the topological space $(\operatorname{Lip}(\bar{D}), \mathcal{X}^\infty)$, and that the set $K^{\mathfrak{L}}$ is compact in $(\mathcal{M}^D, \mathcal{X}^{\mathfrak{L}})$. This means that $K_\varepsilon^\infty \times K^{\mathfrak{L}}$ is compact in $(X^{\mathfrak{P}}, \mathcal{X}^{\mathfrak{P}})$. As in the proof of the upper bound of probabilities, we observe that $\tilde{\gamma}_n^*$ is supported on $K_\varepsilon^\infty \times K^{\mathfrak{L}}$ for n sufficiently large. This completes the proof; the compact set that we have found is independent of the choice of α that appeared in the original formulation of exponential tightness. $\square$

3.12 Proof of strict convexity

3.12.1 The product setting

For the proof of strict convexity of σ, it is useful to work in the product setting $\Omega \times \Omega$, because one is then able to study the difference $\phi_1 - \phi_2$ of a pair of height functions (ϕ_1, ϕ_2) and apply the theory of moats from Section 3.5. Almost all constructions and results in the previous sections generalise to the product setting. An alternative way of viewing the product setting is by considering a height function to take values in the two-dimensional space E^2 rather than E. This section gives an overview of the definitions and results for the product setting as required for the proof of strict convexity of σ.

Write $\mathcal{P}^2(X, \mathcal{X})$ for the set of probability measures on $(X, \mathcal{X})^2$ whenever $(X, \mathcal{X})$ is a measurable space. If $\mu \in \mathcal{P}^2(X, \mathcal{X})$, then write μ_1 and μ_2 for the marginals of μ on the first and second space respectively.

Definition 3.12.1. The *topology of weak local convergence* is the coarsest topology on $\mathcal{P}^2(\Omega, \mathcal{F}^\nabla)$ that makes the evaluation map $\mu \mapsto \mu(f)$ continuous for any bounded continuous cylinder function f on Ω^2, that is, a bounded function $f : \Omega^2 \to \mathbb{R}$ which is $\mathcal{F}^\nabla_\Lambda \times \mathcal{F}^\nabla_\Lambda$-measurable for some $\Lambda \subset\subset \mathbb{Z}^d$, and continuous with respect to the topology of uniform convergence on Ω^2—the set of functions from $\mathbb{Z}^d$ to E^2.

Definition 3.12.2. Write $\mathcal{P}^2_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ for the set of $\mathcal{L}$-invariant probability measures in $\mathcal{P}^2(\Omega, \mathcal{F}^\nabla)$; a measure $\mu \in \mathcal{P}^2(\Omega, \mathcal{F}^\nabla)$ is called $\mathcal{L}$-*invariant* if $\mu(A \times B) = \mu(\theta A \times \theta B)$ for any $A, B \in \mathcal{F}^\nabla$ and $\theta \in \Theta$. This is equivalent to asking that (ϕ_1, ϕ_2) and $(\theta\phi_1, \theta\phi_2)$ have the same distribution under μ.

Definition 3.12.3. By the *slope* of $\mu \in \mathcal{P}^2_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ we simply mean the pair of slopes of the two marginals of μ; $S^2(\mu) := (S(\mu_1), S(\mu_2))$. The slope functional S^2 is clearly strongly affine, as in the non-product setting.

Definition 3.12.4. For $\mu \in \mathcal{P}^2(\Omega, \mathcal{F}^\nabla)$ and $\Lambda \subset\subset \mathbb{Z}^d$, define the *free energy* of μ in Λ by

$$\mathcal{H}^2_\Lambda(\mu|\Phi) := \mathcal{H}_{\mathcal{F}^\nabla_\Lambda \times \mathcal{F}^\nabla_\Lambda}(\mu|\lambda^{\Lambda-1} \times \lambda^{\Lambda-1}) + \mu(H^{0,\Phi}_\Lambda(\phi_1) + H^{0,\Phi}_\Lambda(\phi_2)).$$

Note that we immediately have

$$\mathcal{H}^2_\Lambda(\mu|\Phi) \geq \mathcal{H}_\Lambda(\mu_1|\Phi) + \mathcal{H}_\Lambda(\mu_2|\Phi), \tag{3.12.5}$$

with equality if and only if the restriction of μ to $\mathcal{F}^\nabla_\Lambda \times \mathcal{F}^\nabla_\Lambda$ decomposes as the product of μ_1 and μ_2, or if either side equals ∞. If μ is $\mathcal{L}$-invariant, then define the *specific free energy* of μ by

$$\mathcal{H}^2(\mu|\Phi) = \lim_{n\to\infty} n^{-d} \mathcal{H}^2_{\Pi_n}(\mu|\Phi).$$

It follows immediately from (3.12.5) that $\mathcal{H}^2(\mu|\Phi) \geq \mathcal{H}(\mu_1|\Phi) + \mathcal{H}(\mu_2|\Phi)$. In particular, this implies that $\mathcal{H}^2(\mu|\Phi) \geq \sigma(S(\mu_1)) + \sigma(S(\mu_2))$. For convenience, we shall write $\sigma^2(u, v) := \sigma(u) + \sigma(v)$. Note that

$$\sigma^2(u, v) := \inf_{\mu \in \mathcal{P}^2_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla) \text{ with } S^2(\mu) = (u,v)} \mathcal{H}^2(\mu|\Phi).$$

With these definitions, the following results generalise naturally to the product setting:

1. Theorem 3.4.1 for existence of the specific free energy,
2. Theorem 3.4.2 for finite energy, where the result applies if
$$\mathcal{H}^2(\mu|\Phi) = \sigma^2(S^2(\mu)) < \infty,$$
3. Theorem 3.9.1, Proposition 3.9.3, and Theorem 3.9.4 for ergodic decompositions,
4. Theorem 3.10.5 for limit equalities and Theorem 3.11.5 for the large deviations principle.

Rather than repeating each result here, we state clearly the generalised result that is used whenever referring to it.

3.12.2 Moats in the empirical limit

In this section, we suppose that σ is not strictly convex, and construct the pathological measure which derives from this assumption. Let K denote the smallest real number such that $Kd_1 \geq q$, and write ρ for the uniform probability measure on the set $E \cap [0, 4K)$, with random variable U. For fixed $(\phi_1, \phi_2, U) \in \Omega \times \Omega \times E$, we shall write $\xi = \xi(\phi_1, \phi_2, U)$ for the function

$$\xi := \left\lfloor \frac{1}{4K}(\phi_1 - \phi_1(0) - \phi_2 + \phi_2(0) - U) \right\rfloor : \mathbb{Z}^d \to \mathbb{Z}.$$

We will refer to ξ as the *difference function* associated to the triplet (ϕ_1, ϕ_2, U). Remark that the law of $\nabla\xi$ is $\mathcal{L}$-invariant in $\mu \times \rho$ for any $\mu \in \mathcal{P}^2_{\mathcal{L}}(\Omega, \mathcal{F})$; the random variable U makes the rounding operation shift-invariant as in the proof of Lemma 3.10.32.

Theorem 3.12.6. *Let Φ denote a potential which is monotone and in $\mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$. Assume that σ is affine on the line segment $[u_1, u_2]$ connecting two distinct slopes $u_1, u_2 \in U_\Phi$, and set $u = (u_1 + u_2)/2$. Select two vertices $x \in \mathcal{L}$ and $y \in \mathbb{Z}^d$ subject only to $(u_1 - u_2)(x) \neq 0$. Then there exists a product measure $\mu \in \mathcal{P}^2_{\mathcal{L}}(\Omega, \mathcal{F}^\nabla)$ such that $S^2(\mu) = (u, u)$ and $\mathcal{H}^2(\mu|\Phi) = \sigma^2(u, u) = 2\sigma(u)$, and such that with positive $\mu \times \rho$-probability, the following two events occur simultaneously:*

1. *The function ξ is not constant on the set $y + \mathbb{Z}x$,*
2. *The set $\{\xi = 0\} \subset \mathbb{Z}^d$ has at least three distinct infinite connected components.*

In the next section, we discuss rigorously how to derive a contradiction from this theorem (under the additional condition whenever $E = \mathbb{Z}$), using Theorem 3.4.2 and the argument for uniqueness of the infinite cluster of Burton and Keane [4]. The purpose of the remainder of this section is to prove Theorem 3.12.6.

Let us assume the setting of Theorem 3.12.6: Φ is a monotone potential in $\mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$, u_1 and u_2 are distinct slopes in U_Φ such that σ is affine on $[u_1, u_2]$, and $u := (u_1 + u_2)/2$. In the proof of the theorem, we shall suppose that $y = 0$, without loss of generality. Fix $0 < \varepsilon < K$ so small that $u_1, u_2 \in U_{q_\varepsilon}$. We shall use the large deviations principle with the good asymptotic profile (D, b) where $D := (0, 1)^d \subset \mathbb{R}^d$ and $b := u|_{\partial D}$, and with the good approximation $(D_n, b_n)_{n \in \mathbb{N}}$ of (D, b) defined by $D_n := \Pi_n$ and $b_n := \phi^u$ for all $n \in \mathbb{N}$. As per usual, we write $\gamma_n := \gamma_{D_n}(\cdot, b_n)$, and we shall also write $\gamma_n^2 := \gamma_n \times \gamma_n$.

Set $t = 1/2$, and recall the definitions of the slope v and the functions p and p_α from the proof of Lemma 3.10.24 (Page 114). Fix ε_1, ε_2, and ε_3 strictly positive and consider $n \in \mathbb{N}$. When taking limits we shall take first $n \to \infty$, then $\varepsilon_3 \to 0$, then $\varepsilon_2 \to 0$, and finally $\varepsilon_1 \to 0$; it is again convenient to work on different scales. Define $D' := (\varepsilon_1, 1 - \varepsilon_1)^d \subset D$ and $D'_n := nD' \cap \mathbb{Z}^d$. Write H_k for the affine hyperplane $\{2v = k\varepsilon_2\} \subset \mathbb{R}^d$. Note that the sets $(H_k)_{k \in \mathbb{Z}}$ correspond to the hyperplanes where the gradient of p_{ε_2} changes. For k even, p_{ε_2} equals u on H_k. For k odd, p_{ε_2} equals $u + \varepsilon_2/4$ on H_k. Finally, write

$$\begin{aligned}
H_{n,k} &:= \{x \in \mathbb{Z}^d : d_2(x, nH_k) \leq n\varepsilon_3\}, \\
D_n^0 &:= (\cup_{k \in 2\mathbb{Z}} H_{n,k}) \cap D_n, \\
D_n^+ &:= (\cup_{k \in 2\mathbb{Z}+1} H_{n,k}) \cap D'_n.
\end{aligned}$$

See Figure 3.6 for an overview of this construction.

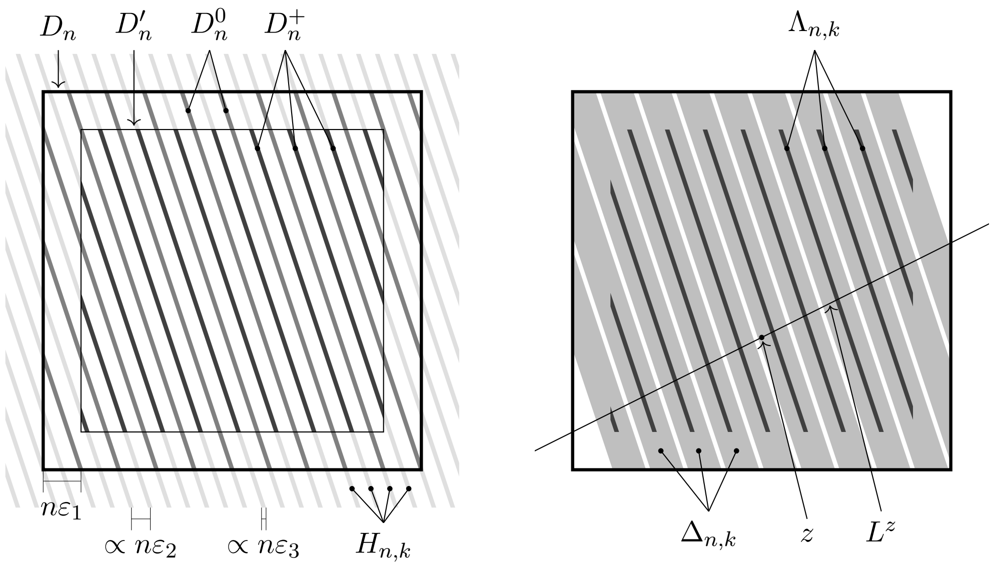


Figure 3.6: Several constructions in Subsection 3.12.2

Proposition 3.12.7. *Assume the setting of Theorem 3.12.6. If $E = \mathbb{Z}$, then there is a $\delta > 0$ such that*

$$n^{-d} \log \gamma_n(\phi_{D_n^0} = \phi^u_{D_n^0} \text{ and } \phi_{D_n^+} \geq \phi^u_{D_n^+} + n\delta\varepsilon_2) = o(1)$$

in the limit of n, ε_3, ε_2, and ε_1. If $E = \mathbb{R}$ and $\varepsilon > 0$, then there is a $\delta > 0$ such that

$$n^{-d} \log \gamma_n(|\phi_{D_n^0} - \phi^u_{D_n^0}| \leq \varepsilon \text{ and } \phi_{D_n^+} \geq \phi^u_{D_n^+} + n\delta\varepsilon_2) = o(1)$$

in the limit of n, ε_3, ε_2, and ε_1.

Proof. In fact, we shall demonstrate that any $\delta < 1/4$ works. Write f for the smallest $\|\cdot\|_{q_\varepsilon}$-Lipschitz function which satisfies $f \geq u$ and which equals p_{ε_2} on D'. This function is well-defined and equals u on $\mathbb{R}^d \smallsetminus D$ for ε_2 sufficiently small (depending only on ε_1).

The pressure $P_\Phi(D, b)$ is equal to $\sigma(u)$ because σ is convex, $b = u|_{\partial D}$, and $\mathrm{Vol}(D) = 1$. Moreover, $\int_D \sigma(\nabla f(x))dx$ tends to $\sigma(u)$ in the limit of ε_1 and ε_2, because σ is affine on the line segment connecting u_1 and u_2, and because the gradient of f equals u_1 on roughly half of D, and u_2 on roughly the other half of D with respect to Lebesgue measure. Note that $\sigma(\nabla f)$ is bounded uniformly as f is $\|\cdot\|_{q_\varepsilon}$-Lipschitz. This means that for any $\varepsilon' > 0$, which is allowed to depend arbitrarily on ε_1 and ε_2, we have

$$n^{-d} \log \gamma_n(\mathfrak{G}_n^{-1}(B^\infty_{\varepsilon'}(f))) = o(1)$$

in the limit of n, ε_2, and ε_1.

Note that for ε_3 and ε' sufficiently small depending on ε_1 and ε_2, all height functions $\phi \in \mathfrak{G}_n^{-1}(B^\infty_{\varepsilon'}(f))$ satisfy $\phi_{D_n^+} \geq \phi^u_{D_n^+} + n\delta\varepsilon_2$ (by virtue of the choice of f). Moreover, $\phi_{D_n^0}$ and $\phi^u_{D_n^0}$ must be close for such ϕ. By repeating arguments of the proof of the lower bound on probabilities in the large deviations principle, it is straightforward to see that conditioning further on the exact values of $\phi_{D_n^0}$ (up to ε in the continuous case) does not decrease the value of the limit of the normalised probabilities. In particular, this implies the proposition. ∎

By interchanging the role of u_1 and u_2, one obtains the same result as in Proposition 3.12.7, now with the inequality sign $\geq$ replaced by $\leq$, and with $n\delta\varepsilon_2$ replaced by $-n\delta\varepsilon_2$. By appealing to both the original proposition and the version with replacements, one deduces immediately the following proposition.

Proposition 3.12.8. *Assume the setting of Theorem 3.12.6. If $E = \mathbb{Z}$, then there exists a $\delta > 0$ such that*

$$n^{-d} \log \gamma_n^2((\phi_1 - \phi_2)_{D_n^0} = 0 \text{ and } (\phi_1 - \phi_2)_{D_n^+} \geq n\delta\varepsilon_2) = o(1)$$

in the limit of n, ε_3, ε_2, and ε_1. If $E = \mathbb{R}$ and $\varepsilon > 0$, then there is a $\delta > 0$ such that

$$n^{-d} \log \gamma_n^2(|(\phi_1 - \phi_2)_{D_n^0}| \leq 2\varepsilon \text{ and } (\phi_1 - \phi_2)_{D_n^+} \geq n\delta\varepsilon_2) = o(1)$$

in the limit of n, ε_3, ε_2, and ε_1.

Recall Section 3.5 on moats; we are now ready to apply the theory developed there. If k is odd with $H_{n,k} \cap D_n'$ nonempty, then write $\Lambda_{n,k} := H_{n,k} \cap D_n'$. Note that $D_n^+ = \cup_k \Lambda_{n,k}$. Write also $\Delta_{n,k}$ for the connected component of $D_n \smallsetminus D_n^0$ containing $\Lambda_{n,k}$; see Figure 3.6 for an example of the sets $\Lambda_{n,k}$ and $\Delta_{n,k}$. Write $E_n^a(m)$ for the event that each connected component $\Delta_{n,k}$ contains a sequence of $\lceil m \rceil$ nested $4K, 4K + a$-moats of $(\phi_1 - \phi_2, \Lambda_{n,k})$.

Lemma 3.12.9. *Assume the setting of Theorem 3.12.6. For any $a \geq 4K$, there is a $\delta > 0$ such that*

$$n^{-d} \log \gamma_n^2(E_n^a(n\delta\varepsilon_2)) = o(1)$$

in the limit of n, ε_3, ε_2, and ε_1.

Proof. This follows immediately from the previous proposition and from Proposition 3.5.11. Note that the prefactor which appears on the left in (3.5.5) is of order $n^{O(1/\varepsilon_2) \cdot O(n\delta\varepsilon_2)}$, because distances are bounded by n, there are at most $O(1/\varepsilon_2)$ sets $\Lambda_{n,k}$, and because we enforce $n\delta\varepsilon_2$ moats around each set $\Lambda_{n,k}$. In particular, keeping all constants other than n fixed, the logarithm of this term is of order $O(n \log n)$, which disappears in the normalisation because we normalise by n^{-d} with $d \geq 2$. $\square$

Proof of Theorem 3.12.6. Let us consider a configuration $(\phi_1, \phi_2) \in E_n^a(n\delta\varepsilon_2)$, and focus on the collection of moats of $f := \phi_1 - \phi_2$. Fix $x \in \mathcal{L}$ with $u_1(x) - u_2(x) \neq 0$, and define $L := \mathbb{Z}x$ and $L_N := \{-N, \ldots, N\}x$. Write $\bar{L}_N$ for a path through the square lattice of minimal length traversing all the vertices in L_N. Draw some vertex z from $\mathcal{L} \cap D_n$ uniformly at random, and write $L^z := L + z$, $L_N^z := L_N + z$, and $\bar{L}_N^z := \bar{L}_N + z$. We are interested in the line L^z, and the way this line intersects the moats of f. We make a series of important geometrical observations. By saying that a quantity is *uniformly positive*, we mean that it has a strictly positive lower bound which is independent of the four parameters, for n sufficiently large and for ε_3, ε_2, and ε_1 sufficiently small.

1. If a is at least $(4 \vee 2m)K$, then the d_1-distance from the inside to the outside of a fixed climbing or descending $4K, 4K + a$-moat is at least $\lfloor m \rfloor + 1$, as f is $2K$-Lipschitz (See Proposition 3.5.3, Statement 3). If $m \geq d_1(0, x)$ and if L_N^z intersects both the inside and outside of some moat, then L_N^z must also intersect that moat. In particular, if L^z intersects $\Lambda_{n,k}$, then L^z must necessarily also intersect all moats surrounding $\Lambda_{n,k}$. In the sequel, we choose $a' := (4 \vee 2d_1(0, x))K$ and $a = 3a'$.

2. With uniformly positive probability, z lies in $\Delta_{n,k}$ with L^z intersecting $\Lambda_{n,k}$, for some odd integer k. This is illustrated by Figure 3.6; it is important here that $(u_1 - u_2)(x) \neq 0$ so that x does not lie in the hyperplane $\{u_1 - u_2 = 0\}$. Let us suppose that such an odd integer k indeed exists. Write $m^\pm$ for the smallest and largest integer respectively such that $z + m^\pm x \in \Delta_{n,k}$. Then $m^+ - m^- \leq O(n\varepsilon_2)$, where the constant is independent of all four parameters. But $\Delta_{n,k}$ contains a sequence of $\lceil n\delta\varepsilon_2 \rceil$ nested $4K, 4K + a$-moats of $\Lambda_{n,k}$; L^z intersects each one of them. These moats thus have a uniformly positive density in the set $z + \{m^-, \ldots, m^+\}x$. But z was chosen uniformly random from $\mathcal{L} \cap D_n$ and therefore we may rerandomise its position within $z + \{m^-, \ldots, m^+\}x$. Since the moats are disjoint from one another and have a positive density within this set, we observe there exists a fixed constant $N \in \mathbb{N}$ such that L_N^z intersects at least five distinct nested moats with uniformly positive probability. In fact, each $4K, 4K + a$-moat contains a $4K, 4K + a'$-moat (Proposition 3.5.3, Statement 7), and z is contained in such a moat with uniformly positive probability. Therefore, the event that L_N^z intersects at least five distinct nested $4K, 4K + a$-moats, and simultaneously $f(z) \in [4K, 4K + a')$, has uniformly positive probability.

3. Let us mention a first consequence of the event described above. Since L_N^z intersects more than three distinct $4K, 4K + a$-moats, it must intersect both the inside and outside of the middle moat. This moat contains both a $4K, 4K + a'$-moat, as well as a $4K + 2a', 4K + 3a'$-moat, which L_N^z must both intersect. The value of f differs by at least $a' \geq 4K$ on these two moats. In particular, $\xi = \xi(\phi_1, \phi_2, U)$ cannot be constant on L_N^z, regardless of the value of U. Similarly, $\xi(\theta_z\phi_1, \theta_z\phi_2, U)$ cannot be constant on L_N.

4. Let us mention a second consequence. Since the set L_N^z intersects five distinct nested $4K, 4K + a$-moats, it must intersect both the inside and the outside of the three middle moats. Fix $U \in [0, 4K)$, and write $a'' := f(z) + U \in [4K, 4K + 2a')$. The set $\bar{L}_N^z$ must intersect three $a'', a'' + 4K$-moats: each of the three middle $4K, 4K + a$-moats contains a $a'', a'' + 4K$-moats which $\bar{L}_N^z$ must also intersect. But these three moats correspond exactly to connected components of $\{\xi = 0\}$ for $\xi := \xi(\theta_z\phi_1, \theta_z\phi_2, U)$, which are intersected by $\bar{L}_N$. We must however limit ourselves to local observations, as we always work in the topology of (weak) local convergence. Write therefore $\Sigma_m := \{-m, \ldots, m\}^d \subset\subset \mathbb{Z}^d$; we only consider m so large that $\bar{L}_N \subset \Sigma_m$. The previous observation means that for any $m \in \mathbb{N}$, $\{\xi = 0\} \cap \Sigma_m$ has three connected components which intersect both $\bar{L}_N$ and $\partial^1\Sigma_m$, at least if n is sufficiently large—this is because each moat must surround some set $\Lambda_{n,k}$, which grows large whenever n is large.

Let us summarise what we have done so far. We proved that there exist constants $N \in \mathbb{N}$ and $\delta' > 0$ with the following properties. Choose $(\phi_1, \phi_2) \in E_n^a(n\delta\varepsilon_2)$, and choose $z \in \mathcal{L} \cap D_n$ uniformly at random. Then for fixed $m \in \mathbb{N}$, the probability that for any $U \in [0, 4K)$,

1. $\xi := \xi(\theta_z\phi_1, \theta_z\phi_2, U)$ is not constant on L_N,

2. $\{\xi = 0\} \cap \Sigma_m$ has three connected component which intersect both $\bar{L}_N$ and $\partial^1\Sigma_m$,

is at least δ', for n sufficiently large depending on m, and for ε_3, ε_2, and ε_1 small.

In the final part of the proof, we use this intermediate result, as well as the large deviations principle and compactness of the lower level sets M_C of the specific free energy, to construct the desired measure for Theorem 3.12.6.

Let us first consider the case $E = \mathbb{Z}$. Consider $m \in \mathbb{N}$ so large that $\bar{L}_N \subset \Sigma_m$, and write $A_m \in \mathcal{F}^\nabla_{\Sigma_m} \times \mathcal{F}^\nabla_{\Sigma_m}$ for the event that for any $U \in [0, 4K)$, the function $\xi := \xi(\phi_1, \phi_2, U)$ is not constant on L_N, and that $\{\xi = 0\} \cap \Sigma_m$ has three connected components intersecting both $\partial^1 \Sigma_m$ and $\bar{L}_N$. Write B_m for the set of measures $\mu \in \mathcal{P}^2(\Omega, \mathcal{F}^\nabla)$ such that $\mu(A_m) > \delta'/2$. Note that B_m is in the basis for the tolopogy of weak local convergence on the space of product measures $\mathcal{P}^2(\Omega, \mathcal{F}^\nabla)$. Recall the definition of $\mathfrak{L}_n(\phi)$ in Subsection 3.11.1, and define, for the product setting,

$$\mathfrak{L}^2_n(\phi_1, \phi_2) := \int_D \delta_{(x, \theta_{[nx]_\mathcal{L}} \phi_1, \theta_{[nx]_\mathcal{L}} \phi_2)} dx \in \mathcal{M}^D_2,$$

where by $\mathcal{M}^D_2$ we mean the set of measures in $\mathcal{M}(D \times \Omega \times \Omega, \mathcal{D} \times \mathcal{F}^\nabla \times \mathcal{F}^\nabla)$ for which the first marginal equals the Lebesgue measure on D. By Lemma 3.12.9 and the intermediate result, we know that

$$n^{-d} \log \gamma^2_n(\mathfrak{L}^2_n(D, \cdot) \in B_m) = o(1)$$

as $n \to \infty$. It therefore follows from the large deviations principle that $\bar{B}_m$ contains a shift-invariant measure $\mu_m \in \mathcal{P}^2_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with $S^2(\mu_m) = (u, u)$ and $\mathcal{H}^2(\mu_m|\Phi) \leq 2\sigma(u)$. In particular, this means that $\mu_m(A_m) \geq \delta'/2$, and in fact $\mu_m(A_{m'}) \geq \delta'/2$ for all $m' \leq m$ because $A_m \subset A_{m'}$ for $m' \leq m$. By compactness of the lower level sets of the specific free energy, the sequence $(\mu_m)_{m \in \mathbb{N}}$ has a subsequential limit $\mu \in \mathcal{P}^2_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ in the topology of local convergence which satisfies $S^2(\mu) = (u, u)$ and $\mathcal{H}^2(\mu|\Phi) \leq 2\sigma(u)$. In particular, $\mu(A_m) \geq \delta'/2$ for all m, which means that μ satisfies all the requirements of Theorem 3.12.6; the intersection $\cap_m A_m$ of the decreasing sequence $(A_m)_{m \in \mathbb{N}}$ is precisely the event that $\{\xi = 0\}$ has three infinite level sets which intersect $\bar{L}_N$, regardless of the value of U.

In the case that $E = \mathbb{R}$, there is a slight complication. If $E = \mathbb{R}$, then the indicator 1_{A_m} is not continuous with respect to the topology of uniform convergence on Ω^2, and therefore the sets B_m as defined above are not in the basis of the topology of weak local convergence. Introduce therefore the sequence of functions $(f_{m,k})_{k \in \mathbb{N}}$ where each function $f_{m,k} : \Omega^2 \to [0, 1]$ is defined by $f_{m,k}(\phi_1, \phi_2) := 0 \vee (1 - k d_\infty(A_m, (\phi_1, \phi_2)))$; here d_∞ denotes the metric corresponding to the norm $\|\cdot\|_\infty$ on Ω^2. Write $B_{m,k}$ for the set of product measures μ such that $\mu(f_{m,k}) > \delta'/2$. Then $B_{m,k}$ is in the basis of the topology of weak local convergence, and we have

$$n^{-d} \log \gamma^2_n(\mathfrak{L}^2_n(D, \cdot) \in B_{m,k}) = o(1)$$

as $n \to \infty$. Therefore $\bar{B}_{m,k}$ contains a measure $\mu_{m,k}$ with $S^2(\mu_{m,k}) = (u, u)$ and $\mathcal{H}^2(\mu_{m,k}|\Phi) \leq 2\sigma(u)$. Moreover, the sequence of measures $(\mu_{m,k})_{k \in \mathbb{N}}$ must have a subsequential limit μ_m in the topology of local convergence, and this limit must satisfy $\mathcal{H}^2(\mu_m|\Phi) \leq 2\sigma(u)$, $S^2(\mu_m) = (u, u)$, and $\mu_m(f_{m,k}) \geq \delta'/2$ for all k. The dominated convergence theorem says that $\mu_m(\bar{A}_m) = \mu_m(1_{\bar{A}_m}) = \mu_m(\lim_k f_{m,k}) \geq \delta'/2$. But $\mu_m(\partial A_m) = 0$, since μ_m has finite specific free energy and is therefore locally absolutely continuous with respect to the Lebesgue measure. In particular, $\mu_m(A_m) \geq \delta'/2$. Take now a subsequential limit of the sequence $(\mu_m)_{m \in \mathbb{N}}$ for the desired measure. For this last step, it is important that the topology of local convergence and the topology of weak local convergence coincide on the lower level sets of the specific free energy. $\square$

3.12.3 Application of the argument of Burton and Keane

In this subsection we prove Theorem 3.4.12, which is equivalent to the conjunction of Theorem 3.12.13 and Theorem 3.12.14. Recall the definition of ρ and ξ in the previous subsection.

Lemma 3.12.10. *Let Φ denote any potential in $\mathcal{S}_\mathcal{L}+\mathcal{W}_\mathcal{L}$, and consider a measure $\mu\in\mathcal{P}^2_\mathcal{L}(\Omega,\mathcal{F}^\nabla)$. Then one of the following properties must fail:*

1. *μ is ergodic and at least one of $S(\mu_1)$ and $S(\mu_2)$ lies in U_Φ,*
2. *μ is a minimiser in the sense that $\mathcal{H}^2(\mu|\Phi)=\sigma^2(S^2(\mu))<\infty$,*
3. *With positive $\mu\times\rho$-probability, $\{\xi=0\}$ has at least three infinite components.*

The proof uses a construction which also appears in the part in Chapter 4 on strict convexity.

Proof of Lemma 3.12.10. For a fixed configuration (ϕ_1,ϕ_2,U), a *trifurcation box* is a finite set $\Lambda\subset\subset\mathbb{Z}^d$ such that for some $a\in\mathbb{Z}$, the set $\{\xi=a\}\smallsetminus\Lambda$ has three infinite connected components, which are contained in a single connected component of $\{\xi=a\}$. If μ is shift-invariant then almost surely $\mu\times\rho$ has no trifurcation boxes, due to the argument of Burton and Keane [4]. Note that it is important for this statement that the gradient of ξ is shift-invariant in $\mu\times\rho$. To arrive at the desired contradiction, we aim to prove that trifurcation boxes occur with positive probability for the measure μ described in the statement of the lemma.

Write Ω_q^2 for the set of pairs of q-Lipschitz height functions. The natural adaptation of Theorem 3.4.2 to the product setting asserts that

$$1_{\Omega_q^2}(\lambda^\Lambda\times\lambda^\Lambda\times\mu\pi_{\mathbb{Z}^d\smallsetminus\Lambda})\times\rho\ll\mu\times\rho \tag{3.12.11}$$

for any $\Lambda\subset\subset\mathbb{Z}^d$, where by $\mu\pi_{\mathbb{Z}^d\smallsetminus\Lambda}$ we mean the product measure μ restricted to the vertices in the complement of Λ, as in the non-product setting. Therefore it suffices to demonstrate that trifurcation boxes occur with positive measure in the measure on the left in the display, for some $\Lambda\subset\subset\mathbb{Z}^d$.

Suppose, without loss of generality, that $S(\mu_1)\in U_\Phi$. Write Σ_n for the set $\{-n,\dots,n\}^d\subset\subset\mathbb{Z}^d$, for any $n\in\mathbb{N}$. Then for some fixed $n\in\mathbb{N}$, three infinite components of $\{\xi=0\}$ intersect Σ_n with positive $\mu\times\rho$-probability. Moreover, as μ is ergodic with $S(\mu_1)\in U_\Phi$, we observe that the two functions

$$(\phi_1\pm 8nK)|_{\Sigma_n}\phi_1|_{\mathbb{Z}^d\smallsetminus\Sigma_N} \tag{3.12.12}$$

are q-Lipschitz with high μ-probability as $N\to\infty$. This is due to Lemma 3.6.1, Theorem 3.10.15 and because $S(\mu_1)\in U_\Phi$—recall for comparison the pyramid construction from the proof of Lemma 3.10.14. In particular, for $N\geq n$ sufficiently large, the $\mu\times\rho$-probability that three infinite components of $\{\xi=0\}$ intersect Σ_n and simultaneously the two functions in (3.10.14) are q-Lipschitz, is positive. Now choose $x\in\mathcal{L}$ such that $0\not\in\Sigma_N+x$, and write $\Sigma_n':=\Sigma_n+x$ and $\Sigma_N':=\Sigma_N+x$. Due to shift-invariance, have now proven that with positive $\mu\times\rho$-probability, Σ_n' intersects three connected components of $\{\xi=a\}$ for some $a\in\mathbb{Z}$, and the two functions in (3.12.12) are q-Lipschitz for Σ_n and Σ_N replaced by Σ_n' and Σ_N' respectively. Let us write A for this event.

Let us first discuss the discrete setting $E = \mathbb{Z}$. If $(\phi_1, \phi_2, U) \in A$, then there exists another q-Lipschitz function $\phi_1' \in \Omega$ which equals ϕ_1 on the complement of Σ_N', and such that $\{\xi = a\} \cup \Sigma_n' \subset \{\xi' = a\}$ where $\xi' := (\phi_1', \phi_2, U)$. In particular, this means that Σ_N' is a trifurcation box for ξ'. For example, one can take ϕ_1' to be the smallest q-Lipschitz extension of $\phi_1|_{\mathbb{Z}^d \smallsetminus \Sigma_N'}$ to $\mathbb{Z}^d$ which equals at least

$$\phi_2 + 4Ka + U + (\phi_1(0) - \phi_2(0))$$

on $\{\xi = a\} \cup \Sigma_n'$. This proves that the event that Σ_N' is a trifurcation box has positive measure in the measure on the left in (3.12.11) if we choose $\Lambda = \Sigma_n'$. If $E = \mathbb{R}$, then we must show that not only such a q-Lipschitz function ϕ_1' exists, but also that the set of such functions ϕ_1' has positive Lebesgue measure. The original measure μ has finite specific free energy and therefore almost surely the height functions ϕ_1 and ϕ_2 are not taut, that is, for every $\Lambda \subset\subset \mathbb{Z}^d$ there almost surely exists a positive constant $\varepsilon > 0$ such that the restriction of ϕ_1 and ϕ_2 to Λ are q_ε-Lipschitz. Now choose Λ so large that $\Sigma_N' \subset \Lambda^{-R}$, choose ε at least so small that $S(\mu_1) \in U_{q_\varepsilon}$, and construct the initial height function ϕ_1' such that it is also q_ε-Lipschitz. It is easy to see that one can employ the remaining flexibility granted by Proposition 3.6.5, Statement 3 to demonstrate that the set of of suitable height functions has positive Lebesgue measure. □

Theorem 3.12.13. *Let Φ denote a potential which is monotone and in $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$. If $E = \mathbb{R}$, then σ is strictly convex on U_Φ.*

Proof. Let μ denote the measure from Theorem 3.12.6, and write w_μ for its ergodic decomposition. The measure μ satisfies $\mathcal{H}^2(\mu|\Phi) = \sigma^2(S^2(\mu)) < \infty$, and both $\mathcal{H}^2(\cdot|\Phi)$ and $S^2(\cdot)$ are strongly affine. This implies that w_μ-almost every measure ν satisfies $\mathcal{H}^2(\nu|\Phi) = \sigma^2(S^2(\nu)) < \infty$. Since $E = \mathbb{R}$, this implies also that $S(\nu_1), S(\nu_2) \in U_\Phi$.

With positive w_μ-probability, the $\nu \times \rho$-probability that $\{\xi = 0\}$ has at least three distinct infinite connected components, is positive. We have now proven the existence of a measure which satisfies all criteria of Lemma 3.12.10. This is the desired contradiction. □

Theorem 3.12.14. *Let Φ denote a potential which is monotone and in $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$. Consider now the discrete case $E = \mathbb{Z}$. Suppose that σ satisfies the following property: for any affine map $h : (\mathbb{R}^d)^* \to \mathbb{R}$ such that $h \leq \sigma$, the set $\{h = \sigma\} \cap \partial U_\Phi$ is convex. Then σ is strictly convex on U_Φ. In particular, σ is strictly convex on U_Φ if at least one of the following conditions is satisfied:*

1. *σ is affine on ∂U_Φ, but not on $\bar{U}_\Phi$,*

2. *σ is not affine on $[u_1, u_2]$ for any distinct $u_1, u_2 \in \partial U_\Phi$ such that $[u_1, u_2] \not\subset \partial U_\Phi$.*

Proof. Suppose that σ satisfies the property in the statement. Let $h : (\mathbb{R}^d)^* \to \mathbb{R}$ denote an affine map such that $h \leq \sigma$, and such that the set $\{h = \sigma\} \cap U_\Phi$ contains at least two slopes. We aim to derive a contradiction.

Let us first cover the case that $\{h = \sigma\} \subset U_\Phi$. Let μ denote the measure from Theorem 3.12.6, with slope $S(\mu) = (u, u)$ for some $u \in \{h = \sigma\}$. Write w_μ for the ergodic decomposition of μ. Then w_μ-almost surely $S(\nu_1), S(\nu_2) \in \{h = \sigma\} \subset U_\Phi$, and therefore the proof is the same as for the real case.

Let us now discuss the case that $\{h = \sigma\}$ intersects ∂U_Φ. Recall Lemma 3.6.1. Since $\{h = \sigma\} \cap \partial U_\Phi$ is convex, this intersection must be contained in the boundary of one of the half-spaces $H = H(p)$ contributing to the intersection in Lemma 3.6.1, where $p = (p_k)_{0\leq k\leq n}$ is a path of finite length through $(\mathbb{Z}^d, \mathbb{A})$ with $p_n - p_0 \in \mathcal{L}$. Set $y := p_0$ and $x := p_n - p_0$. If a shift-invariant measure in $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ has finite specific free energy and its slope in $\partial H(p)$, then the random function ϕ must satisfy

$$\phi(y + kx) - \phi(y) = kq(p) := k \sum_{k=1}^{n} q(p_{k-1}, p_k) \tag{3.12.15}$$

for any $k \in \mathbb{Z}$ almost surely. As x is orthogonal to $\partial H(p)$, it is straightforward to find two distinct slopes $u_1, u_2 \in \{h = \sigma\} \cap U_\Phi$ such that $(u_1 - u_2)(x) \neq 0$.

Let μ denote the measure from Theorem 3.12.6, and write w_μ for its ergodic decomposition. The measure μ satisfies $\mathcal{H}^2(\mu|\Phi) = \sigma^2(S^2(\mu))$, and both $\mathcal{H}^2(\cdot|\Phi)$ and $S^2(\cdot)$ are strongly affine. This implies that w_μ-almost every measure ν satisfies $\mathcal{H}^2(\nu|\Phi) = \sigma^2(S^2(\nu)) < \infty$. We know that w_μ-almost surely $S(\nu_1)$ and $S(\nu_2)$ lie in $\{h = \sigma\} \subset \bar{U}_\Phi$, but it is not guaranteed that these slopes lie in U_Φ.

With positive w_μ-probability, the $\nu \times \rho$-probability that ξ is not constant on $y + \mathbb{Z}x$ and that $\{\xi = 0\}$ has at least three distinct infinite connected components, is positive. But if ξ is not constant on $y + \mathbb{Z}x$, then (3.12.15) is false for ϕ having the distribution of either ν_1 or ν_2, or both, and consequently at least one of $S(\nu_1)$ and $S(\nu_2)$ does not lie in $\partial H(p)$. Conclude that with positive w_μ-probability, at least one of $S(\nu_1)$ and $S(\nu_2)$ lies in U_Φ, and the $\nu \times \rho$-probability that $\{\xi = 0\}$ has three or more infinite connected components, is positive. We have now proven the existence of a measure which satisfies all criteria of Lemma 3.12.10. This is the desired contradiction. □

3.13 Applications

3.13.1 The Holley criterion

Each time we apply the theory, we must verify that the specification associated to the model of interest is monotone. An interesting property of stochastic monotonicity is that it does not depends on any formalism and can be checked through the Holley criterion. This criterion is usually stated in the context of the Ising model or Fortuin-Kasteleyn percolation (see for example [25]) but can be extended to random surfaces in a straightforward way. Throughout this section, we will use this criterion in combination with Theorem 3.4.12 to prove the strict convexity of the surface tension for various interesting models.

Theorem 3.13.1 (Holley criterion)**.** *The potential $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ is monotone if and only if for any two q-Lipschitz functions $\phi, \psi \in \Omega$ with $\phi \leq \psi$ and for any $x \in \mathbb{Z}^d$, we have*

$$\gamma_{\{x\}}(\cdot, \phi) \preceq \gamma_{\{x\}}(\cdot, \psi).$$

Proof. Choose ϕ and ψ as in the statement of the theorem, and consider $\Lambda \subset\subset \mathbb{Z}^d$. We aim to demonstrate that

$$\gamma_\Lambda(\cdot, \phi) \preceq \gamma_\Lambda(\cdot, \psi).$$

Write κ_Λ for the probability kernel associated with Glauber dynamics, that is,

$$\kappa_\Lambda := |\Lambda|^{-1} \sum_{x\in\Lambda} \gamma_{\{x\}}.$$

It is clear under the assumption of the theorem that κ_Λ preserves the partial order $\preceq$ on q-Lipschitz measures. Claim now that

$$\mu\kappa_\Lambda^n \to \mu\gamma_\Lambda$$

in the strong topology as $n \to \infty$ for any q-Lipschitz probability measure μ; this would indeed imply the theorem. This is a standard fact in probability theory. The only detail requiring attention is that it is necessary for any q-Lipschitz function ϕ, that $\gamma_\Lambda(\cdot, \phi)$-almost every height function ψ is accessible from ϕ by local moves, that is, by updating the value of ϕ by one vertex in Λ at a time, and such that all intermediate functions are also q-Lipschitz. This is straightforward to check from the definition of q—in particular, it is important that $q(x, y) + q(y, x) > 0$ for any $x, y \in \mathbb{Z}^d$ distinct. □

3.13.2 Submodular potentials

A potential Φ is said to be *submodular* if for every $\Lambda \subset\subset \mathbb{Z}^d$, Φ_Λ has the property that

$$\Phi_\Lambda(\phi \wedge \psi) + \Phi_\Lambda(\phi \vee \psi) \leq \Phi_\Lambda(\phi) + \Phi_\Lambda(\psi).$$

Sheffield proposes this family of potentials as a natural generalisation of simply attractive potentials, and asks if similar results as the ones proved for simply attractive potentials in [54] could be proved for finite-range submodular potentials. It is easy to see that submodular potentials generate monotone specifications.

Lemma 3.13.2. *A submodular potential is monotone.*

Proof. Let $\phi_1, \phi_2 \in \Omega$ denote q-Lipschitz functions with $\phi_1 \leq \phi_2$. It suffices to check the Holley criterion (Theorem 3.13.1). Write f_i for the Radon-Nikodym derivative of $\gamma_{\{x\}}(\cdot, \phi_i)\pi_{\{x\}}$ with respect to λ, for $i \in \{1, 2\}$. It suffices to demonstrate that $f_1\lambda \preceq f_2\lambda$ as measures on $(E, \mathcal{E})$. Submodularity of Φ implies that $f_1(b)f_2(a) \leq f_1(a)f_2(b)$ for $\lambda \times \lambda$-almost every $a, b \in E$ with $a \leq b$. It is a simple exercise to see that this implies the desired stochastic domination. □

If $E = \mathbb{R}$ and Φ a submodular Lipschitz potential fitting the framework of this thesis (which is a very mild requirement), then we derive immediately from Theorem 3.4.12 that the surface tension is strictly convex.

Corollary 3.13.3. *Suppose that $E = \mathbb{R}$ and consider a submodular Lipschitz potential $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$. Then σ is strictly convex on U_Φ.*

In the remainder of this section, we focus on the case $E = \mathbb{Z}$. If $E = \mathbb{Z}$, then we cannot immediately conclude that the surface tension is strictly convex, because we must fulfill the additional condition in Theorem 3.4.12. We demonstrate how to derive this extra condition for many natural discrete models. Let $(\mathbb{A}, q)$ denote the local Lipschitz constraint associated with the potential of interest and fix $R \in \mathbb{N}$ minimal subject to $d_1(x, y) \leq R$ for all $\{x, y\} \in \mathbb{A}$.

A measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ is called *frozen* if for any $\Lambda \subset\subset \mathbb{Z}^d$, the values of the random function ϕ_Λ in μ depend deterministically on the boundary values $\phi_{\partial^R\Lambda}$. Call a local Lipschitz constraint *freezing* if any measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ which is supported on q-Lipschitz functions, and which has $S(\mu) \in \partial U_\Phi$, is frozen. This condition on the local Lipschitz constraint implies that any such measure has zero specific entropy, that is, $\mathcal{H}(\mu|\lambda) = 0$. Indeed, deterministic dependence implies that

$$\mathcal{H}_{\mathcal{F}^{\nabla}_{\Pi_n}}(\mu|\lambda^{\Pi_n - 1}) = \mathcal{H}_{\mathcal{F}^{\nabla}_{\partial^R \Pi_n}}(\mu|\lambda^{\partial^R \Pi_n - 1}) = O(n^{d-1}) = o(n^d)$$

as $n \to \infty$.

Lemma 3.13.4. *If the local Lipschitz constraint $(\mathbb{A}, q)$ is invariant by the full lattice $\mathcal{L} = \mathbb{Z}^d$, then it is freezing. In particular, the local Lipschitz constraints corresponding to dimer models, the six-vertex model, and Kd_1-Lipschitz functions for $K \in \mathbb{N}$, are freezing.*

Proof. Fix $\mu \in \mathcal{P}_{\mathcal{L}}(\Omega, \mathcal{F}^{\nabla})$ with $S(\mu) \in \partial U_\Phi$ and supported on q-Lipschitz functions. As in the proof of Theorem 3.12.14, there is a path $p = (p_k)_{0 \leq k \leq n}$ of finite length through $(\mathbb{Z}^d, \mathbb{A})$ with $x := p_n - p_0 \in \mathcal{L} \smallsetminus \{0\}$, such that

$$\phi(p_0 + y + kx) - \phi(p_0 + y)$$

is deterministic in μ for any $y \in \mathcal{L}$ and $k \in \mathbb{Z}$. Moreover, this path is a *cycle lift* as defined in the proof of Lemma 3.6.1. Since $\mathcal{L} = \mathbb{Z}^d$, this means that $\phi(y + kx) - \phi(y)$ is deterministic for any $y \in \mathbb{Z}^d$, and that $d_1(0, x) \leq R$. In particular, ϕ_Λ depends deterministically on $\phi_{\partial^R \Lambda}$ in μ for any $\Lambda \subset\subset \mathbb{Z}^d$. $\square$

The final goal of this section is to prove the following theorem.

Theorem 3.13.5. *Suppose that $E = \mathbb{Z}$, and that $\Phi \in \mathcal{S}_{\mathcal{L}} + \mathcal{W}_{\mathcal{L}}$ is a submodular Lipschitz potential with a freezing local Lipschitz constraint. Then the associated surface tension σ is strictly convex on U_Φ.*

We first prove two auxiliary lemmas.

Lemma 3.13.6. *If $E = \mathbb{Z}$ and Φ a submodular gradient potential, then*

$$\Phi_\Lambda(\lceil \tfrac{\phi_1 + \phi_2}{2} \rceil) + \Phi_\Lambda(\lfloor \tfrac{\phi_1 + \phi_2}{2} \rfloor) \leq \Phi_\Lambda(\phi_1) + \Phi_\Lambda(\phi_2)$$

for any $\phi_1, \phi_2 \in \Omega$ and $\Lambda \subset\subset \mathbb{Z}^d$.

Proof. Write $\xi^{\pm} := \phi_1 \pm \phi_2$, so that $\phi_1 = (\xi^+ + \xi^-)/2$ and $\phi_2 = (\xi^+ - \xi^-)/2$. Write

$$F(\psi^+, \psi^-) := \Phi_\Lambda(\tfrac{\psi^+ + \psi^-}{2}) + \Phi_\Lambda(\tfrac{\psi^+ - \psi^-}{2})$$

for any $\psi^+, \psi^- \in \Omega$ with $\psi^+ + \psi^- \equiv 0 \mod 2$. For example, the right hand side of the display in the statement of the lemma equals $F(\xi^+, \xi^-)$, and the left hand side equals $F(\xi^+, p \circ \xi^-)$, where $p : \mathbb{Z} \to \{0, 1\}$ is the *parity function* which maps even integers to 0 and odd integers to 1. Therefore it suffices to demonstrate that

$$F(\psi^+, p \circ \psi^-) \leq F(\psi^+, \psi^-)$$

for any $\psi^+, \psi^- \in \Omega$ with $\psi^+ + \psi^- \equiv 0 \mod 2$.

Observe that F has the following four properties:

1. *Translation invariance*: $F(\psi^+ + a_1, \psi^- + a_2) = F(\psi^+, \psi^-)$ for any $a_1, a_2 \in \mathbb{Z}$ with $a_1 + a_2$ even, because Φ is a gradient specification,
2. *Inversion invariance*: $F(\psi^+, -\psi^-) = F(\psi^+, \psi^-)$; replacing ψ^- by $-\psi^-$ corresponds to interchanging the sum and difference of ψ^+ and ψ^-,

3. *Submodularity*: $F(\psi^+, |\psi^-|) \leq F(\psi^+, \psi^-)$; equivalent to submodularity of Φ,

4. *Locally measurable*: $F(\psi^+, \psi^-)$ depends on $\psi_\Lambda^\pm$ only.

By applying the three operations on the pair (ψ^+, ψ^-) finitely many times, one can turn the original pair into a new pair $(\psi^+, \hat{\psi}^-)$, where $\hat{\psi}_\Lambda^- = (p \circ \psi^-)_\Lambda$. In particular, since each operation can only decrease the value of F, we have

$$F(\psi^+, p \circ \psi^-) = F(\psi^+, \hat{\psi}^-) \leq F(\psi^+, \psi^-)$$

as desired. $\square$

Corollary 3.13.7. *Suppose that $E = \mathbb{Z}$ and that $\Phi \in \mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ is submodular. If $\mu_1, \mu_2 \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ are ergodic, then there exists an ergodic measure $\nu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with*

$$S(\nu) = \frac{S(\mu_1) + S(\mu_2)}{2} \qquad \text{and} \qquad \langle \nu | \Phi \rangle \leq \frac{\langle \mu_1 | \Phi \rangle + \langle \mu_2 | \Phi \rangle}{2}.$$

Proof. Write $\hat{\mu} \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ for the following measure: to sample from $\hat{\mu}$, sample first a pair (ϕ_1, ϕ_2) from $\mu_1 \times \mu_2$, and sample X from $\{0, 1\}$ independently and uniformly at random; the final sample ψ from $\hat{\mu}$ is now defined by

$$\psi := \begin{cases} \lceil \frac{\phi_1 - \phi_1(0) + \phi_2 - \phi_2(0)}{2} \rceil & \text{if } X = 0, \\ \lfloor \frac{\phi_1 - \phi_1(0) + \phi_2 - \phi_2(0)}{2} \rfloor & \text{if } X = 1. \end{cases}$$

Since $\phi_1 - \phi_1(0)$ and $\phi_2 - \phi_2(0)$ are asymptotically close to $S(\mu_1)$ and $S(\mu_2)$ respectively in the measure $\mu_1 \times \mu_2$ in the sense of Theorem 3.10.15, it is clear that ψ is asymptotically close to $(S(\mu_1) + S(\mu_2))/2$ in $\hat{\mu}$ (see also the proof of Lemma 3.10.32). In particular, $S(\nu) = (S(\mu_1) + S(\mu_2))/2$ for $w_{\hat{\mu}}$-almost every ν in the ergodic decomposition of $\hat{\mu}$. By the previous lemma, we have

$$\langle \hat{\mu} | \Phi \rangle \leq \frac{\langle \mu_1 | \Phi \rangle + \langle \mu_2 | \Phi \rangle}{2}.$$

As $\langle \cdot | \Phi \rangle$ is strongly affine, we have $\langle \nu | \Phi \rangle \leq \langle \hat{\mu} | \Phi \rangle$ with positive $w_{\hat{\mu}}$-probability. This proves the existence of the desired measure ν. $\square$

Lemma 3.13.8. *Consider the case that $E = \mathbb{Z}$, Φ a potential in $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$, and μ an ergodic minimiser with $S(\mu) \in U_\Phi$. Then $\mathcal{H}(\mu|\lambda) < 0$.*

Proof. Suppose that μ does have zero combinatorial entropy; we aim to derive a contradiction. Write $u := S(\mu)$, and write $\hat{\mu} \in \mathcal{P}_\mathcal{L}^2(\Omega, \mathcal{F}^\nabla)$ for the unique measure which has μ as its first marginal, and in which ϕ_1 and ϕ_2 are equal almost surely. Then $S^2(\hat{\mu}) = (u, u)$ and $\mathcal{H}^2(\hat{\mu}|\Phi) = 2\langle \mu | \Phi \rangle = 2\mathcal{H}(\mu|\Phi) = \sigma^2(S^2(\hat{\mu})) < \infty$, that is, $\hat{\mu}$ is a minimiser in the product setting. The adaptation of Theorem 3.4.2 to the product setting implies that

$$1_{\Omega_q^2}(\hat{\mu}\pi_{\mathbb{Z}^d \smallsetminus \Lambda} \times \lambda^\Lambda \times \lambda^\Lambda) \ll \hat{\mu}$$

for any $\Lambda \subset\subset \mathbb{Z}^d$, where Ω_q^2 is the set of pairs of q-Lipschitz height functions. However, since μ is ergodic with slope in U_Φ, we can find some $\Lambda \subset\subset \mathbb{Z}^d$ such that with positive $\hat{\mu}$-probability $\phi_1|_{\mathbb{Z}^d \smallsetminus \Lambda}$ has more than a single q-Lipschitz extension to $\mathbb{Z}^d$. This contradicts that ϕ_1 and ϕ_2 are almost surely equal in $\hat{\mu}$. $\square$

We are now ready to prove the second main theorem of this section.

Proof of Theorem 3.13.5. Recall Theorem 3.4.12. If σ is not strictly convex, then there is an affine map $h : (\mathbb{R}^d)^* \to \mathbb{R}$ with $h \leq \sigma$ and such that $\{h = \sigma\} \cap \partial U_\Phi$ is not convex. Write H for the exposed points of $\{h = \sigma\} \subset (\mathbb{R}^d)^*$ which are also in ∂U_Φ. Then the convex envelope of H intersects U_Φ.

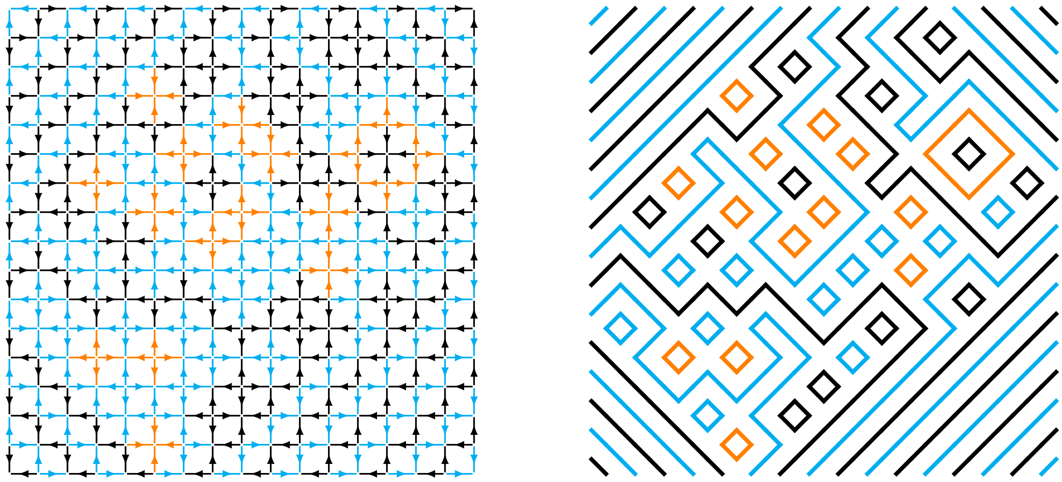
Note that each slope in H is also an exposed point of σ. This means that for each slope in H, there is an ergodic minimiser μ of that slope. Moreover, since $\mathcal{H}(\mu|\lambda) = 0$ for any μ with $S(\mu) \in H \subset \partial U_\Phi$, we must have $\langle \mu|\Phi\rangle = h(S(\mu)) = \sigma(S(\mu))$ for any such measure μ. The fact that the convex envelope of H intersects U_Φ, together with Corollary 3.13.7, implies that there exists an ergodic measure $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\nabla)$ with $S(\mu) \in U_\Phi$ and $\langle \mu|\Phi\rangle \leq h(S(\mu)) \leq \sigma(S(\mu))$. But it is only possible that $\langle \mu|\Phi\rangle \leq \sigma(S(\mu))$ if $\langle \mu|\Phi\rangle = \sigma(S(\mu))$ and if μ is a minimiser with $\mathcal{H}(\mu|\lambda) = 0$. This contradicts Lemma 3.13.8. $\square$

3.13.3 Tree-valued graph homomorphisms

The flexibility of the main theorem in this chapter can also be used to prove statements about the behaviour of random functions taking values in target spaces other than $\mathbb{Z}$ and $\mathbb{R}$. A noteworthy example is the model of tree-valued graph homomorphisms described in [44]. Let $k \geq 2$ denote a fixed integer, and let $\mathcal{T}_k$ denote the k-regular tree, that is, a tree in which every vertex has exactly k neighbours. In this context, tree-valued graph homomorphisms are functions from $\mathbb{Z}^d$ to the vertices of $\mathcal{T}_k$ which also map the edges of the square lattice to the edges of the tree. Regular trees are natural objects in several fields of mathematics: in group theory, for example, they arise as Cayley graphs of free groups on finitely many generators. As a significant result in [44], the authors characterise the surface tension for the model (there named *entropy*) and show that it is equivalent to the number of graph homomorphisms with nearly-linear boundary conditions. In this section we will confirm the conjecture from [44], which states that this entropy function is strictly convex. We must first show how the model and the corresponding surface tension fit into the framework of this thesis. A tree-valued graph homomorphism can be represented by an integer-valued graph homomorphism after introducing an infinite-range potential to compensate for the "loss of information".

Let us first introduce some definitions. Write $d_{\mathcal{T}_k}$ for the graph metric on $\mathcal{T}_k$. Let g denote a fixed *bi-infinite geodesic* through $\mathcal{T}_k$, that is, a $\mathbb{Z}$-indexed sequence of vertices $g = (g_n)_{n\in\mathbb{Z}} \subset \mathcal{T}_k$ such that $d_{\mathcal{T}_k}(g_n, g_m) = |m - n|$ for any $n, m \in \mathbb{Z}$. Let $p : \mathcal{T}_k \to \mathbb{Z}$ denote the projection of the tree onto g, defined such that $p(x)$ minimises $d_{\mathcal{T}_k}(x, g_{p(x)})$ for any $x \in \mathcal{T}_k$. Write h for the *horocyclic height function* on $\mathcal{T}_k$; this is the function $h : \mathcal{T}_k \to \mathbb{Z}$ defined by $h(x) := p(x) + d_{\mathcal{T}_k}(x, g_{p(x)})$ (see also [26]). In other words, if $x = g_n$ for some $n \in \mathbb{Z}$, then $h(x) = n$, and h increases by one every time one moves away from the geodesic g. The function h can also be characterised as follows: each vertex $x \in \mathcal{T}_k$ has a unique neighbour y such that $h(y) = h(x) - 1$, and $h(z) = h(x) + 1$ for every other neighbour z of x.

The graphs $\mathbb{Z}^d$, $\mathbb{Z}$, and $\mathcal{T}_k$ are bipartite, we shall call the two parts the *even vertices* and *odd vertices* respectively; the set of even vertices is the part containing 0 if the graph is $\mathbb{Z}^d$ or $\mathbb{Z}$, and the part containing g_0 if the graph is $\mathcal{T}_k$. By a *graph homomorphism* we mean a map from $\mathbb{Z}^d$ to $\mathbb{Z}$ or $\mathcal{T}_k$ which preserves the parity of the vertices, and which maps edges to edges. Write Ω and $\tilde{\Omega}$ respectively for the set of graph homomorphisms from $\mathbb{Z}^d$ to either $\mathbb{Z}$ or $\mathcal{T}_k$. For fixed $\phi \in \Omega$ and $n \in \mathbb{Z}$,



The gradient of the graph homomorphism — The boundaries of the upper level sets

Figure 3.7: A random $\mathcal{T}_3$-valued graph homomorphism

we call some set $\Lambda \subset\subset \mathbb{Z}^d$ an *n-upper level set* if Λ is a connected component of $\{\phi \geq n\} \subset \mathbb{Z}^d$ in the square lattice graph. An n-upper level set is also called an *n-level set* or simply a *level set*.

Write U for the set of slopes $u \in (\mathbb{R}^d)^*$ such that $|u(e_i)| < 1$ for each element e_i in the natural basis of $\mathbb{R}^d$. For fixed $u \in \bar{U}$, write $\phi^u \in \Omega$ for the graph homomorphism defined by

$$\phi^u(x) := \lfloor u(x) \rfloor + \begin{cases} 0 & \text{if } d_1(0,x) \equiv \lfloor u(x) \rfloor \mod 2, \\ 1 & \text{if } d_1(0,x) \equiv \lfloor u(x) \rfloor + 1 \mod 2, \end{cases}$$

and write $\tilde{\phi}^u \in \tilde{\Omega}$ for the graph homomorphism defined by $\tilde{\phi}^u(x) = g_{\phi^u(x)}$.

It is shown in Section 3 of [44] that the entropy function $\mathrm{Ent} : \bar{U} \to [-\log k, 0]$ associated to the model of graph homomorphisms from $\mathbb{Z}^d$ to $\mathcal{T}_k$ can be estimated by counting for each slope $u \in \bar{U}$ the number of graph homomorphisms $\phi : \mathbb{Z}^d \to \mathcal{T}_k$ which equal $\tilde{\phi}^u$ on the complement of Π_n. More precisely, for $u \in \bar{U}$, we have

$$\mathrm{Ent}(u) = \lim_{n\to\infty} -n^{-d} \log |\{\tilde{\phi} \in \tilde{\Omega} : \tilde{\phi}_{\mathbb{Z}^d \smallsetminus \Pi_n} = \tilde{\phi}^u_{\mathbb{Z}^d \smallsetminus \Pi_n}\}|.$$

Notice that counting the number of functions in this set is similar to considering the normalising constant in the definition of the specification, as we frequently do in this chapter. Before proceeding, let us already remark that $\mathrm{Ent}(u) = 0$ for $u \in \partial U$. Indeed, for such u, the set in the display contains only a single element: the original function $\tilde{\phi}^u$. It is also easy to see that Ent is not identically zero on $\bar{U}$. Consider, for example, the slope $u = 0$, and consider the set of all graph homomorphisms $\tilde{\phi}$ which equal $\tilde{\phi}^u$ on the complement of Π_n and which map all the even vertices of the square lattice to $g_0 \in \mathcal{T}_k$. Then this set contains at least $k^{\lfloor n^d/2 \rfloor}$ functions, proving that $\mathrm{Ent}(u) \leq -\frac{1}{2} \log k < 0$.

We now get to the heart of the case. Let us use the horocyclic height function to count the set in the previous display in a different way. Suppose that some graph homomorphism $\phi \in \Omega$ equals ϕ^u on the complement of Π_n. How many graph homomorphisms $\tilde{\phi} \in \tilde{\Omega}$ do there exist which satisfy $h \circ \tilde{\phi} = \phi$ and equal $\tilde{\phi}^u$ on the complement of Π_n? It turns out that this number must be precisely $(k-1)^{F_{\Pi_n}(\phi)}$, where $F_\Lambda(\phi)$ denotes the number of level sets of ϕ which are entirely contained in Λ, for any $\Lambda \subset\subset \mathbb{Z}^d$. Indeed, each time we see an n-level set of ϕ, the function $\tilde{\phi}$ must be constant on the outer boundary of that n-level set—say with value $x \in \mathcal{T}_k$—and there are $k-1$ neighbours of x which lead to an increase of the horocyclic height function by exactly one. In particular, we have

$$\operatorname{Ent}(u) = \lim_{n\to\infty} -n^{-d} \log \sum_{\phi\in\Omega,\, \phi_{\mathbb{Z}^d\smallsetminus\Pi_n}=\phi^u_{\mathbb{Z}^d\smallsetminus\Pi_n}} (k-1)^{F_{\Pi_n}(\phi)}. \tag{3.13.9}$$

See Figure 3.7 for a sample of the model, with the gradient of the graph homomorphism on the left, and with the boundaries of the level sets of the horocyclic height function on the right. We have now reduced to a problem expressed entirely in terms of integer-valued functions. In fact, we do no longer require k to be an integer, although we do require that $k \geq 2$. In the remainder of this section, we construct a potential Φ which fits into our class $\mathcal{S}_\mathcal{L} + \mathcal{W}_\mathcal{L}$ and which is monotone, and such that $U_\Phi = U$ and $\sigma = \operatorname{Ent}$. This proves that σ and Ent are strictly convex on $U_\Phi = U$. In fact, the specification induced by the potential that we construct is not perfectly monotone, but we shall demonstrate that it is sufficiently monotone for us to deduce that σ is strictly convex.

Unfortunately, we cannot hope to use a potential that counts the level sets directly. The reason is that there is no upper bound on the number of level sets containing a single point; such a potential would always sum to infinity. However, each finite level set has a uniquely defined outer boundary, and each vertex is contained in only finitely many outer boundaries. This means that counting outer boundaries of finite level sets is equivalent to counting finite level sets, and the potential that does so is well-defined and fits our framework, as we will show. It is not possible through this method to count infinite level sets, but we shall demonstrate how to work around this apparent difficulty.

We shall now describe how to characterise the outer boundary of a finite level set. This is not entirely straightforward due to the connectivity properties of the square lattice. By the $*$-graph on $\mathbb{Z}^d$, we mean the graph in which two vertices x and y are neighbours if and only if $\|x-y\|_\infty = 1$. For example, each vertex has $3^d - 1$ distinct $*$-neighbours. On every single occasion that we mention a graph-related notion, we mean the usual square lattice graph, unless we explicitly mention the $*$-graph. Due to the connectivity properties of the square lattice, we have the following proposition.

Proposition 3.13.10. *Suppose that $\Lambda \subset\subset \mathbb{Z}^d$ is finite and connected, and that its complement $\Delta := \mathbb{Z}^d \smallsetminus \Lambda$ is $*$-connected. Define $\partial^*\Delta$ to be the set of vertices $x \in \mathbb{Z}^d$ such that:*

1. *Either $x \in \Lambda = \mathbb{Z}^d \smallsetminus \Delta$ and $*$-adjacent to Δ,*
2. *Or $x \in \Delta$ and adjacent to $\Lambda = \mathbb{Z}^d \smallsetminus \Delta$.*

Then $\partial^*\Delta \cap \Lambda = \partial^*\Delta \cap (\mathbb{Z}^d \smallsetminus \Delta)$ is connected, and so is $\partial^*\Delta$.

Consider a finite nonempty connected set $\Lambda \subset\subset \mathbb{Z}^d$. Write Λ^∞ for the *outside* of Λ, that is, the unique unbounded $*$-connected component of the complement of Λ. Write also $\bar{\Lambda}$ for the complement of Λ^∞: this set is finite and connected, and contains Λ. The pair $(\bar{\Lambda}, \Lambda^\infty)$ will play the role of (Λ, Δ) in the previous proposition. The set $\partial^*\Lambda^\infty$ can obviously be written as the disjoint union of $\partial^*\Lambda^\infty \cap \Lambda^\infty$ and $\partial^*\Lambda^\infty \cap \bar{\Lambda}$. Claim that $\partial^*\Lambda^\infty \cap \bar{\Lambda} = \partial^*\Lambda^\infty \cap \Lambda$. Indeed, if $x \in \partial^*\Lambda^\infty \cap \bar{\Lambda}$ is not in Λ, then it should be in Λ^∞ as it is $*$-adjacent to Λ^∞; this proves the claim. This also means that all vertices in $\partial^*\Lambda^\infty \cap \Lambda^\infty$ are adjacent to Λ.

Suppose now that Λ is also an n-level set of some graph homomorphism $\phi \in \Omega$. Then ϕ must equal exactly $n-1$ on $\partial^*\Lambda^\infty \cap \Lambda^\infty$, and ϕ must be at least n on $\partial^*\Lambda^\infty \cap \bar{\Lambda} = \partial^*\Lambda^\infty \cap \Lambda$. We have now proven the following lemma.

Lemma 3.13.11. *Suppose that $\Delta \subsetneq \mathbb{Z}^d$ is $*$-connected and cofinite, with its complement connected. Then*

$$\begin{aligned}\{\phi \in \Omega : \Delta \text{ is the outside of a } n\text{-level set of } \phi \text{ for some } n \in \mathbb{Z}\} \\ = \{\phi \in \Omega : \phi_{\partial^*\Delta \cap \Delta} = n-1 \text{ and } \phi_{\partial^*\Delta \smallsetminus \Delta} \geq n \text{ for some } n \in \mathbb{Z}\} \in \mathcal{F}^\nabla_{\partial^*\Delta}.\end{aligned}$$

Moreover, no two level sets of ϕ produce the same outside boundary $\partial^*\Delta$.

Define the potential $\Xi = (\Xi_\Lambda)_{\Lambda \subset\subset \mathbb{Z}^d}$ by

$$\Xi_\Lambda(\phi) = -\log(k-1)$$

if $\Lambda = \partial^*\Delta^\infty$ for some finite level set Δ of ϕ, and $\Xi_\Lambda(\phi) = 0$ otherwise. For fixed $x \in \mathbb{Z}^d$ and $\phi \in \Omega$, there are at most 3^d finite level sets Δ of ϕ such that $x \in \partial^*\Delta^\infty$. In particular, this means that $\|\Xi\| \leq 3^d \log(k-1)$. Moreover, since $\Xi_\Lambda \equiv 0$ whenever $\Lambda \subset\subset \mathbb{Z}^d$ is not connected, it is clear that $e^-(\Lambda) \leq |\partial\Lambda| \cdot \|\Xi\|$. In particular, e^- is an amenable function, which means that $\Xi \in \mathcal{W}_\mathcal{L}$. Remark that $H^\Xi_\Lambda(\phi)$ equals $-\log(k-1)$ times the number of finite level sets Δ of ϕ for which $\partial^*\Delta^\infty$ intersects Λ. Unfortunately, it is not possible to count infinite level sets with this construction; this is a small inconvenience that we must circumvent.

Write Ψ for the potential which forces graph homomorphisms, that is, $\Psi_\Lambda(\phi) = \infty$ if $\Lambda = \{x, y\}$ is an edge of the square lattice and $|\phi(y) - \phi(x)| \neq 1$, and $\Psi_\Lambda(\phi) = 0$ otherwise. This potential belongs to $\mathcal{S}_\mathcal{L}$, modulo the detail explained in Subsection 3.4.4, which we shall simply ignore here.

Lemma 3.13.12. *For any integer $k \geq 2$, the surface tension σ associated to the potential $\Phi := \Psi + \Xi$ equals the entropy function* Ent*.*

Proof. We prove that $\sigma(u) = \mathrm{Ent}(u)$ for $u \in U_\Phi$, the result extends to all $u \in \bar{U}_\Phi$ because both σ and Ent are continuous on $\bar{U}_\Phi$. Due to Theorem 3.4.10, we know that

$$\begin{aligned}\sigma(u) = P_\Phi((0,1)^d, u|_{\partial(0,1)^d}) &= \lim_{n\to\infty} -n^{-d} \log \int_{E^{\Pi_n}} e^{-H^\Phi_\Lambda(\psi\phi^u_{\mathbb{Z}^d\smallsetminus\Pi_n})} d\lambda^{\Pi_n}(\psi) \\ &= \lim_{n\to\infty} -n^{-d} \log \sum_{\phi\in\Omega,\, \phi_{\mathbb{Z}^d\smallsetminus\Pi_n} = \phi^u_{\mathbb{Z}^d\smallsetminus\Pi_n}} e^{-H^\Xi_{\Pi_n}(\phi)}.\end{aligned}$$

But the logarithm of the ratio of $e^{-H^\Xi_{\Pi_n}(\phi)}$ with $(k-1)^{F_{\Pi_n}(\phi)}$ is of order $O(n^{d-1}) = o(n^d)$ uniformly over ϕ as $n \to \infty$, so that the equality $\sigma(u) = \mathrm{Ent}(u)$ follows from (3.13.9). $\square$

Definition 3.13.13. Write Ω_- for the set of graph homomorphisms $\phi \in \Omega$ which have no infinite level sets.

Lemma 3.13.14. *The specification induced by the potential $\Phi := \Psi + \Xi$ is stochastically monotone over Ω_- for any $k \geq 2$.*

Proof. We use the Holley criterion (Theorem 3.13.1) to prove that γ_Λ preserves $\preceq$; we suppose that $\Lambda = \{0\}$ without loss of generality. Let $\phi_1, \phi_2 \in \Omega_-$ denote graph homomorphisms without infinite level sets, and which satisfy $\phi_1 \leq \phi_2$. Notice that the only case where the local Gibbs measure $\gamma_\Lambda(\cdot, \phi)$ is not a Dirac measure, is if there exist a $n \in \mathbb{Z}$ such that $\phi(x) = n$ for any neighbour x of 0. If this is not the case for ϕ_1 or ϕ_2 then the proof is trivial; we reduce to the case that $\phi_1(x) = \phi_2(x) = 1$ for any neighbour x of 0 in $\mathbb{Z}^d$. It remains to show that $\gamma_\Lambda(\cdot, \phi_1) \preceq \gamma_\Lambda(\cdot, \phi_2)$. Without loss of generality, $\phi_1(0) = \phi_2(0) = 0$.

Write ψ for the random function in either local Gibbs measure. Since $\phi_i(x) = 1$ for any neighbour x of 0 and for $i \in \{1, 2\}$, the function ψ can only take two values with positive probability: they are 0 and 2. What we thus must show is that the quantity

$$a_i := \frac{\gamma_\Lambda(\psi(0) = 2, \phi_i)}{\gamma_\Lambda(\psi(0) = 0, \phi_i)}$$

satisfies $a_1 \leq a_2$. Claim that $a_i = (k-1)^{2-X_i}$, where X_i is the number of 1-level sets of ϕ_i which are adjacent to 0. If $\psi(0) = 0$, then all 1-level sets adjacent to 0 are counted separately, and $\{0\}$ is not a level set. If $\psi(0) = 2$, then we count two level sets: the set $\{0\}$ is a 2-level set, and all neighbours of 0 are contained in the same 1-level set. All other level sets remain unaffected. This proves the claim. We must therefore prove that $X_1 \geq X_2$. This is clear: increasing the values of ϕ can only increase the size of the 1-level set containing a fixed vertex x, and potentially merge several 1-level sets. In particular, it can only decrease the number of 1-level sets adjacent to 0. $\square$

Theorem 3.13.15. *The surface tension σ associated to the potential Φ defined above, is strictly convex on U_Φ whenever $k \geq 2$.*

Proof. We must circumvent the problem that the specification γ induced by Φ is monotone only after restricting it to the set Ω_-. Remark that Theorem 3.5.4 and Proposition 3.5.11 remain true in this context if the measure μ in the statement of Theorem 3.5.4 is supported on Ω_-. The only time that monotonicity is used in the proof for strict convexity of σ, is in the application of these two results in Lemma 3.12.9. Recall that the local Gibbs measure γ_n in the statement of Lemma 3.12.9 was defined to be $\gamma_{\Pi_n}(\cdot, \phi^u)$; this is now problematic because ϕ^u does have infinite level sets. This can be easily solved by the following modification. Define ϕ_n^u to be the smallest graph homomorphism which equals ϕ^u on the set $\Pi_n \cup \partial\Pi_n$. It is easy to check that $\{\phi_n^u \geq m\}$ is finite for any $n \in \mathbb{N}$ and $m \in \mathbb{Z}$; in particular, $\phi_n^u \in \Omega_-$. Moreover, the sequence $(\Pi_n, \phi_n^u)_{n\in\mathbb{N}}$ is as much an approximation of $((0,1)^d, u|_{\partial(0,1)^d})$ as the original sequence $(\Pi_n, \phi^u)_{n\in\mathbb{N}}$. In particular, all of the same arguments apply if we simply replace each local Gibbs measure $\gamma_n = \gamma_{\Pi_n}(\cdot, \phi^u)$ by $\gamma_{\Pi_n}(\cdot, \phi_n^u)$. We had already seen that $\sigma = 0$ on ∂U_Φ and $\sigma(0) < 0$, which proves that σ is strictly convex. $\square$

3.13.4 Stochastic monotonicity in the six-vertex model

Consider the two-dimensional square lattice. An *arrow configuration* is an orientation of each edge of the square lattice, in such a way that each vertex has exactly two incoming edges and two outgoing edges. This means that there are six configurations for the four edges incident to a fixed vertex; see Figure 3.8. Each of these six types receives a weight, and one studies the probability measure where the probability of observing an arrow configuration is proportional to the product of the weights over

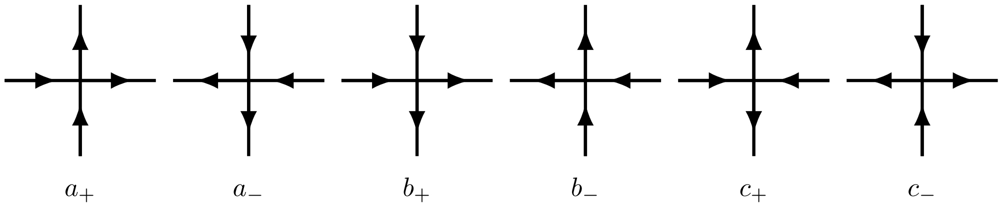


Figure 3.8: The six types of arrow configurations and their weights

the vertices in that configuration. This is the six-vertex model, which is the subject of an extensive literature. Each arrow configuration has an associated height function, which assigns integers to the faces of the square lattice, and is defined as follows: the height of the face to the right of an arrow is always exactly one more than the height of the face to the left of it, and the height of a fixed reference face is set to zero. It is straightforward to see that this uniquely defines the height functions associated to an arrow configuration. The six-vertex model can thus be considered a Lipschitz random surface. Our main theorem asserts that the surface tension of this random surface model is strictly convex, if the specification is monotone. It is a straightforward exercise to demonstrate that the specification is monotone if and only if

$$c_+c_- \geq \max\{a_+a_-, b_+b_-\};$$

this is verified through checking the Holley criterion (Theorem 3.13.1). Informally, this means that the specification is monotone if the model prefers vertices for which the four values of the adjacent faces are as close to each other as possible. Finally, we should mention that from the perspective of the specification, there is some gauge equivalence in the choice of the six weights; for details we refer to the work of Sridhar [55, Section 2.2].

Theorem 3.13.16. *The potential $\Phi \in \mathcal{S}_{\mathcal{L}}$ corresponding to the six-vertex model is monotone if and only if $c_+c_- \geq \max\{a_+a_-, b_+b_-\}$, in which case σ is strictly convex on U_Φ.*

Although it is not directly stated in *Random Surfaces* [54], the potential Φ can be written as a simply attractive potential whenever $c_+c_- \geq \max\{a_+a_-, b_+b_-\}$. Therefore this theorem should be considered an alternative proof rather than a novel result.

Chapter 4

A generalisation of the honeycomb dimer model to higher dimensions

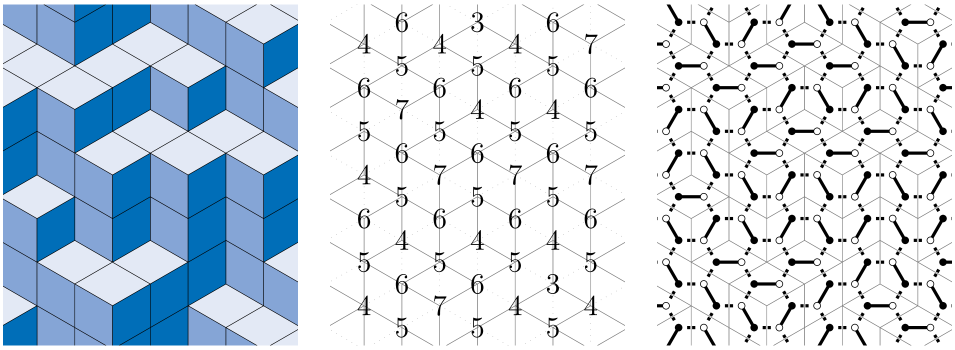
Linde, Moore, and Nordahl introduced a generalisation of the honeycomb dimer model to higher dimensions. The purpose of this chapter is to describe a number of structural properties of this generalised model. First, it is shown that the samples of the model are in one-to-one correspondence with the perfect matchings of a hypergraph. This leads to a generalised Kasteleyn theory: the partition function of the model equals the Cayley hyperdeterminant of the adjacency hypermatrix of the hypergraph. Second, we prove an identity which relates the covariance matrix of the random height function directly to the random geometrical structure of the model. This identity is known in the planar case but is new for higher dimensions. It relies on a more explicit formulation of Sheffield's *cluster swap* which is made possible by the structure of the honeycomb dimer model. Finally, we use the special properties of this explicit cluster swap to give a new and simplified proof of strict convexity of the surface tension in this case.

4.1 Introduction

4.1.1 Background

Random models on shift-invariant Euclidean graphs such as the square lattice and the hexagonal lattice form a well-known subject of study in both combinatorics and statistical physics [25]. There are several integrable models which allow for a quantitative analysis. In the integrable setting, the focus is on deriving quantitative results concerning (asymptotics of) partition functions and correlation functions. Examples of such models are the Ising model [45, 61], ice-type models [40, 57, 62], and the dimer model [28, 59], see also [2]. These quantitative estimates in turn imply qualitative results, for example (non)uniqueness of shift-invariant Gibbs measures, and (when the samples are height functions) strict convexity of the surface tension. Sometimes it is possible to derive qualitative results even in the absence of quantitative estimates. Georgii [20] provides an excellent overview for the theory of Gibbs measures for general (non-integrable) models, and Sheffield [54] derives many key results for models of height functions in the gradient setting, including strict convexity of the surface tension in any dimension.

The focus of this chapter is a natural generalisation of the hexagonal dimer model (Figure 4.1c) to arbitrary dimension. The generalised model first appeared in the



(a) (b) (c)

Figure 4.1: Dimension $d = 2$; projections onto H of several representations:

(a) As a stepped surface or lozenge tiling,

(b) As a height function on the simplicial lattice (X^d, E^d),

(c) As a dimer cover or perfect matching of the dual hexagonal lattice.

work of Linde, Moore, and Nordahl [41]. It also belongs to the category of models under consideration in the thesis of Sheffield [54]. A great deal is known about the original two-dimensional dimer model: we mention three pivotal developments. The mathematical study of dimers was initiated by Kasteleyn, Temperley, and Fisher. Kasteleyn [28] and Temperley and Fisher [59] independently calculated the number of perfect matchings of an $n \times m$ grid or, equivalently, the number of domino tilings of an $n \times m$ rectangle. Kasteleyn [29, 30] later showed that the number of perfect matchings of any bipartite planar graph equals the determinant of a matrix that is closely related to the adjacency matrix of the concerned graph. Cohn, Kenyon, and Propp [5] proved the variational principle for domino tilings; the scope of their article includes the hexagonal dimer model. Remarkably, a closed-form solution for the surface tension is found, something that is not to be expected in higher dimensions. Their derivation of the closed-form formula relies on a bijection between dimer configurations and height functions, together with an original application of the Kasteleyn theory. Kenyon, Okounkov, and Sheffield [34] establish a bijection between the set of accessible slopes and the set of ergodic Gibbs measures. They furthermore classify the ergodic Gibbs measures into three categories (frozen, liquid, and gaseous) which describe qualitatively the behaviour of the random surface. Their paper contains many more qualitative and quantitative results.

4.1.2 The honeycomb dimer model in dimension $d \geq 2$

Let (X^d, E^d) denote the graph obtained from the square lattice $\mathbb{Z}^{d+1}$ by identifying vertices which differ by an integer multiple of the vector $\mathbf{n} := \mathbf{e}_1 + \cdots + \mathbf{e}_{d+1}$. This graph is called the *simplicial lattice*; its vertices are equivalence classes of vertices of the square lattice. Let Ω denote the set of functions $f : X^d \to \mathbb{Z}$ which have the property that $f(\mathbf{0}) \in (d+1)\mathbb{Z}$ and $f([\mathbf{x}+\mathbf{e}_i]) - f([\mathbf{x}]) \in \{-d, 1\}$ for any $\mathbf{x} \in \mathbb{Z}^{d+1}$ and $1 \leq i \leq d+1$. Functions in Ω are called *height functions*—see Figure 4.1b. The set Ω

is in bijection with the set of *stepped surfaces* in $\mathbb{R}^{d+1}$. Informally, a stepped surface is a union of unit hypercubes with integer coordinates, such that each hypercube is well-supported, and such that there is no overhang. Stepped surfaces are related directly to the three-dimensional interpretation of the familiar picture of lozenge tilings for $d = 2$, see Figure 4.1a. Each stepped surface is furthermore associated with a *tiling* of the hyperplane orthogonal to $\mathbf{n}$; this tiling is essentially obtained by projecting the exposed faces of the hypercubes of the stepped surface onto this hyperplane. These bijections are all introduced in the work of Linde, Moore, and Nordahl [41].

In this chapter we are interested in the model of uniformly random height functions whenever fixed or periodic boundary conditions are enforced. For fixed boundary conditions, it is also possible to consider more general Boltzmann measures, in the spirit of the classical planar dimer model. The purpose of this chapter is to point out a number of structural properties of the generalised model which lead to new results.

4.1.3 The double dimer model and the cluster swap

For the double dimer model, one superimposes two dimer configurations of the same graph. The union of two such dimer covers decomposes into a number of isolated edges which appear in both dimer covers, called *double edges*, and a number of closed loops of even length, where each edge of the loop is contained in exactly one of the original dimer covers. If the distribution of the two dimer configurations is uniformly random, subject say to fixed boundary conditions, then the orientation of each loop is uniformly random in its two states. More precisely, this means that for each loop, one can flip a fair coin to decide if each dimer should change the configuration that it belongs to, without changing the distribution of the product measure. Many results have been obtained for the double dimer model: see the work of Kenyon and Pemantle [31, 35] for the relevant literature.

Sheffield introduces *cluster swapping* in the seminal monograph *Random surfaces* [54]. The technique employs the same idea of considering two configurations—height functions, in this case—at once, then identifying and resampling independent structures. The cluster swap applies to *simply attractive potentials*, that is, models of height functions which are induced by a convex nearest-neighbour potential. The setup is much more intricate than for the original double dimer model. To compensate for the possible change in potential, one first identifies a ferromagnetic Ising model, then achieves independence through the Edwards-Sokal representation [13] which in turn derives from the Swendsen-Wang update [58]. The cluster swap applies directly to the heights of the two height functions; it is not an operation on their gradients.

4.1.4 Main results

First, we develop a new construction, the *cluster boundary swap*, for the generalised model. This cluster boundary swap is entirely analogous to the resampling operation in the double dimer model. In particular, the difference of two uniformly random height functions decomposes as a geometrical structure consisting of boundaries—which generalise loops—and double edges, and, conditional on this structure, a number of fair coin flips, one for the orientation of each boundary. Moreover, the operation directly manipulates the gradients of the two height functions, which works to our advantage. The name *cluster boundary swap* is chosen intentionally, as it should be considered a special case of the cluster swap, adapted and optimised for the special geometrical structure that is present in the generalisation of the honeycomb dimer model to higher dimensions. The author is not aware of a similar construction for any other model in dimension $d > 2$; in particular, no generalisation is known for the dimer model on the two-dimensional square lattice.

Second, we use the cluster boundary swap to obtain an identity which relates the covariance matrix of the random height function f to the geometrical structure of the model. We prove that the variance of f at a vertex $\mathbf{x} \in X^d$ is exactly $\frac{1}{2}(d+1)^2$ times the expectation of the number of boundaries separating $\mathbf{x}$ from infinity in the product measure. The identity makes sense for fixed boundary conditions only, but it does apply to general Boltzmann measures (which includes the uniform probability measure). Similarly, for $\mathbf{x}, \mathbf{y} \in X^d$, we prove that the covariance between f at $\mathbf{x}$ and $\mathbf{y}$ equals $\frac{1}{2}(d+1)^2$ times the number of boundaries that separate both $\mathbf{x}$ and $\mathbf{y}$ from infinity. This identity is new and was previously known, to our knowledge, only for two-dimensional dimer models.

Third, we prove that the set of tilings, as introduced in [41], are in one-to-one correspondence with the perfect matchings of a hypergraph which is the natural dual of the simplicial lattice (X^d, E^d). This hypergraph is $d!$-regular and $d!$-partite, in the sense that there is a partition of its vertex set into $d!$-parts, such that each hyperedge contains exactly one vertex in each part. We derive a generalised Kasteleyn theory: we prove that the partition function of any Boltzmann measure equals the Cayley hyperdeterminant of the adjacency hypermatrix of this hypergraph. It is well-known that no Kasteleyn weighting is necessary for the Kasteleyn theory on the hexagonal lattice. We prove the same here: we may take the adjacency hypermatrix without modifications as our Kasteleyn hypermatrix.

Finally, we use the special properties of the cluster boundary swap—relative to the more general cluster swap—to greatly streamline the proof in [54] for strict convexity of the surface tension. Note that strict convexity of the surface tension for $d = 2$ was first derived in [5] by means of an explicit calculation. Strict convexity of the surface tension is important, because it implies that the model is stable on a macroscopic scale. The macroscopic behaviour of the model is described by a large devations principle, which in turn implies a variational principle—these are due to generic arguments which do not require that the surface tension is strictly convex. If the surface tension is strictly convex, however, then the rate function in the large deviations principle has a unique minimiser, and consequently the random height function concentrates around a unique limit shape. See [5, 33, 32] for the variational principle in the original dimer setting.

4.1.5 Structure of the chapter

Section 4.2 formally introduces the model, and gives an overview of the several representations of each sample. These constructions derive directly from [41]. Section 4.3 describes the probability measures under consideration, by enforcing fixed or periodic boundary conditions. We construct the cluster boundary swap in Section 4.4. In Section 4.5, we state and prove the identity for the covariance structure of the random height function, and Section 4.6 contains the generalised Kasteleyn theory. The purpose of the remainder of the chapter is to motivate and prove the result of strict convexity of the surface tension. In Section 4.7 we introduce gradient Gibbs measures, which play an important role in the proof, but which are also interesting in their own right. Section 4.8 motivates the result: it introduces the surface tension, and describes how it is related to the large deviations principle and the variational principle. Strict convexity of the surface tension is finally proven in Section 4.9.

4.2 Stepped surfaces, tilings, height functions

Linde, Moore, and Nordahl [41] observed that each sample of the model has three natural representations. The purpose of this section is to give an overview of these representations, and to state or derive some basic properties. For details, refer to [41]. Throughout this chapter, $d \geq 2$ denotes the fixed dimension that we work in.

4.2.1 Stepped surfaces

Informally, a stepped surface is a union of unit hypercubes with corners in $\mathbb{Z}^{d+1}$. The hypercubes must be properly stacked without overhang—recall for this notion the familiar picture of lozenge tilings, which corresponds to $d = 2$. If a hypercube is present at some coordinate $\mathbf{x} \in \mathbb{Z}^{d+1}$, then we require the presence of a hypercube at $\mathbf{x} - \mathbf{e}_i$ for every $1 \leq i \leq d+1$; this indeed ensures that every hypercube is well-supported, and excludes overhang.

The formal definition is as follows. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{d+1}$, then write $\mathbf{x} \leq \mathbf{y}$ whenever $\mathbf{x}_i \leq \mathbf{y}_i$ for all i. A set $A \subset \mathbb{R}^{d+1}$ is called *closed under* $\leq$ whenever $\mathbf{x} \in A$ and $\mathbf{y} \leq \mathbf{x}$ implies $\mathbf{y} \in A$. Write $L(A)$ for the closure of a set $A \subset \mathbb{R}^{d+1}$ under $\leq$, that is,

$$L(A) := \{\mathbf{y} \in \mathbb{R}^{d+1} : \mathbf{y} \leq \mathbf{x} \text{ for some } \mathbf{x} \in A\}.$$

A *stepped surface* is a strict nonempty subset of $\mathbb{R}^{d+1}$ of the form $L(A)$ for some $A \subset \mathbb{Z}^{d+1}$. Let Ψ denote the set of stepped surfaces. If F is a stepped surface, then write $V(F)$ for the set $\partial F \cap \mathbb{Z}^{d+1}$. The set $V(F)$ should be thought of as the discrete boundary of F. It is a simple exercise to work out that $F = L(V(F))$. In particular, each stepped surface is characterised by this discrete boundary.

4.2.2 The height function of a stepped surface

Consider a stepped surface F. The height function associated to F is essentially a function that has the discrete boundary $V(F)$ of F as its graph. The value of this function at each vertex $\mathbf{x} \in V(F)$ is given by $\mathbf{x}_1 + \cdots + \mathbf{x}_{d+1} = (\mathbf{x}, \mathbf{n})$, where $\mathbf{n}$ is the vector $\mathbf{e}_1 + \cdots + \mathbf{e}_{d+1}$ and $(\cdot, \cdot)$ the natural inner product. This value is also called the *height* of the vertex $\mathbf{x}$.

Write (X^d, E^d) for the graph obtained from the square lattice $\mathbb{Z}^{d+1}$ after identifying all vertices which differ by an integer multiple of $\mathbf{n}$. In particular, each vertex $[\mathbf{x}] \in X^d$ is an equivalence class of the form $[\mathbf{x}] := \mathbf{x} + \mathbb{Z}\mathbf{n}$ for some $\mathbf{x} \in \mathbb{Z}^{d+1}$. The $2d + 2$ neighbours of $[\mathbf{x}]$ are of the form $[\mathbf{x} \pm \mathbf{e}_i]$. The graph (X^d, E^d) is called the *simplicial lattice*. It is a simple exercise to derive from the definition of a stepped surface that $|[\mathbf{x}] \cap V(F)| = 1$ for any $[\mathbf{x}] \in X^d$. To derive this, observe first that $\mathbf{x} + \mathbb{R}\mathbf{n}$ intersects ∂F once because F is closed under $\leq$, then note that this point of intersection has integral coordinates whenever $\mathbf{x}$ does.

The *height function* associated to F is given by the function

$$f : X^d \to \mathbb{Z},\ [\mathbf{x}] \mapsto (\mathbf{y}, \mathbf{n}) \text{ where } \{\mathbf{y}\} = [\mathbf{x}] \cap V(F).$$

This is the desired parametrisation of $V(F)$; the set $V(F)$ is equal to

$$V(f) := \{\mathbf{x} \in \mathbb{Z}^{d+1} : (\mathbf{x}, \mathbf{n}) = f([\mathbf{x}])\}.$$

The assignment $F \mapsto f$ is injective because $F \mapsto V(F)$ is injective.

Let us now identify the image of the map $F \mapsto f$. The function f satisfies $f([\mathbf{x}]) \equiv (\mathbf{x}, \mathbf{n}) \mod d+1$; we call this the *parity condition.* Assert furthermore that $f([\mathbf{x}+\mathbf{e}_i]) \leq f([\mathbf{x}])+1$. This assertion is called the *Lipschitz condition.*

Suppose that the assertion is false, that is, that instead $f([\mathbf{x}+\mathbf{e}_i]) > f([\mathbf{x}])+1$. Write $\mathbf{y}$ for the unique vertex in $[\mathbf{x}] \cap V(F)$, and write $\mathbf{z}$ for the unique vertex in $[\mathbf{x}+\mathbf{e}_i] \cap V(F)$. Then $\mathbf{z}_j > (\mathbf{y}+\mathbf{e}_i)_j \geq \mathbf{y}_j$ for all j. In this case, it is impossible that both $\mathbf{z}$ and $\mathbf{y}$ are contained in ∂F, because F is closed under $\leq$. This proves the assertion.

The Lipschitz condition also implies that $f([\mathbf{x}+\mathbf{e}_i]) \geq f([\mathbf{x}]) - d$, since $[\mathbf{e}_1 + \cdots + \mathbf{e}_{d+1}] = [\mathbf{n}] = [\mathbf{0}]$. The Lipschitz condition, this new inequality, and the parity condition imply jointly that the gradient ∇f of f satisfies

$$\nabla f([\mathbf{x}], [\mathbf{x}+\mathbf{e}_i]) := f([\mathbf{x}+\mathbf{e}_i]) - f([\mathbf{x}]) \in \{-d, 1\}.$$

On the other hand, if the flow ∇f satisfies this equation for any $\mathbf{x}$ and i and if $f([\mathbf{0}])$ is an integer multiple of $d+1$, then f also clearly satisfies the parity condition.

A *height function* is a function $f : X^d \to \mathbb{Z}$ which satisfies $f([\mathbf{0}]) \in (d+1)\mathbb{Z}$ and

$$\nabla(f([\mathbf{x}]), f([\mathbf{x}+\mathbf{e}_i])) = f([\mathbf{x}+\mathbf{e}_i]) - f([\mathbf{x}]) \in \{1, -d\}$$

for all $\mathbf{x}$ and i; write Ω for the set of height functions. If F is a stepped surface, then the associated height function f is indeed an element of Ω. The injective map $\Psi \to \Omega$, $F \mapsto f$ is in fact a bijection; its inverse is given by the map $\Xi : \Omega \to \Psi$, $f \mapsto L(V(f))$. For details, refer to [41].

4.2.3 The tiling associated to a stepped surface

If f is a height function, then write

$$T(f) := \{\{\mathbf{x}, \mathbf{y}\} \in E^d : \nabla f(\mathbf{x}, \mathbf{y}) = -d\}.$$

The set $T(f)$ characterises ∇f, and therefore it characterises the function f up to constants. The first goal is to characterise the image of the map $f \mapsto T(f)$ over Ω.

A path $(\mathbf{s}_k)_{0 \leq k \leq n} \subset X^d$ of length $n = d+1$ is called a *rooted simplicial loop* or simply a *simplicial loop* if there exists a permutation $\xi \in S_{d+1}$ such that $\mathbf{s}_k = \mathbf{s}_{k-1} + \mathbf{e}_{\xi(k)}$ for any $1 \leq k \leq d+1$. This implies that $\mathbf{s}$ is closed because $[\mathbf{x}] + \mathbf{e}_1 + \cdots + \mathbf{e}_{d+1} = [\mathbf{x}+\mathbf{n}] = [\mathbf{x}]$. Write R^d for the set of rooted simplicial loops. Sometimes we are not concerned with the starting points of the loops. In those cases, two loops are considered equal if they are equal up to indexation—this is equivalent to requiring that the two loops traverse the same set of edges. Write U^d for the set of *unrooted simplicial loops.*

Let $\mathbf{s}$ denote a simplicial loop. Knowing that ∇f must integrate to zero along this simplicial loop, it is immediate that exactly one of the edges of $\mathbf{s}$ is contained in $T(f)$. Write Θ for the set of *tilings*, that is, the set of subsets $T \subset E^d$ with the property that $|T \cap \mathbf{s}| = 1$ for any simplicial loop $\mathbf{s}$. To see that the map $\Omega \to \Theta$, $f \mapsto T(f)$ is surjective, let $T \in \Theta$, and write α_T for the unique flow on (X^d, E^d) such that

$$\alpha_T([\mathbf{x}], [\mathbf{x}+\mathbf{e}_i]) = \begin{cases} 1 & \text{if } \{[\mathbf{x}], [\mathbf{x}+\mathbf{e}_i]\} \not\in T, \\ -d & \text{if } \{[\mathbf{x}], [\mathbf{x}+\mathbf{e}_i]\} \in T. \end{cases} \tag{4.2.1}$$

Then $\alpha_T = \nabla f$ for some height function $f \in \Omega$ if and only if α_T is conservative. The flow α_T integrates to zero along any simplicial loop, by definition of a tiling. It is a simple exercise in group theory to see that this implies that α_T integrates to zero along any closed path, since the graph (X^d, E^d) can be written as the Cayley graph on the generators $\mathbf{e}_1, \ldots, \mathbf{e}_{d+1}$ subject to the relators which are given by all possible permutations of these $d+1$ elements—these relators correspond exactly to the simplicial loops. This proves that $T = T(f)$ for some height function $f \in \Omega$, and therefore the map $\Omega \to \Theta$, $f \mapsto T(f)$ is surjective. Since $T(f)$ characterises f up to constants, this also implies that the map $\Phi : \Omega \to (d+1)\mathbb{Z} \times \Theta$, $f \mapsto (f(0), T(f))$ is a bijection.

4.2.4 The geometrical intuition behind tilings

Let us connect the combinatorial tilings $T \in \Theta$ to the familiar geometric picture of lozenge tilings. If $d = 2$, then (X^d, E^d) is the triangular lattice, and the set $T \subset E^d$ has the property that T contains exactly one edge of every triangle. Thus, if we remove the set T from this triangular lattice, then we are left with a collection of lozenges.

The boundary $\partial F \subset \mathbb{R}^{d+1}$ of a stepped surface is a union of hypercubes of codimension one with integral coordinates. Write H for the orthogonal complement of $\mathbf{n}$, and write $P : \mathbb{R}^{d+1} \to H$ for orthogonal projection onto H. For each stepped surface F, the projection map P restricts to a bijection from ∂F to H. If $d = 2$, then P maps each two-dimensional square contained in ∂F, to a lozenge embedded in H, and jointly these lozenges partition H—we ignore here the fact that the topological boundaries of the lozenges overlap. We observed in the previous paragraph that each edge in T encodes exactly one of these lozenges. The same reasoning applies to higher dimensions: the map P maps the hypercubes of codimension 1 which make up ∂F to H, and these projected hypercubes partition the d-dimensional real vector space H up to overlapping boundaries. Finally, each edge in T encodes exactly one of these projected hypercubes.

4.2.5 Lemmas for analysing stepped surfaces

The map Ξ preserves the lattice structure

Lemma 4.2.2. *Let f_1 and f_2 be height functions and let $F_1 = \Xi(f_1)$ and $F_2 = \Xi(f_2)$. Then*

1. *$f_1 \leq f_2$ if and only if $F_1 \subset F_2$,*
2. *$f_1 \vee f_2$ is a height function, $F_1 \cup F_2$ is a stepped surface, and $\Xi(f_1 \vee f_2) = F_1 \cup F_2$,*
3. *$f_1 \wedge f_2$ is a height function, $F_1 \cap F_2$ is a stepped surface, and $\Xi(f_1 \wedge f_2) = F_1 \cap F_2$.*

Of course, 2 and 3 extend to finite unions, intersections, maxima, and minima.

See Lemma 2 in [41] for a proof.

Local moves

If two height functions agree at all but one vertex, then it is said that they differ by a *local move*. A local move is equivalent to adding or removing a single unit hypercube to the corresponding stepped surface.

Lemma 4.2.3. *Suppose that $R \subset X^d$ is finite and that f and g are height functions that are equal outside R and satisfy $f \leq g$ on R. Then there exists a sequence of height functions $(f_k)_{0 \leq k \leq n} \subset \Omega$ with*

1. *$f_0 = f$ and $f_n = g$,*

2. *for every $0 \leq k < n$, there is a unique $\mathbf{x} \in R$ such that $f_{k+1} = f_k + (d+1) \cdot 1_{\mathbf{x}}$.*

The sequence is increasing and all functions f_k agree to f and g on $X^d \smallsetminus R$.

See Lemma 3 in [41] for a proof. If f and g agree outside R but neither $f \leq g$ nor $f \geq g$, then one can first go down from f to $f \wedge g$ and then up from $f \wedge g$ to g. Lemma 4.2.2 ensures that $f \wedge g$ is a height function.

The Kirszbraun theorem

Consider a vertex $[\mathbf{x}] \in X^d$. Since all elements in $[\mathbf{x}]$ differ from one another by multiples of $\mathbf{n}$, and since $\mathbf{n} \in \operatorname{Ker} P$, there is a unique element $\mathbf{y} \in H$ such that $\{\mathbf{y}\} = P[\mathbf{x}]$. Let us identify each vertex $[\mathbf{x}] \in X^d$ with this corresponding point $\mathbf{y} \in H$. Write $\mathbf{g}_i := P\mathbf{e}_i = \mathbf{e}_i - \mathbf{n}/(d+1)$. The neighbours of $\mathbf{x} \in X^d \subset H$ are then given by $\mathbf{x} \pm \mathbf{g}_i$.

Define the asymmetric norm $\|\cdot\|_+$ on H by $\|\mathbf{x}\|_+ := -(d+1)\min_i \mathbf{x}_i$. Remark that $\|\cdot\|_+$ is the largest asymmetric norm on H subject to $\|\mathbf{g}_i\|_+ \leq 1$ for all i. In other words, a function $f : X^d \to \mathbb{R}$ satisfies the Lipschitz condition introduced in Subsection 4.2.2 if and only if

$$f(\mathbf{y}) - f(\mathbf{x}) \leq \|\mathbf{y} - \mathbf{x}\|_+ \tag{4.2.4}$$

for all $\mathbf{x}, \mathbf{y} \in X^d$. A function $f : D \to \mathbb{R}$ defined on a subset $D \subset H$ is called *Lipschitz* whenever f satisfies (4.2.4) for all $\mathbf{x}, \mathbf{y} \in D$. This definition is consistent with the previous definition of the Lipschitz property for height functions.

Lemma 4.2.5. *If a function $f : H \to \mathbb{R}$ is Lipschitz then there exists a unique largest height function g subject to $g \leq f|_{X^d}$; the value of g at each $\mathbf{x} \in X^d$ is given by $g(\mathbf{x}) := k$, where k is the largest integer which makes g satisfy the parity condition at $\mathbf{x}$, and which does not exceed $f(\mathbf{x})$.*

If $f : H \to \mathbb{R}$ is a Lipschitz function, then write $\lfloor f \rfloor$ for the largest height function subject to $\lfloor f \rfloor \leq f|_{X^d}$; its function values are given by the previous lemma. We leave its proof to the reader. The previous lemma results in a discrete analogue of the Kirszbraun theorem for the current setting.

Lemma 4.2.6. *If $R \subset H$ and $f : R \to \mathbb{R}$ is Lipschitz, then f extends to a Lipschitz function $\bar{f} : H \to \mathbb{R}$. If $R \subset X^d$ and $f : R \to \mathbb{Z}$ is Lipschitz and satisfies the parity condition for every $\mathbf{x} \in R$, then f extends to a height function $\bar{f} \in \Omega$.*

The first assertion is the original Kirszbraun theorem. For the second assertion, let $R \subset X^d$ and $f : R \to \mathbb{Z}$ be as in the lemma. The function f extends to some Lipschitz function $g : H \to \mathbb{R}$ by the Kirszbraun theorem. The previous lemma states that $\lfloor g \rfloor$ is a height function that equals f on R. This proves the second assertion.

4.3 Random height functions

In this section, we introduce boundary conditions (fixed or periodic) in order to reduce Ω to a finite set, and define and study probability measures on these finite sets. We assert that the model is monotone in boundary conditions, and state an immediate corollary of this fact by employing the Azuma-Hoeffding inequality.

4.3.1 Fixed boundary conditions

We reduce Ω to a finite set by applying fixed boundary conditions. One can study the uniform probability measure on this finite set. One can also define more general Boltzmann measures. If $R \subset X^d$, then write R^{c} for $X^d \smallsetminus R$. Write $E^d(R)$ for the set of edges in E^d that are incident to at least one vertex in R.

Definition 4.3.1. Define, for any height function f and for any tiling T,

$$\begin{aligned}\Omega(R, f) &:= \{g \in \Omega : g|_{R^{\mathsf{c}}} = f|_{R^{\mathsf{c}}}\},\\ \Theta(R, T) &:= \{Y \in \Theta : Y \smallsetminus E^d(R) = T \smallsetminus E^d(R)\}.\end{aligned}$$

Call a set $R \subset X^d$ a *region* if R is finite and if R^{c} is connected.

Lemma 4.3.2. *Let $R \subset X^d$ be a finite set, f a height function, and $T := T(f)$. Then*

1. *$\Omega(R, f)$ and $\Theta(R, T)$ are finite sets,*
2. *The map $g \mapsto T(g)$ restricts to an injection from $\Omega(R, f)$ to $\Theta(R, T)$,*
3. *If R is a region, then the restricted map in the previous statement is a bijection.*

Proof. Without loss of generality, R does not contain $\mathbf{0}$. If $g \in \Omega(R, f)$ then we must have $g(\mathbf{0}) = f(\mathbf{0})$ and $T(g) \smallsetminus E^d(R) = T(f) \smallsetminus E^d(R)$. The map $g \mapsto T(g)$ restricts to an injection because the gradient ∇g can be reconstructed from $T(g)$, which is sufficient for reconstructing g as the constant is determined from $g(\mathbf{0}) = f(\mathbf{0})$. The logarithm of $|\Theta(R, T)|$ is bounded by $|E^d(R)| \log 2 < \infty$. We have now proven the first two assertions of the lemma.

Next, we prove that the same restriction map is also surjective whenever R is a region. Fix $Y \in \Theta(R, T)$ and define $g := \Phi^{-1}(f(\mathbf{0}), Y)$. It suffices to show that $g \in \Omega(R, f)$, that is, that $g(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x} \in R^{\mathsf{c}}$. Let $\mathbf{p}$ denote a path from $\mathbf{0}$ to $\mathbf{x}$ through $(X^d \smallsetminus R, E^d \smallsetminus E^d(R))$; such a path exists by definition of a region. Then $g(\mathbf{x}) - f(\mathbf{0})$ and $f(\mathbf{x}) - f(\mathbf{0})$ can both be calculated in terms of integrals of ∇g and ∇f respectively over the path $\mathbf{p}$. Since $Y \smallsetminus E^d(R) = T \smallsetminus E^d(R)$, these gradients are equal on the edges of $\mathbf{p}$, which proves that $g(\mathbf{x}) = f(\mathbf{x})$. $\square$

Fix a finite set $R \subset X^d$ and a height function $f \in \Omega$. Write $f^{\pm}$ for the pointwise minimum and maximum over all height functions in the finite set $\Omega(R, f)$. These are also height function by virtue of Lemma 4.2.2, and clearly $\Omega(R, f) = \{g \in \Omega : f^- \leq g \leq f^+\}$. The same lemma implies the following result.

Lemma 4.3.3. *Fix a finite set $R \subset X^d$ and a height function f. Write $f^{\pm}$ as in the preceding discussion and define $F^{\pm} := \Xi(f^{\pm})$. Then Ξ restricts to a bijection from $\Omega(R, f)$ to $\{F \in \Psi : F^- \subset F \subset F^+\}$.*

Next, we define Boltzmann measures. The uniform probability measures on $\Omega(R, f)$ and $\Theta(R, T)$ are examples of Boltzmann measures. The introduction of Boltzmann measures allows us to increase the relative probability of tilings containing certain edges.

Definition 4.3.4. Let R be a region, f a height function, and $T := T(f)$ a tiling. A positive real function w on $E^d(R)$ is called a *weight function.* Let $\mathbb{P}_w$ be the probability measure on the set $\Theta(R, T)$ such that $\mathbb{P}_w(Y) \propto \prod_{e \in Y \cap E^d(R)} w(e)$ for any $Y \in \Theta(R, T)$, that is,

$$\mathbb{P}_w(Y) := \frac{1}{Z_w} \prod_{e \in Y \cap E^d(R)} w(e) \qquad \text{where} \qquad Z_w := \sum_{Y \in \Theta(R,T)} \prod_{e \in Y \cap E^d(R)} w(e).$$

The probability measure $\mathbb{P}_w$ is called a *Boltzmann measure* and the normalising constant Z_w is called the *partition function.* The measure $\mathbb{P}_w$ is also considered a probability measure on the sample space $\Omega(R, f)$ by defining $\mathbb{P}_w(g) := \mathbb{P}_w(T(g))$. Write $\mathbb{P}$ for $\mathbb{P}_w$ with w identically equal to 1, and write Z for the corresponding partition function. Observe that $Z = |\Omega(R, f)| = |\Theta(R, T)|$. The definition of Z_w makes sense also when w takes complex values.

We prove that Z_w equals the Cayley hyperdeterminant of a suitable hypermatrix in Section 4.6.

4.3.2 The periodic setting

Write H^* for the natural dual space of the d-dimensional real vector space H. A linear form $s \in H^*$ is called a *slope.*

Definition 4.3.5. Periodic boundary conditions are characterised by a pair (L, s), where $L \subset X^d$ is a full-rank sublattice and $s \in H^*$ a slope. A function $f : X^d \to \mathbb{R}$ is called (L, s)-*periodic* if, for any $\mathbf{x} \in X^d$ and $\mathbf{y} \in L$,

$$f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + s(\mathbf{y}).$$

Write $\Omega(L, s)$ for the set of (L, s)-periodic height functions that map $\mathbf{0}$ to 0. Call a pair (L, s) *valid* if $\Omega(L, s)$ is nonempty.

It is not a priori clear which periodic boundary conditions (L, s) are valid. If f is a height function, then write $f|_L$ and $s|_L$ for the restrictions of f and s to $L \subset X^d \subset H$ respectively. First, every function $f \in \Omega(L, s)$ must satisfy $f|_L = s|_L$, so if $s|_L$ does not extend to a height function then $\Omega(L, s)$ is empty. On the other hand, if $s|_L$ extends to a height function, then the minimum amongst all possible extensions is (L, s)-periodic. Therefore (L, s) is valid if and only if $s|_L$ extends to a height function. Lemma 4.2.6 imposes a Lipschitz condition and a parity condition on $s|_L$. Clearly $s|_L$ is Lipschitz if and only if s is Lipschitz.

Definition 4.3.6. Write $\mathcal{S}$ for the set of slopes in H^* which are Lipschitz, that is,

$$\mathcal{S} := \{s \in H^* : \max_i s(\mathbf{g}_i) \leq 1\}.$$

Proposition 4.3.7. *The set $\mathcal{S} \subset H^*$ is a closed d-simplex with extreme points $(s^i)_{1 \leq i \leq d+1}$, where each slope s^i satisfies, for any j,*

$$s^i(\mathbf{g}_j) := \begin{cases} -d & \text{if } i = j, \\ 1 & \text{if } i \neq j. \end{cases}$$

Introduce also the parity condition for the following result.

Lemma 4.3.8. *Suppose given periodic boundary conditions (L, s). Then (L, s) is valid if and only if $s \in \mathcal{S}$ and $s([\mathbf{x}]) \equiv (\mathbf{x}, \mathbf{n}) \mod d+1$ for every $[\mathbf{x}] \in L \subset X^d$. In particular, if $L \subset (d+1)X^d$, then $(\mathbf{x}, \mathbf{n}) \in (d+1)\mathbb{Z}$ for every $[\mathbf{x}] \in L$, and under this extra condition, (L, s) is valid if and only if $s \in \mathcal{S}$ and $s(\mathbf{x}) \in (d+1)\mathbb{Z}$ for every $\mathbf{x} \in L$.*

If $L = n(d+1)X^d$ for some $n \in \mathbb{N}$, then (L, s) is valid if and only if $s \in \mathcal{S}$ and

$$s(n(d+1)\mathbf{g}_i) = n(d+1)s(\mathbf{g}_i) \in (d+1)\mathbb{Z}$$

for every $1 \leq i \leq d+1$, that is, $s(\mathbf{g}_i) \in \mathbb{Z}/n$.

Definition 4.3.9. Write $L_n := n(d+1)X^d$ for every $n \in \mathbb{N}$. Define

$$\mathcal{S}_n := \{s \in \mathcal{S} : s(\mathbf{g}_i) \in \mathbb{Z}/n \text{ for every } 1 \leq i \leq d+1\}.$$

In other words, $\mathcal{S}_n$ is precisely the set of slopes s such that (L_n, s) is valid.

Remark that $\mathcal{S}_n$ converges to $\mathcal{S}$ in the Hausdorff metric, in the sense that every slope $s \in \mathcal{S}$ can be approximated by a sequence of slopes $(s_n)_{n\in\mathbb{N}}$ where $s_n \in \mathcal{S}_n$ for each n.

4.3.3 Symmetries of the periodic setting

For $\mathbf{x} \in X^d$, write $\theta_{\mathbf{x}}$ for the map $H \to H$, $\mathbf{y} \mapsto \mathbf{y} + \mathbf{x}$. This map is called a *shift*, and it is also clearly a symmetry of X^d. Write $\Theta(L)$ for the group $\{\theta_{\mathbf{x}} : \mathbf{x} \in L\}$ for any sublattice L of X^d, and write $\Theta := \Theta(X^d)$.

If f is a function defined on either H or X^d, then write θf for the function defined by $\theta f(\mathbf{x}) := f(\theta\mathbf{x})$. If f is a height function and $\theta \in \Theta$, then define the height function $\tilde{\theta} f$ by $\tilde{\theta} f := \theta f - f(\theta\mathbf{0}) + f(\mathbf{0})$. In other words, $\tilde{\theta} f$ is the unique height function such that $(\tilde{\theta} f)(\mathbf{0}) = f(\mathbf{0})$ and $\theta T(\theta f) = T(f)$. The map $\tilde{\theta} : \Omega \to \Omega$ is bijective, and $\tilde{\theta}$ restricts to a bijection from $\Omega(L, s)$ to $\Omega(L, s)$ for any periodic boundary conditions (L, s).

Each height function $f \in \Omega(L, s)$ is characterised by the $\Theta(L)$-invariant set $T(f) \subset E^d$. This implies in particular that the set $\Omega(L, s)$ is finite, because we have $|\Omega(L, s)| \leq 2^{|E^d/\Theta(L)|} < \infty$.

Lemma 4.3.10. *Pick valid periodic boundary conditions (L, s), write $\mathbb{P}$ for the uniform probability measure on $\Omega(L, s)$, and write f for the random function in $\Omega(L, s)$. Then $\tilde{\theta} f \sim f$ in $\mathbb{P}$ for any $\theta \in \Theta$. Moreover, $\mathbb{E} f(\mathbf{x}) = s(\mathbf{x})$ for every $\mathbf{x} \in X^d$.*

Proof. The first assertion is obvious as $\tilde{\theta} : \Omega(L, s) \to \Omega(L, s)$ is a bijection and $\mathbb{P}$ is uniform on this set. The first assertion implies that the map $\mathbb{E}(f(\cdot)) : X^d \to \mathbb{R}$ is additive. Therefore it must extend to a linear form in H^*. Now L is full-rank and $\mathbb{P}(f(\mathbf{x}) = s(\mathbf{x})) = 1$ for every $\mathbf{x} \in L$, and therefore $\mathbb{E}(f(\cdot))$ must extend to the linear form $s \in \mathcal{S} \subset H^*$. $\square$

4.3.4 Monotonicity of the random function

Monotonicity in boundary conditions is often an essential property for understanding the macroscopic behaviour of the system. See Lemmas 3 and 4 in [41] for a proof of the following theorem—note that these lemmas apply both to fixed boundary conditions in the general Boltzmann setting, as well as to periodic boundary conditions. The proof simply says that the Glauber dynamics preserve the monotonicity and mix to the correct distribution.

Theorem 4.3.11 (Monotonicity). *Let R be a region, let $b_1, b_2 \in \Omega$, and fix a weight function w. Write $\mathbb{P}_1$ and $\mathbb{P}_2$ for the Boltzmann measures with weight w on $\Omega(R, b_1)$ and $\Omega(R, b_2)$ respectively. Write a_- and a_+ for the infimum and supremum of $(b_1 - b_2)|_{R^c}$ respectively. Then there exists a probability distribution $\mathbb{P}$ on the pair $(f_1, f_2) \in \Omega(R, b_1) \times \Omega(R, b_2)$ with marginals $\mathbb{P}_1$ and $\mathbb{P}_2$ such that $a_- \leq f_1 - f_2 \leq a_+$ almost surely.*

Now let (L, s) denote a valid pair of periodic boundary conditions, let $b_1, b_2 \in \Omega(L, s)$, and fix $R \subset X^d$. Write $\mathbb{P}_1$ and $\mathbb{P}_2$ for the uniform probability measure on $\Omega(L, s)$ conditioned on $f|_R = b_1|_R$ and $f|_R = b_2|_R$ respectively. Write a_- and a_+ for the infimum and supremum of $(b_1 - b_2)|_R$ respectively. Then there exists a probability distribution $\mathbb{P}$ on the pair $(f_1, f_2) \in \Omega(L, s)^2$ with marginals $\mathbb{P}_1$ and $\mathbb{P}_2$ such that $a_- \leq f_1 - f_2 \leq a_+$ almost surely.

4.3.5 Pointwise concentration inequalities

Theorem 4.3.12 (Azuma-Hoeffding inequality). *Let $R \subset X^d$ denote a finite set, let $b \in \Omega$, and fix a weight function w. Write $\mathbb{P}$ for the Boltzmann measure with weight w on $\Omega(R, b)$. Fix $\mathbf{x} \in X^d$. Then the following inequalities hold true:*

1. $\operatorname{Var} f(\mathbf{x}) \leq (d+1)^2 n$,
2. $\mathbb{P}(f(\mathbf{x}) - \mu \geq (d+1)a) \leq \exp -\frac{a^2}{2n}$ *for all $a \geq 0$ whenever $n > 0$,*
3. $\mathbb{P}(f(\mathbf{x}) - \mu \leq (d+1)a) \leq \exp -\frac{a^2}{2n}$ *for all $a \leq 0$ whenever $n > 0$,*

where $n = d_{(X^d, E^d)}(\mathbf{x}, R^c)$ and $\mu = \mathbb{E} f(\mathbf{x})$.

Now let (L, s) denote a valid pair of periodic boundary conditions, and write $\mathbb{P}$ for the uniform probability measure on $\Omega(L, s)$. Then (1)–(3) remain true for the choices $n = d_{(X^d, E^d)}(\mathbf{x}, L) \leq d_{(X^d, E^d)}(\mathbf{x}, \mathbf{0})$ and $\mu = \mathbb{E} f(\mathbf{x}) = s(\mathbf{x})$.

Proof. This is a consequence of monotonicity (Theorem 4.3.11) and the Azuma-Hoeffding inequality. Focus on fixed boundary conditions; the proof for periodic boundary conditions is entirely analogous. Let $\mathbf{p} = (\mathbf{p}_k)_{0 \leq k \leq n}$ denote a path of shortest length from R^c to $\mathbf{x}$. Write $(f_k)_{0 \leq k \leq n}$ for the martingale such that f_k equals the expectation of $f(\mathbf{x})$ conditional on the values of f on the vertices in $\{\mathbf{p}_0, \ldots, \mathbf{p}_k\}$. By Theorem 4.3.11, we have $|f_k - f_{k+1}| \leq d + 1$. The theorem now follows from the Azuma-Hoeffding inequality. □

4.4 The cluster boundary swap

In the seminal work [54], Sheffield introduces *cluster swapping*. This technique is related to the double dimer model, where for uniform probability measures, the orientation of each loop is uniformly random in the two states, independently of any

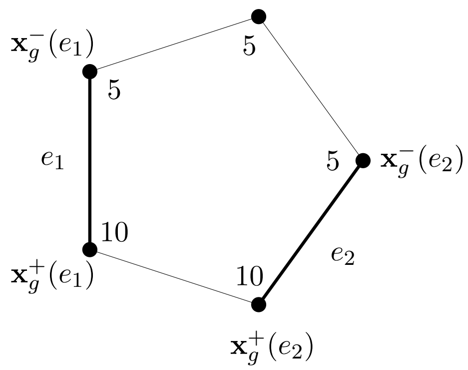


Figure 4.2: The values of $g = f_1 - f_2$ along a simplicial loop; $d = 4$

other structure that is present. In this section, we introduce the *cluster boundary swap*. The cluster boundary swap is a special case of the cluster swap, adapted and optimised for the special nature of the model under consideration. Its properties are reminiscent of the double dimer model, because, conditional on the geometrical structure involving boundaries and double edges, the orientation of each boundary is uniformly random in its two states.

4.4.1 The boundary graph and the level set decomposition

In Subsection 4.2.3 it was proven that every height function f is characterised by the pair $\Phi(f) = (f(\mathbf{0}), T(f))$. In this section, f_1 and f_2 denote height functions, and we write $(a_i, T_i) := \Phi(f_i)$ for $i \in \{1, 2\}$. The difference function $f_1 - f_2$ is denoted by g. The goal of this subsection is to understand the level set structure of g.

Lemma 4.4.1. *Let $\mathbf{s} \in R^d$ be a rooted simplicial loop. As one walks along $\mathbf{s}$,*

1. *The function f_1 moves up by 1 exactly d times,*
2. *The function f_1 moves down by d exactly once,*
3. *Either g remains constant, or it changes value twice,*
4. *If g is not constant, then the difference between its two values is $d + 1$.*

The lemma follows immediately from the observations in Subsection 4.2.3.

Write $A \ominus B$ for the symmetric difference of arbitrary sets A and B.

Definition 4.4.2. Define the graph $G_g = (V_g, E_g)$ as follows. Its vertex set V_g is given by

$$V_g := T_1 \ominus T_2 = \{e \in E^d : g \text{ is not constant on } e\} \subset E^d,$$

and two vertices $e_1, e_2 \in V_g \subset E^d$ are neighbours if some simplicial loop travels through both e_1 and e_2. The graph G_g is called the *boundary graph*. For $e \in V_g$, write $\mathbf{x}_g^-(e), \mathbf{x}_g^+(e) \in X^d$ for the vertices contained in $e \subset X^d$ on which g takes the smaller value and the larger value respectively—see Figure 4.2. For any $C \subset V_g$, we write $\mathbf{x}_g^\pm(C) := \{\mathbf{x}_g^\pm(e) : e \in C\}$.

For example, if $d = 2$, then one can identify each edge of the triangular lattice with the lozenge obtained by removing that edge. The set V_g is then precisely the set of edges of lozenges which appear in exactly one of the two configurations. Two edges in V_g are neighbours if they belong to the same triangle of the triangular lattice. Each connected component of G_g corresponds to the set of edges of the triangular lattice crossed by a nontrivial loop of the double dimer model.

Lemma 4.4.3. *Let $C \subset V_g$ be a connected component of G_g. Then $(X^d, E^d \smallsetminus C)$ consists of two connected components, one containing $\mathbf{x}_g^-(C)$, and the other containing $\mathbf{x}_g^+(C)$. Moreover, each of $\mathbf{x}_g^-(C)$ and $\mathbf{x}_g^+(C)$ is contained in a connected component of the graph $(X^d, E^d \smallsetminus V_g)$.*

Proof. Suppose that the G_g-vertices e_1 and e_2 are neighbours in the graph G_g; write $\mathbf{s}$ for a simplicial loop passing through both e_1 and e_2. Then $\mathbf{s}$ contains no other edges in V_g by Proposition 4.4.1, 3, and therefore $\mathbf{x}_g^-(e_1)$ and $\mathbf{x}_g^-(e_2)$ are connected in the graph $(X^d, E^d \smallsetminus V_g)$; see also Figure 4.2. Induct on this argument to see that $\mathbf{x}_g^-(C)$ is contained in a connected component of $(X^d, E^d \smallsetminus V_g)$. Identical reasoning applies to the set $\mathbf{x}_g^+(C)$, and we also learn that each of $\mathbf{x}_g^\pm(C)$ is contained in a connected component of $(X^d, E^d \smallsetminus C)$.

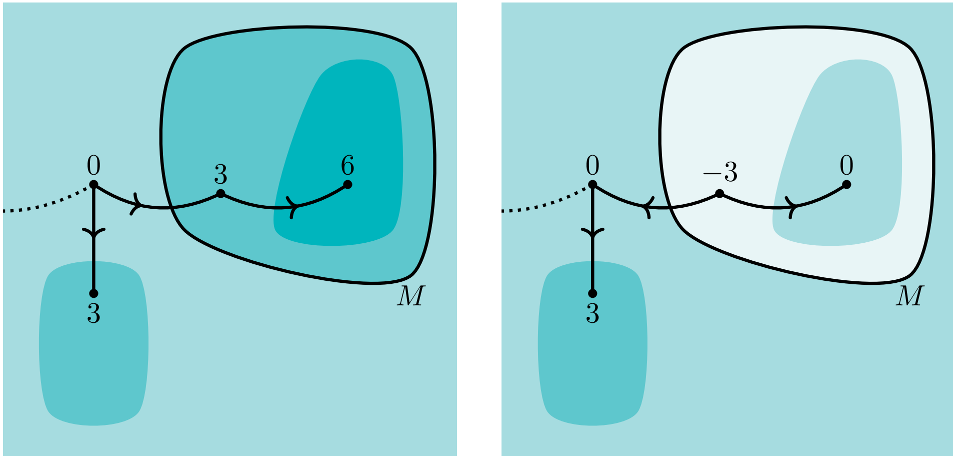
The sets $\mathbf{x}_g^\pm(C)$ cover all the endpoints of edges in C, and therefore two possibilities remain: either the graph $(X^d, E^d \smallsetminus C)$ is connected, or it consists of two connected components, with one containing $\mathbf{x}_g^-(C)$, and the other containing $\mathbf{x}_g^+(C)$. To establish the lemma we must exclude the first possibility. Every simplicial loop intersects C an even number of times. The group theory arguments that proved that the flow in Subsection 4.2.3 was conservative, imply here that any closed walk through (X^d, E^d) intersects C an even number of times. This proves that $(X^d, E^d \smallsetminus C)$ is not connected. □

Definition 4.4.4. A *g-level set* is a connected component of the graph $(X^d, E^d \smallsetminus V_g)$. A *$g$-boundary* is a connected component of the graph $G_g = (V_g, E_g)$. The g-level sets are considered subsets of X^d, and the g-boundaries are considered subsets of $V_g \subset E^d$. If E is a g-boundary, then write $X_g^\pm(E)$ for the g-level set containing $\mathbf{x}_g^\pm(E)$. The *level set decomposition of g* or $\mathrm{LSD}(g)$ is an undirected graph, where the vertices are the g-level sets and the edges are the g-boundaries. The g-boundary E connects the g-level sets $X_g^-(E)$ and $X_g^+(E)$. Write g for the graph homomorphism $g : \mathrm{LSD}(g) \to (d+1)\mathbb{Z}$ that assigns the value $g(X)$ to a g-level set X. The vector field ∇g directs the edges in $\mathrm{LSD}(g)$: it orients each g-boundary E from $X_g^-(E)$ to $X_g^+(E)$. Write $(\mathrm{LSD}(g), \nabla g)$ for this directed graph.

If $d = 2$, then the g-boundaries correspond exactly to the loops of the double dimer model. The g-level sets correspond to the connected components of $\mathbb{R}^2$ with all loops of the double dimer model removed. In Figure 4.3 we see an example of this graph. Each g-level set contracts into a single $\mathrm{LSD}(g)$-vertex. The $\mathrm{LSD}(g)$-edges are comprised of the g-boundaries separating the g-level sets.

Lemma 4.4.5. *The graph $\mathrm{LSD}(g)$ is well-defined and a tree.*

Proof. It follows from Lemma 4.4.3 that every ∇g-directed $\mathrm{LSD}(g)$-edge has a well-defined startpoint and endpoint, and that removing an edge disconnects the graph. □



g and $(\mathrm{LSD}(g), \nabla g)$ $\qquad$ g' and $(\mathrm{LSD}(g'), \nabla g')$

Figure 4.3: The level set decomposition and a cluster boundary swap by M; $d = 2$.

4.4.2 The cluster boundary swap

Lemma 4.4.6. *Let $M \subset V_g \subset E^d$ be a union of g-boundaries. Then the sets $T_i' := T_i \ominus M$ are tilings for $i \in \{1,2\}$. Write f_1' and f_2' for the unique height functions such that $\Phi(f_i') = (a_i, T_i')$ and define $g' = f_1' - f_2'$. Then $G_{g'} = G_g$ and* $\mathrm{LSD}(g') = \mathrm{LSD}(g)$*. Moreover, $\nabla g'$ and ∇g are the same except that the g-boundaries contained in M have reversed orientation, that is, $\nabla g' = (-1)^{1_M} \cdot \nabla g$.*

Proof. First claim that T_1' and T_2' are tilings. We focus on T_1'. Let $\mathbf{s}$ be a simplicial loop, and abuse notation by writing $\mathbf{s}$ also for the set of edges crossed by this loop. It suffices to prove that $|T_1' \cap \mathbf{s}| = 1$. Now either $\mathbf{s} \cap M$ is empty, or contains two edges, one from T_1 and one from T_2. In the former case we have $T_1' \cap \mathbf{s} = T_1 \cap \mathbf{s}$ and consequently $|T_1' \cap \mathbf{s}| = 1$. In the latter case, we have $T_1' \cap \mathbf{s} = T_2 \cap \mathbf{s}$ and consequently $|T_1' \cap \mathbf{s}| = 1$, as desired. This proves the claim. The appropriate functions f_1' and f_2' exist because Φ is a bijection from Ω to $(d+1)\mathbb{Z} \times \Theta$. Next,

$$V_{g'} := T_1' \ominus T_2' = (T_1 \ominus M) \ominus (T_2 \ominus M) = T_1 \ominus T_2 = V_g,$$

and consequently $G_{g'} = G_g$ and $\mathrm{LSD}(g') = \mathrm{LSD}(g)$. Recall the definition of α_T in terms of T in (4.2.1). We have

$$\nabla g' = \alpha_{T_1'} - \alpha_{T_2'} = \alpha_{T_1 \ominus M} - \alpha_{T_2 \ominus M} = (-1)^{1_M} \cdot (\alpha_{T_1} - \alpha_{T_2}) = (-1)^{1_M} \cdot \nabla g;$$

this follows directly from the fact that $M \subset T_1 \ominus T_2$ and from the definition of α_T. $\square$

Definition 4.4.7. Define

$$\begin{aligned}(T_1, T_2) \ominus M &:= (T_1', T_2') = (T_1 \ominus M, T_2 \ominus M),\\ (f_1, f_2) \ominus M &:= (f_1', f_2') = \left(\Phi^{-1}(a_1, T_1 \ominus M), \Phi^{-1}(a_2, T_2 \ominus M)\right),\end{aligned}$$

whenever these are related as in the previous lemma. Write $(f_1, f_2) \sim (f_1', f_2')$ whenever $(f_1', f_2') = (f_1, f_2) \ominus M$ for some union of g-level sets M, in which case we say that the two pairs *differ by a cluster boundary swap*. The relation $\sim$ is an equivalence relation; write $[(f_1, f_2)]$ for the equivalence class of (f_1, f_2).

Remark 4.4.8. 1. If $(f_1', f_2') = (f_1, f_2) \ominus M$, then $f_1' + f_2' = f_1 + f_2$; a cluster boundary swap does not change the sum of the two involved height functions. To see that this is the case, observe that M is a subset of $T_1 \ominus T_2$, and therefore $1_{T_1 \ominus M} + 1_{T_2 \ominus M} = 1_{T_1} + 1_{T_2}$ and

$$\nabla f_1 + \nabla f_2 = \alpha_{T_1} + \alpha_{T_2} = \alpha_{T_1 \ominus M} + \alpha_{T_2 \ominus M} = \nabla f_1' + \nabla f_2'.$$

2. The cluster boundary swap was formalised in terms of the height functions f_1 and f_2. The cluster boundary swap should however be understood as an operation on the gradients ∇f_1 and ∇f_2 of these height functions. This gradient operation is made into an operation on the non-gradient height functions by choosing the vertex $\mathbf{0} \in X^d$ as a reference vertex at which the height is held constant.

Figure 4.3 illustrates a cluster boundary swap. The thick contour is the set M, and the two difference functions g and g' are related by $g = f_1 - f_2$ and $g' = f_1' - f_2'$ where $(f_1', f_2') := (f_1, f_2) \ominus M$. Swapping by M effectively inverts the orientation of the corresponding g-boundary. One can swap any union of g-boundaries. Therefore one can direct the edges of $\mathrm{LSD}(g)$ in any desired way. We obtain the following theorem.

Theorem 4.4.9. *The relation $\sim$ is an equivalence relation on Ω^2. The elements in the equivalence class of (f_1, f_2) correspond naturally to the graph homomorphisms from the tree $\mathrm{LSD}(g)$ to $(d+1)\mathbb{Z}$ that map $\mathbf{0}$ to $g(\mathbf{0})$.*

4.5 The variance and covariance structure

This section is dedicated to a straightforward application of Theorem 4.4.9 in the fixed boundary setting.

Theorem 4.5.1. *Let R be a region not containing $\mathbf{0}$, let b be a height function, and let $w : E^d(R) \to (0, \infty)$ be a weight function. Denote the Boltzmann measure on $\Omega(R, b)$ with weight w by $\mathbb{P}_w$. Abuse notation by writing $\mathbb{P}_w$ for $\mathbb{P}_w \times \mathbb{P}_w$; write (f_1, f_2) for the pair of random functions in this measure, and write $g := f_1 - f_2$. Also write f for f_1. Then for any $\mathbf{x} \in X^d$, we have*

$$\mathrm{Var}_w\, f(\mathbf{x}) = \frac{1}{2}(d+1)^2 \mathbb{E}_w d_{\mathrm{LSD}(g)}(\mathbf{0}, \mathbf{x}).$$

In other words, the variance of $f(\mathbf{x})$ in $\mathbb{P}_w$ equals $\frac{1}{2}(d+1)^2$ times the $\mathbb{P}_w$-expectation of the number of g-boundaries that separate $\mathbf{x}$ from $\mathbf{0}$.

Proof. The random variables $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are i.i.d., and therefore

$$\mathrm{Var}_w\, f(\mathbf{x}) = \frac{1}{2} \mathrm{Var}_w(f_1(\mathbf{x}) - f_2(\mathbf{x})) = \frac{1}{2}\mathbb{E}_w(f_1(\mathbf{x}) - f_2(\mathbf{x}))^2 = \frac{1}{2}\mathbb{E}_w g(\mathbf{x})^2.$$

It suffices to prove that $\mathbb{E}_w g(\mathbf{x})^2 = (d+1)^2 \mathbb{E}_w d_{\mathrm{LSD}(g)}(\mathbf{0}, \mathbf{x})$. In fact, we make the stronger claim that

$$\mathbb{E}_w(g(\mathbf{x})^2 | [(f_1, f_2)]) = (d+1)^2 d_{\mathrm{LSD}(g)}(\mathbf{0}, \mathbf{x}).$$

The left hand side is $\sigma([(f_1, f_2)])$-measurable by definition. For the right hand side, observe that the graph $\mathrm{LSD}(g)$ is constant on each equivalence class $[(f_1, f_2)]$, which means that $d_{\mathrm{LSD}(g)}(\mathbf{0},\mathbf{x})$ is also $\sigma([(f_1,f_2)])$-measurable. The proof of the claim relies on Theorem 4.4.9.

Assert first that $[(f_1,f_2)] \subset \Omega(R,b)^2$ whenever $(f_1,f_2) \in \Omega(R,b)^2$. The set R^{c} is connected by the definition of a region, and it contains $\mathbf{0}$. Therefore R^{c} is contained in the g-level set containing $\mathbf{0}$. A cluster boundary swap does not alter the values of f_1 and f_2 on this g-level set, and therefore f_1, f_2, f_1', f_2', and b all assume the same values on R^{c} provided that $(f_1',f_2') \sim (f_1,f_2)$ and $(f_1,f_2) \in \Omega(R,b)^2$. This proves the assertion. Next, assert that $\mathbb{P}_w$ conditioned on $[(f_1,f_2)]$ is uniform on $[(f_1,f_2)]$. To see that this is the case, observe that

$$\mathbb{P}_w((f_1,f_2)) \propto \prod_{e \in E^d(R)} w(e)^{(1_{T(f_1)}+1_{T(f_2)})(e)}.$$

Now $1_{T(f_1)} + 1_{T(f_2)} = 1_{T(f_1')} + 1_{T(f_2')}$ whenever $(f_1',f_2') \sim (f_1,f_2)$, which proves the assertion.

Theorem 4.4.9 now provides the distribution of the function g in the measure $\mathbb{P}_w$ conditioned on $[(f_1,f_2)]$. In particular, as $\mathrm{LSD}(g)$ is a tree, the distribution of $g(\mathbf{x})$ is given by summing the outcomes of $d_{\mathrm{LSD}(g)}(\mathbf{0},\mathbf{x})$ fair coin flips, each worth $\pm(d+1)$. It is well-known that the expectation of the square of this random variable is $(d+1)^2 d_{\mathrm{LSD}(g)}(\mathbf{0},\mathbf{x})$, which proves the claim. □

In fact, the exact same calculation works for the covariance of $f(\mathbf{x})$ with $f(\mathbf{y})$.

Theorem 4.5.2. *Work in the setting of the previous theorem. Then for any* $\mathbf{x},\mathbf{y} \in X^d$*, we have*

$$\operatorname{Cov}_w(f(\mathbf{x}),f(\mathbf{y})) = \frac{1}{2}(d+1)^2 \mathbb{E}_w d_{\mathrm{LSD}(g)}(\mathbf{0},\mathbf{z})$$

where $\mathbf{z}$ *is the last* $\mathrm{LSD}(g)$*-vertex of the* $\mathrm{LSD}(g)$*-path from* $\mathbf{0}$ *to* $\mathbf{x}$ *that also appears in the* $\mathrm{LSD}(g)$*-path from* $\mathbf{0}$ *to* $\mathbf{y}$*. In other words, the covariance of* $(f(\mathbf{x}),f(\mathbf{y}))$ *in* $\mathbb{P}_w$ *equals* $\frac{1}{2}(d+1)^2$ *times the expectation of the number of* g*-boundaries that separate both* $\mathbf{x}$ *and* $\mathbf{y}$ *from* $\mathbf{0}$*.*

Proof. Again, we have $\operatorname{Cov}_w(f(\mathbf{x}),f(\mathbf{y})) = \frac{1}{2}\operatorname{Cov}_w(g(\mathbf{x}),g(\mathbf{y}))$, and we prove that

$$\mathbb{E}_w(g(\mathbf{x})g(\mathbf{y})|[(f_1,f_2)]) = (d+1)^2 d_{\mathrm{LSD}(g)}(\mathbf{0},\mathbf{z}).$$

The conditioned measure $\mathbb{P}_w$ directs the edges of $\mathrm{LSD}(g)$ independently and uniformly at random, as in the previous theorem. Thus, under the conditioned measure $\mathbb{P}_w$, we have

$$(g(\mathbf{x}),g(\mathbf{y})) \sim (A+X, A+Y),$$

where A, X, and Y are independent, where A is determined by summing the outcome of $d_{\mathrm{LSD}(g)}(\mathbf{0},\mathbf{z})$ fair independent coin flips each valued $\pm(d+1)$, where X is determined by flipping $d_{\mathrm{LSD}(g)}(\mathbf{z},\mathbf{x})$ coins, and where Y is determined by flipping $d_{\mathrm{LSD}(g)}(\mathbf{z},\mathbf{y})$ coins. This proves the assertion. □

4.6 Generalisation of the Kasteleyn theory

Consider fixed boundary conditions (R,f) and (R,T) with R a region and $T = T(f)$. The goal of this section is to show that $Z = |\Omega(R,f)| = |\Theta(R,T)|$ equals the Cayley hyperdeterminant of the adjacency hypermatrix of a suitably defined hypergraph. In fact, we have no trouble in generalising to Boltzmann measures; we show that one can insert the weights w into the adjacency hypermatrix so that the Cayley hyperdeterminant equals Z_w. The hypergraph, which we shall denote by (U^d, H^d), is dual to the simplicial lattice (X^d, E^d). Recall that U^d is the set of unrooted simplicial loops that was introduced earlier. In dimension $d = 2$ we recover exactly the theory of the dimer model on the hexagonal lattice. (See Figure 4.1c.)

4.6.1 The dual of the simplicial lattice

In this subsection we define the hypergraph (U^d, H^d). Consider first the collection of simplicial loops. If $\mathbf{s} = (\mathbf{s}_k)_{0 \leq k \leq d+1} \in R^d$ is a rooted simplicial loop then $\mathbf{s}$ is characterised by its starting point $\mathbf{s}_0 \in X^d$ and the permutation $\xi \in S_{d+1}$ which describes in which order the increments $(\mathbf{g}_i)_i$ appear. This automatically gives a bijection from R^d to $X^d \times S_{d+1}$. Let us agree to index each unrooted loop $\mathbf{s} \in U^d$ (by default) such that $\mathbf{s}_1 = \mathbf{s}_0 + \mathbf{g}_{d+1}$. There is a unique way of doing so, because the increment $\mathbf{g}_{d+1}$ appears exactly once in each loop. With this convention, each unrooted loop $\mathbf{s} \in U^d$ is characterised by its starting point $\mathbf{s}_0$ and the order $\xi \in S_d$ in which the remaining increments $\{\mathbf{g}_1, \dots, \mathbf{g}_d\}$ appear in the path after the first increment. By adopting the convention we obtain a bijection from U^d to $X^d \times S_d$. We identify the unrooted loop $\mathbf{s}$ with its image under the bijection, so that every pair $(\mathbf{x}, \xi) \in X^d \times S_d$ denotes also an unrooted simplicial loop.

Definition 4.6.1. For any $e \in E^d$, write $h(e)$ for the set of unrooted simplicial loops that traverse e.

Write $e = \{\mathbf{x}, \mathbf{x} + \mathbf{g}_j\} \in E^d$ and let us make a number of observations about the set $h(e)$. First, the assignment $e \mapsto h(e)$ is injective, because the edge e is the only edge that is traversed by all loops in $h(e)$. Secondly, there are precisely $d!$ unrooted simplicial loops that traverse e, since they correspond to the $d!$ ways that we can order the d increments $(\mathbf{g}_i)_{i \neq j}$ that we need to walk back to $\mathbf{x}$ from $\mathbf{x} + \mathbf{g}_j$. Therefore $h(e)$ contains $d!$ unrooted loops. Finally, if $\mathbf{s}^1, \mathbf{s}^2 \in h(e)$ are distinct loops identified with the pairs $(\mathbf{x}^1, \xi^1), (\mathbf{x}^2, \xi^2) \in X^d \times S_d$, then the permutations $\xi^1, \xi^2 \in S_d$ must be distinct. Conclude that for every $\xi \in S_d$, there is a unique $\mathbf{x} \in X^d$ such that $(\mathbf{x}, \xi) \in h(e)$.

The reason that we introduced the map h is the following. A set $T \subset E^d$ is a tiling if and only if $h(T)$ is a partition of U^d, the set of simplicial loops. Once could rephrase this statement by saying that $h(T)$ is a perfect matching of the hypergraph $(U^d, h(E^d))$.

Definition 4.6.2. Write H^d for the set $\{h(e) : e \in E^d\}$. The hypergraph (U^d, H^d) is called the *dual hyperlattice* or simply the *dual* (of the simplicial lattice).

Lemma 4.6.3. *The map $h : T \mapsto \{h(e) : e \in T\}$ is a bijection from Θ to the set of perfect matchings of the hypergraph (U^d, H^d).*

Note that (U^d, H^d) is really dual to (X^d, E^d) because the map h is a bijection from E^d to H^d. The hyperlattice is $d!$-uniform, because every hyperedge $h(e)$ contains $d!$ elements. It is also $d!$-partite with the partition $\{X^d \times \{\xi\} : \xi \in S_d\}$, because every hyperedge $h(e)$ contains one loop in each member $X^d \times \{\xi\}$. The $d!$-partite structure of the dual of the simplicial lattice is special and it is a feature that distinguishes the simplicial lattice from other lattices (in particular, the author is not aware of a similar construction for the square lattice in dimension larger than two). The $d!$-partite structure enables us to generalise the Kasteleyn theory.

4.6.2 The approach suggested by the classical dimer theory

The purpose of this section is to demonstrate that the size of $Z = |\Omega(R, f)| = |\Theta(R, T)|$ equals the Cayley hyperdeterminant of a suitable adjacency hypermatrix. First recall how this works in the Kasteleyn theory for the dimer model on the hexagonal lattice. If $d = 2$ then $d! = 2$, that is, (U^d, H^d) is a regular bipartite graph. In fact, it is really the planar dual of the triangular lattice: the hexagonal lattice. The vertex set U^d is split into its two parts: a set of black and a set of white vertices. A dimer cover (that is, a perfect matching of the graph) matches each black vertex to one white vertex. This is illustrated by Figure 4.1c. The dimer cover is thus encoded by a bijective map σ from the set of black vertices to the set of white vertices; each dimer is of the form $\{b, \sigma(b)\}$ with b ranging over the set of black vertices. To calculate the number of dimer covers, one needs to count the bijections σ from the black vertices to the white that produce a dimer cover. If K is an $n \times n$ matrix, then $\operatorname{Det} K$ is defined as (this is the Leibniz formula)

$$\operatorname{Det} K = \sum_{\sigma \in S_n} \left[\operatorname{Sign} \sigma \prod_{k=1}^{n} K(k, \sigma(k)) \right]. \tag{4.6.4}$$

If the matrix K is suitably chosen, then the term in the square brackets reduces to the indicator function of the event that σ encodes a dimer cover, in which case $\operatorname{Det} K$ equals the number of dimer covers. This is the Kasteleyn theory for dimer models. These observations suggest the following approach, consisting of four steps:

1. First, encode each tiling as a tuple of bijections. It turns out that we need $d! - 1$ bijections in each tuple, because the graph (U^d, H^d) is $d!$-partite, and because we need one bijection for each colour beyond the first. This is Lemma 4.6.8.

2. Second, we show that applying fixed boundary conditions fixes the bijections at all but a finite number of points. Each tiling $Y \in \Theta(R, T)$ is thus encoded as a $(d! - 1)$-tuple of bijections between finite sets. This is Lemma 4.6.10.

3. Third, we define a rank $d!$ adjacency hypermatrix K and construct the Cayley hyperdeterminant for this hypermatrix, such that the each nonzero term in the sum in the definition of $\operatorname{Det} K$ corresponds to a tiling $Y \in \Theta(R, T)$. These are Definitions 4.6.11 and 4.6.13.

4. Finally, each nonzero term in this sum takes value 1 or -1. This is due to the signs that appear in the determinant formula (note that the sign also appears in (4.6.4)). It takes some effort to show that all nonzero terms have the same sign. This is Lemma 4.6.16.

Once this has all been done, it is clear that $Z = |\Omega(R, f)| = |\Theta(R, T)| = \pm \operatorname{Det} K$. The dual hyperlattice (U^d, H^d) plays a crucial role in the analysis. Fix, throughout the remainder of this section, an enumeration $\{\xi^1, \ldots, \xi^{d!}\} = S_d$.

4.6.3 The Kasteleyn theory in dimension $d \geq 2$

We start with Step 1 of the proposed approach. Let $Y \in \Theta$ be a tiling of (X^d, E^d), so that $h(Y)$ is a perfect matching of (U^d, H^d). Each hyperedge $h(e) \in h(Y)$ contains one simplicial loop in each of the $d!$ parts of the partition of U^d. The bijections corresponding to Y are the unique maps

$$\sigma_i : X^d \times \{\xi^1\} \to X^d \times \{\xi^i\} \tag{4.6.5}$$

such that

$$\{\mathbf{s}, \sigma_2(\mathbf{s}), \sigma_3(\mathbf{s}), \ldots, \sigma_{d!}(\mathbf{s})\} \in h(Y) \tag{4.6.6}$$

for every unrooted simplicial loop $\mathbf{s} \in X^d \times \{\xi^1\}$. All elements of $h(Y)$ are given by ranging $\mathbf{s}$ over $X^d \times \{\xi^1\}$ in (4.6.6). This is completely analogous to the dimer model.

Suppose given arbitrary bijections $(\sigma_i)_{2 \le i \le d!}$ as in (4.6.5). Then the set of sets of simplicial loops

$$\Big\{\{\mathbf{s}, \sigma_2(\mathbf{s}), \ldots, \sigma_{d!}(\mathbf{s})\} : \mathbf{s} \in X^d \times \{\xi^1\}\Big\} \tag{4.6.7}$$

is automatically a partition of $U^d = X^d \times S_d$, because each map σ_i is a bijection and therefore each loop $(\mathbf{x}, \xi^i)$ appears precisely once. Conclude that (4.6.7) is a perfect matching of (U^d, H^d) if and only if (4.6.7) is a subset of the hyperedge set H^d. This yields the following result: Step 1 of the suggested approach.

Lemma 4.6.8. *The set of* $(d! - 1)$*-tuples of bijections*

$$\Big(\sigma_i : X^d \times \{\xi^1\} \to X^d \times \{\xi^i\}\Big)_{2 \le i \le d!} \quad \textit{satisfying} \quad \{\mathbf{s}, \sigma_2(\mathbf{s}), \ldots, \sigma_{d!}(\mathbf{s})\} \in H^d$$

for every loop $\mathbf{s} \in X^d \times \{\xi^1\}$ *is in bijection with the perfect matchings of* (U^d, H^d)*. The perfect matching of a tuple (under this bijection) is given by ranging* $\mathbf{s}$ *over* $X^d \times \{\xi^1\}$*; this is precisely the set in* (4.6.7).

Now consider Step 2 of the suggested approach. Suppose given a region R and a tiling T, and consider a tiling $Y \in \Theta(R, T)$. By definition, $Y \in \Theta(R, T)$ if and only if $Y \smallsetminus E^d(R) = T \smallsetminus E^d(R)$. Therefore all loops traversing $T \smallsetminus E^d(R)$ must be matched in the same way as in T, and the loops traversing $T \cap E^d(R)$ can be matched differently. However, the loops that are matched differently are not allowed to produce new hyperedges outside the set $h(E^d(R))$, since we want $Y \smallsetminus E^d(R) = T \smallsetminus E^d(R)$. We first need to identify, for each part of the partition $\{X^d \times \{\xi^i\} : 1 \le i \le d!\}$ of U^d, the set of loops traversing $T \cap E^d(R)$, that is, the loops that are allowed to be matched differently. This motivates the following definition.

Definition 4.6.9. Define, for a fixed region R and a fixed tiling T, and for $1 \le i \le d!$,

$$\begin{aligned} X_i &:= \{(\mathbf{x}, \xi^i) : \text{the loop } (\mathbf{x}, \xi^i) \text{ intersects } T \cap E^d(R)\} \\ &= \{(\mathbf{x}, \xi^i) : \text{the loop } (\mathbf{x}, \xi^i) \text{ does not intersect } T \smallsetminus E^d(R)\} \subset X^d \times \{\xi^i\}. \end{aligned}$$

Observe that $|X_i| = |T \cap E^d(R)|$, and therefore the sets X_i all have the same, finite size. The sets X_i contain the loops that are allowed to match differently. This is Step 2 of the suggested approach.

Lemma 4.6.10. *Let* R *be a region and* T *a tiling. The set of* $(d! - 1)$*-tuples of bijections*

$$(\sigma_i : X_1 \to X_i)_{2 \le i \le d!} \qquad \textit{satisfying} \qquad \{\mathbf{s}, \sigma_2(\mathbf{s}), \ldots, \sigma_{d!}(\mathbf{s})\} \in h(E^d(R))$$

for every loop $\mathbf{s} \in X_1$ *is in bijection with the set of perfect matchings* $h(Y)$ *corresponding to tilings* $Y \in \Theta(R, T)$*. The perfect matching of a tuple (under this bijection) is*

$$\big\{\{\mathbf{s}, \sigma_2(\mathbf{s}), \ldots, \sigma_{d!}(\mathbf{s})\} : \mathbf{s} \in X_1\big\} \cup \big\{h(e) : e \in T \smallsetminus E^d(R)\big\}.$$

The Kasteleyn hypermatrix and its determinant are now straightforwardly defined.

Definition 4.6.11. Let R be a region and T a tiling. Define

$$K : X_1 \times \cdots \times X_{d!} \to \{0,1\},\ (\mathbf{s}_1,\ldots,\mathbf{s}_{d!}) \mapsto 1\left(\{\mathbf{s}_1,\ldots,\mathbf{s}_{d!}\} \in h(E^d(R))\right),$$

where $1(\cdot)$ equals one if the statement inside holds true and zero otherwise. The map K is called the *Kasteleyn hypermatrix.*

From this definition and the previous lemma it follows that

$$|\Theta(R,T)| = \sum_{\sigma_2:X_1\to X_2,\ldots,\sigma_{d!}:X_1\to X_{d!}} \left[\prod_{\mathbf{s}\in X_1} K(\mathbf{s},\sigma_2(\mathbf{s}),\ldots,\sigma_{d!}(\mathbf{s}))\right],$$

where the sum is over bijective maps only. This because the product produces a 1 if the tuple $(\sigma_i)_{2\leq i\leq d!}$ corresponds to an element of $\Theta(R,T)$ and zero otherwise. Recall that $|X_1| = \cdots = |X_{d!}| = |T\cap E^d(R)|$ and write n for this finite number. To simplify notation we identify each set X_i with $[n] := \{1,\ldots,n\}$, so that the previous equality is written

$$|\Theta(R,T)| = \sum_{\sigma_2,\ldots,\sigma_{d!}\in S_n} \left[\prod_{k=1}^{n} K(k,\sigma_2(k),\ldots,\sigma_{d!}(k))\right]. \tag{4.6.12}$$

The expression on the right looks similar to the definition of the determinant of a matrix, and if we insert the signs of the permutations then we obtain precisely the Cayley hyperdeterminant.

Definition 4.6.13. Suppose given a map $A : [n]^m \to \mathbb{C}$ for some $n\in\mathbb{N}$, $m\in 2\mathbb{N}$. Define

$$\begin{aligned}\operatorname{Det} A :=& \sum_{\sigma_2,\ldots,\sigma_m\in S_n}\left(\left[\prod_{i=2}^{m}\operatorname{Sign}\sigma_i\right]\left[\prod_{k=1}^{n} A(\quad k\ ,\sigma_2(k),\ldots,\sigma_m(k))\right]\right) &(4.6.14)\\ =& \frac{1}{n!}\sum_{\sigma_1,\ldots,\sigma_m\in S_n}\left(\left[\prod_{i=1}^{m}\operatorname{Sign}\sigma_i\right]\left[\prod_{k=1}^{n} A(\sigma_1(k),\sigma_2(k),\ldots,\sigma_m(k))\right]\right). &(4.6.15)\end{aligned}$$

This expression is called the *Cayley hyperdeterminant* of A.

The equality in the definition is straightforwardly verified, and it requires m to be even. If we replace A by K in (4.6.14) then (4.6.12) and (4.6.14) are the same, except that some signs appear in (4.6.14) that do not appear in (4.6.12). We conclude that the nonzero terms of the sum in (4.6.14) correspond precisely to the elements of $\Theta(R,T)$. This is Step 3 of the proposed approach. In order to prove that $|\Theta(R,T)| = \pm\operatorname{Det} K$, it suffices to show that all terms of the sum in the definition of $\operatorname{Det} K$ have the same sign (this is Step 4).

Lemma 4.6.16. *Let R be a region and let T be a tiling. Write K for the Kasteleyn hypermatrix. Then all nonzero terms in the sum in the definition of* $\operatorname{Det} K$ *have the same sign.*

Proof. Let R, T and K be as in the lemma. We want to show that all terms of the sum in (4.6.14) (with A replaced with K) have the same sign. The idea is to show that the sign is invariant under making a local move as defined in Subsection 4.2.5. Write f for the unique height function such that $\Phi(f) = (0,T)$. The nonzero terms in (4.6.14) correspond bijectively (through the bijections that we have set up in

Lemma 4.3.2, 3 and in Lemma 4.6.10) to the height functions in $\Omega(R, f)$. We pick two height functions $f', f'' \in \Omega(R, f)$ and prove that the corresponding terms in (4.6.14) have the same sign. By Lemma 4.2.3, we may assume, without loss of generality, that $f'' = f' + (d+1) \cdot 1_{\mathbf{x}}$ for some $\mathbf{x} \in R$. Let $T', T'' \in \Theta(R, T)$ be the tilings corresponding to f', f'' respectively. Recall that

$$T' = \left\{\{\mathbf{y}, \mathbf{y} + \mathbf{g}_i\} \in E^d : \nabla f'(\mathbf{y}, \mathbf{y} + \mathbf{g}_i) = -d\right\},$$

and for T'' we have an identical expression in terms of f''. Remark that $f'' = f'$ except at the point $\mathbf{x}$, and therefore $\nabla f = \nabla f'$ except at the edges incident to $\mathbf{x}$. Since $f'' = f' + (d+1) \cdot 1_{\mathbf{x}}$ and since both f' and f'' are height functions, we must have

$$\begin{aligned}\nabla f'(\mathbf{x}, \mathbf{x} + \mathbf{g}_i) = \nabla f''(\mathbf{x} - \mathbf{g}_i, \mathbf{x}) = 1,\\ \nabla f'(\mathbf{x} - \mathbf{g}_i, \mathbf{x}) = \nabla f''(\mathbf{x}, \mathbf{x} + \mathbf{g}_i) = -d,\end{aligned}$$

and therefore

$$\begin{aligned}T' \smallsetminus T'' = \{\{\mathbf{x}, \mathbf{x} - \mathbf{g}_i\} : 1 \le i \le d+1\},\\ T'' \smallsetminus T' = \{\{\mathbf{x}, \mathbf{x} + \mathbf{g}_i\} : 1 \le i \le d+1\}.\end{aligned} \tag{4.6.17}$$

This means that all loops are matched the same (in the matchings $h(T')$ and $h(T'')$), except for the loops traversing the vertex $\mathbf{x}$. In order to prove the lemma, we work out the effect of this difference on the signs in (4.6.14). Let $(\sigma_i')_{2 \le i \le d!}$ denote the bijections from Lemma 4.6.10 corresponding to T'. This means that

$$\{\{\mathbf{s}, \sigma_2'(\mathbf{s}), \ldots, \sigma_{d!}'(\mathbf{s})\} : \mathbf{s} \in X_1\} = h(T' \cap E^d(R)).$$

Define, for each $1 \le i \le d!$, the bijections

$$\delta_i : X_i \to X_i,\ (\mathbf{x}, \xi^i) \mapsto \begin{cases}(\mathbf{x} + \mathbf{g}_j, \xi^i) & \text{if } (\mathbf{x}, \xi^i) \text{ traverses } \{\mathbf{x} - \mathbf{g}_j, \mathbf{x}\} \text{ for some } j,\\ (\mathbf{x}, \xi^i) & \text{if } (\mathbf{x}, \xi^i) \text{ does not traverse } \mathbf{x}.\end{cases}$$

Note that δ_i is a permutation consisting of one cycle of length $d+1$. Claim that

$$\{\{\delta_1(\mathbf{s}), \delta_2 \circ \sigma_2'(\mathbf{s}), \ldots, \delta_{d!} \circ \sigma_{d!}'(\mathbf{s})\} : \mathbf{s} \in X_1\} = h(T'' \cap E^d(R)). \tag{4.6.18}$$

To support the claim, recall (4.6.17) and observe simply that

$$\begin{aligned}&\{\delta_1(\mathbf{s}), \delta_2 \circ \sigma_2'(\mathbf{s}), \ldots, \delta_{d!} \circ \sigma_{d!}'(\mathbf{s})\}\\ &\quad = \begin{cases}h(\{\mathbf{x}, \mathbf{x} + \mathbf{g}_i\}) & \text{if } \{\mathbf{s}, \sigma_2'(\mathbf{s}), \ldots, \sigma_{d!}'(\mathbf{s})\} = h(\{\mathbf{x}, \mathbf{x} - \mathbf{g}_i\}) \text{ for some } i,\\ \{\mathbf{s}, \sigma_2'(\mathbf{s}), \ldots, \sigma_{d!}'(\mathbf{s})\} & \text{otherwise.}\end{cases}\end{aligned}$$

This proves the claim. Since $\delta_1 : X_1 \to X_1$ is a bijection, the sets in (4.6.18) are equal to

$$\{\{\mathbf{s}, \delta_2 \circ \sigma_2' \circ \delta_1^{-1}(\mathbf{s}), \ldots, \delta_{d!} \circ \sigma_{d!}' \circ \delta_1^{-1}(\mathbf{s})\} : \mathbf{s} \in X_1\}$$

This implies that the bijections from Lemma 4.6.10 corresponding to T'' are, for $2 \le i \le d!$,

$$\sigma_i'' = \delta_i \circ \sigma_i' \circ \delta_1^{-1} : X_1 \to X_i.$$

Now note that $\operatorname{Sign} \delta_i = (-1)^d$ (since δ_i is a cycle of length $d+1$). Conclude that $\operatorname{Sign} \delta_i \cdot \operatorname{Sign} \delta_1^{-1} = 1$, and therefore σ_i' and σ_i'' have the same sign in (4.6.14), for all i. $\square$

We have completed the final step of the approach that was suggested by the Kasteleyn theory for dimer covers. This yields the following theorem.

Theorem 4.6.19. *Let R be a region, let f be a height function, and let $T = T(f)$. Write K for the Kasteleyn hypermatrix. Then $Z = |\Omega(R, f)| = |\Theta(R, T)| = \pm \operatorname{Det} K$.*

4.6.4 Boltzmann measures

Recall the definition of a Boltzmann measure in Subsection 4.3.1. The number $|\Omega(R, f)| = |\Theta(R, T)|$ equals the partition function Z of the uniform probability measures on $\Omega(R, f)$ and $\Theta(R, T)$. The Kasteleyn theory is easily generalised to Boltzmann measures by inserting the weights into the Kasteleyn hypermatrix.

Definition 4.6.20. Let R be a region, T a tiling, and $w : E^d(R) \to \mathbb{C}$ any (complex-valued) weight function. Define

$$\begin{aligned} &K_w : X_1 \times \cdots \times X_{d!} \to \mathbb{C}, \\ &\qquad (\mathbf{s}_1, \ldots, \mathbf{s}_{d!}) \mapsto 1\left(\{\mathbf{s}_1, \ldots, \mathbf{s}_{d!}\} \in h(E^d(R))\right) \cdot w\left(h^{-1}(\{\mathbf{s}_1, \ldots, \mathbf{s}_{d!}\})\right). \end{aligned}$$

The map K_w is called the *weighted Kasteleyn hypermatrix.*

By comparing the definition of Z_w with the definition of the Cayley hyperdeterminant, and taking into account Lemma 4.6.16, it is readily verified that $Z_w = \pm \operatorname{Det} K_w$.

4.7 Gradient Gibbs measures

In previous sections we introduced fixed boundary conditions and periodic boundary conditions, which enabled us to study probability measures on finite subsets of Ω. This section introduces *shift-invariant gradient Gibbs measures*, which prove to be an effective tool for studying the large-scale behaviour of the model. While gradient Gibbs measures are interesting in their own right, their main purpose here are their use in the proof of strict convexity of the surface tension in Section 4.9.

4.7.1 Definition

Write f for the random function in Ω. Define for any $R \subset X^d$,

$$\begin{aligned} \mathcal{F} &:= \sigma(f(\mathbf{x}) : \mathbf{x} \in X^d), & \mathcal{F}_R &:= \sigma(f(\mathbf{x}) : \mathbf{x} \in R), \\ \mathcal{F}^\nabla &:= \sigma(f(\mathbf{x}) - f(\mathbf{y}) : \mathbf{x}, \mathbf{y} \in X^d), & \mathcal{F}_R^\nabla &:= \sigma(f(\mathbf{x}) - f(\mathbf{y}) : \mathbf{x}, \mathbf{y} \in R). \end{aligned}$$

Note that $\mathcal{F}_R^\nabla = \mathcal{F}^\nabla \cap \mathcal{F}_R$ is finite whenever R is finite because it is generated by finitely many random variables, each taking finitely many values. Write $\mathcal{P}(\Omega, \mathcal{X})$ for the collection of probability measures on the measurable space $(\Omega, \mathcal{X})$ for any σ-algebra $\mathcal{X}$ on Ω. Probability measures in $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$ are called *gradient measures.*

A gradient measure $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ is called *shift-invariant* whenever $\mu(\tilde{\theta} A) = \mu(A)$ for any $A \in \mathcal{F}^\nabla$ and $\theta \in \Theta$, where $\tilde{\theta} A := \{\tilde{\theta} f : f \in A\}$. In other words, a gradient measure μ is shift-invariant whenever ∇f and $\theta \nabla f$ have the same law in μ for every $\theta \in \Theta$. The set of shift-invariant gradient measures is denoted by $\mathcal{P}_\Theta(\Omega, \mathcal{F}^\nabla)$. If $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F}^\nabla)$, then it follows from shift-invariance that the map $\mu(f(\cdot) - f(\mathbf{0})) : X^d \to \mathbb{R}$ is additive over X^d. Therefore there exists a unique $s \in H^*$

such that $s(\mathbf{x}) = \mu(f(\mathbf{x}) - f(\mathbf{0}))$ for every $\mathbf{x} \in X^d$, and we must have $s \in \mathcal{S}$ because $s(\mathbf{g}_i) = \mu(f(\mathbf{g}_i) - f(\mathbf{0})) \leq 1$ for every $1 \leq i \leq d+1$. Write $s(\mu)$ for s, the *slope* of $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F}^\nabla)$.

Let (L, s) denote valid periodic boundary conditions and let μ denote the probability measure that is uniformly random in the finite set $\Omega(L, s)$. Lemma 4.3.10 implies that μ restricts to a shift-invariant gradient measure of slope $s(\mu) = s$.

Let us now introduce the notion of a Gibbs measure. Fix a measure $\mu \in \mathcal{P}(\Omega, \mathcal{F})$. The measure μ is called a *Gibbs measure* if for every finite $R \subset X^d$, the distribution of f in μ is the same as the distribution of a sample f obtained by first sampling g from μ, then sampling f from $\Omega(R, g)$ uniformly at random. The definition is formalised in terms of specifications and the Dobrushin-Lanford-Ruelle (DLR) equations. For each finite $R \subset X^d$, let γ_R denote the probability kernel from $(\Omega, \mathcal{F}_{R^c})$ to $(\Omega, \mathcal{F})$ such that for any $f \in \Omega$, the probability measure $\gamma_R(\cdot, f)$ is uniform in $\Omega(R, f)$. It is obvious from the definition that $\Omega(R, f)$ is invariant under changing the values of f on R, so that $\gamma_R(A, \cdot)$ is indeed $\mathcal{F}_{R^c}$-measurable for every $A \in \mathcal{F}$. The kernels γ_R satisfy the consistency condition; if $S \subset R$, then $\gamma_R \gamma_S = \gamma_R$. The collection of probability kernels γ_R is called a specification, and a measure $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ is called a *Gibbs measure* if μ satisfies the DLR equation

$$\mu = \mu \gamma_R \tag{4.7.1}$$

for each finite $R \subset X^d$. This is equivalent to our previous, informal description. By the consistency condition it is sufficient to check the DLR equations for an increasing exhaustive sequence of finite subsets of X^d. Each kernel γ_R restricts to a kernel from $(\Omega, \mathcal{F}^\nabla_{R^c})$ to $(\Omega, \mathcal{F}^\nabla)$. We shall write γ_R^∇ for this restriction. A gradient measure $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ is called a *gradient Gibbs measure* if

$$\mu = \mu \gamma_R^\nabla$$

for each finite subset R of X^d.

4.7.2 Existence and concentration

Theorem 4.7.2. *For each slope $s \in \mathcal{S}$, there is a shift-invariant gradient Gibbs measure $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F}^\nabla)$ of slope s such that, for any $\mathbf{x}, \mathbf{y} \in X^d$, we have the bounds*

1. $\operatorname{Var}_\mu(f(\mathbf{y}) - f(\mathbf{x})) \leq (d+1)^2 n$,
2. $\mu(f(\mathbf{y}) - f(\mathbf{x}) - s(\mathbf{y} - \mathbf{x}) \geq (d+1)a) \leq \exp -\frac{a^2}{2n}$ *for all* $a \geq 0$ *whenever* $n > 0$,
3. $\mu(f(\mathbf{y}) - f(\mathbf{x}) - s(\mathbf{y} - \mathbf{x}) \leq (d+1)a) \leq \exp -\frac{a^2}{2n}$ *for all* $a \leq 0$ *whenever* $n > 0$,

where $n = d_{(X^d, E^d)}(\mathbf{x}, \mathbf{y})$.

The *topology of local convergence* or *$\mathcal{L}$-topology* on $\mathcal{P}(\Omega, \mathcal{X})$ is the coarsest topology that makes the evaluation map $\mu \mapsto \mu(A)$ continuous for every finite $R \subset X^d$ and for any $A \in \mathcal{X} \cap \mathcal{F}_R$. Constructing (gradient) Gibbs measures on $\mathcal{P}(\Omega, \mathcal{X})$ is much easier whenever choosing $\mathcal{X} = \mathcal{F}^\nabla$ and not $\mathcal{X} = \mathcal{F}$, because $\mathcal{F}^\nabla_R$ is finite for finite $R \subset X^d$—see the following lemma.

Lemma 4.7.3. *The set $\mathcal{P}(\Omega, \mathcal{F}^\nabla)$ is compact in the topology of local convergence.*

Proof. The proof is entirely straightforward. Let $(\mu_n)_{n\in\mathbb{N}}$ denote a sequence of measures in $\mathcal{P}(\Omega,\mathcal{F}^\nabla)$ and let $(\Gamma_m)_{m\in\mathbb{N}}$ denote an increasing exhaustive sequence of finite subsets of X^d. Fix $m\in\mathbb{N}$. The σ-algebra $\mathcal{F}^\nabla_{\Gamma_m}$ is finite and therefore there exists a subsequence $(k_n)_{n\in\mathbb{N}}\subset\mathbb{N}$ such that μ_{k_n} converges on $\mathcal{F}^\nabla_{\Gamma_m}$ as $n\to\infty$. By a standard diagonalisation argument we may assume that convergence occurs for all $m\in\mathbb{N}$. The limiting measure exists by the Kolmogorov extension theorem. □

Proof of Theorem 4.7.2. Let $s\in\mathcal{S}$ be the slope of interest. Let $(s_n)_{n\in\mathbb{N}}$ be a sequence of slopes converging to s with $s_n\in\mathcal{S}_n$ for every n. Write μ_n for the uniform probability measure on $\Omega(L_n,s_n)$, for every $n\in\mathbb{N}$. Each measure μ_n restricts to a shift-invariant gradient measure in $\mathcal{P}_\Theta(\Omega,\mathcal{F}^\nabla)$, and $s(\mu_n)=s_n$.

Now apply the previous lemma to obtain a subsequence $(k_n)_{n\in\mathbb{N}}$ along which the sequence of gradient measures $(\mu_n)_{n\in\mathbb{N}}$ converges in the topology of local convergence, say to $\mu\in\mathcal{P}(\Omega,\mathcal{F}^\nabla)$. The limit μ must be shift-invariant as all measures $(\mu_n)_{n\in\mathbb{N}}$ are shift-invariant. At each vertex $\mathbf{x}\in X^d$ we have

$$\mu(f(\mathbf{x})-f(\mathbf{0}))=\lim_{n\to\infty}\mu_{k_n}(f(\mathbf{x})-f(\mathbf{0}))=\lim_{n\to\infty}s_{k_n}(\mathbf{x})=s(\mathbf{x}),$$

which means that $s(\mu)=s$. One shows similarly that (1)–(3) follow from Theorem 4.3.12.

It suffices to prove that the gradient measure μ is a Gibbs measure, that is, that $\mu\gamma_R^\nabla=\mu$ for every finite $R\subset X^d$. Fix a finite subset $R\subset X^d$. Now suppose that $\mu\gamma_R^\nabla$ equals μ on $\mathcal{F}_S^\nabla$ for any finite $S\subset X^d$. Then the two measures must be the same, by the uniqueness statement of the Kolmogorov extension theorem. It thus suffices to prove that $\mu\gamma_R^\nabla$ equals μ on $\mathcal{F}_S^\nabla$ for any finite $S\subset X^d$. We may assume that $R\subset S$ and $\partial R\subset S$ by expanding S if necessary. By using shift-invariance, we may finally assume that $\mathbf{0}\notin S$.

We make the stronger claim that already in the non-gradient setting and before taking limits, we have

$$\mu_n\gamma_R|_{\mathcal{F}_S}=\mu_n|_{\mathcal{F}_S}\tag{4.7.4}$$

for n sufficiently large. The distribution μ_n is not invariant under resampling f on R. However, if $R+\mathbf{x}$ and $R+\mathbf{y}$ are disjoint and not adjacent for any $\mathbf{x},\mathbf{y}\in L_n$ distinct, then μ_n is invariant under resampling f on R, then translating this modification to $R+\mathbf{x}$ for each $\mathbf{x}\in L_n\smallsetminus\{\mathbf{0}\}$. Thus, if n is so large that S and $R+\mathbf{x}$ are disjoint for any $\mathbf{x}\in L_n\smallsetminus\{\mathbf{0}\}$, then (4.7.4) holds true because the additional modifications do not affect the values of μ_n on the σ-algebra $\mathcal{F}_S$. This proves the claim. □

For any $n\in\mathbb{N}$, let Π_n denote a *centred box* of sides $2n$, that is,

$$\Pi_n:=\{a_1\mathbf{g}_1+\cdots+a_d\mathbf{g}_d:-n\leq a_1,\ldots,a_d<n\}\subset X^d.$$

Note that $|\Pi_n|=(2n)^d$.

Proposition 4.7.5. *Let μ denote a measure of Theorem 4.7.2 of slope $s\in\mathcal{S}$. Then μ-almost surely*

$$\lim_{n\to\infty}\frac{1}{n}\|(f-f(\mathbf{0}))|_{\Pi_n}-s|_{\Pi_n}\|_\infty=0.\tag{4.7.6}$$

This follows from a union bound and the inequalities in Theorem 4.7.2.

4.8 The surface tension and the variational principle

The purpose of this section is to give an overview of three closely related concepts which describe the macroscopic behaviour of the model. These motivate the study of strict convexity of the surface tension in Section 4.9. First, there is indeed the surface tension, which describes the asymptotic number of height functions approximating a certain slope. Second, there is the large deviations principle, which describes the asymptotic number of height functions approximating an arbitrary continuous profile. The rate function is the integral of the surface tension over the gradient of the continuous profile of interest. Third, there is the variational principle, which is a direct corollary of the large deviations principle, and describes the typical macroscopic behaviour of the random height function. The surface tension is usually convex, making that the rate function in the large deviations principle is also convex. In the next section, we shall also prove that the surface tension is strictly convex, which implies that the rate function has a unique minimiser, which in turn implies concentration around a single continuous profile in the variational principle. For the results in this section, we refer to [54] and [36].

4.8.1 The surface tension

Definition 4.8.1. Recall the definition of $\lfloor f \rfloor$ for Lipschitz functions $f : H \to \mathbb{R}$ on Page 156. Recall also the definition of $\Pi_n \subset X^d$ at the end of the previous section (Page 173). The *surface tension* is the function $\sigma : \mathcal{S} \to \mathbb{R}$ defined by

$$\sigma(s) := \lim_{n\to\infty} -\frac{1}{|\Pi_n|} \log |\Omega(\Pi_n, \lfloor s \rfloor)|.$$

For convergence of the limit in the definition of $\sigma(s)$, we refer to Section 4 in [36]. The argument is effectively a supermultiplicativity argument: if $A, B \subset X^d$ are finite and disjoint with no vertex of A adjacent to a vertex of B, then $|\Omega(A \cup B, f)| = |\Omega(A, f)| \cdot |\Omega(B, f)|$, and if $A \subset B$, then $|\Omega(A, f)| \leq |\Omega(B, f)|$. In fact, the definition of $\sigma(s)$ is stable under modifications of order $o(n)$ to $\lfloor s \rfloor$ as $n \to \infty$; see Lemma 4.5 in [36] for the following result.

Theorem 4.8.2. *If $s \in \mathcal{S}$, and if $(f_n)_{n\in\mathbb{N}} \subset \Omega$ satisfies $\|f_n|_{\Pi_n} - s|_{\Pi_n}\|_\infty = o(n)$ as $n \to \infty$, then*

$$\sigma(s) = \lim_{n\to\infty} -\frac{1}{|\Pi_n|} \log |\Omega(\Pi_n, f_n)|.$$

The previous result implies immediately that σ is continuous, see also Lemma 4.3 in [36].

Theorem 4.8.3. *The surface tension $\sigma : \mathcal{S} \to \mathbb{R}$ is continuous.*

4.8.2 The large deviations principle

Before stating the large deviations principle, we must introduce a suitable topological space to work in, and we must specify how a sequence of fixed boundary conditions converges to a continuous boundary profile. This is the purpose of the following definition.

Definition 4.8.4. Write $\mathrm{Lip}(D)$ for the collection of real-valued Lipschitz functions on D for any $D \subset H$. A *domain* is a bounded open set $D \subset H$ such that ∂D has zero Lebesgue measure. A *boundary profile* is a pair (D, b) where D is a domain and $b \in \operatorname{Lip}(\partial D)$. An *approximation* of (D, b) is a sequence of pairs $((D_n, b_n))_{n\in\mathbb{N}}$ such that $D_n \subset X^d$ is finite and $b_n \in \Omega$ for any $n \in \mathbb{N}$, and such that

$$\frac{1}{n} D_n \to D, \qquad \frac{1}{n} \operatorname{Graph} b_n|_{\partial D_n} \to \operatorname{Graph} b$$

in the Hausdorff topologies on H and $H \times \mathbb{R}$ respectively as $n \to \infty$.

If (D, b) is a boundary profile with approximation $((D_n, b_n))_{n\in\mathbb{N}}$, then write $(\gamma_n)_{n\in\mathbb{N}}$ for the sequence of measures defined by $\gamma_n := \gamma_{D_n}(\cdot, b_n)$, the uniform probability measure in the finite set $\Omega(D_n, b_n)$. The topological space for the large deviations principle associated with this sequence is the set $\operatorname{Lip}(\bar{D})$ endowed with the topology of uniform convergence—which is equivalent to the topology of pointwise convergence as $\bar{D}$ is compact. We must bring all samples from each measure γ_n to the space $\operatorname{Lip}(\bar{D})$ for the large deviations principle to make sense. For each $n \in \mathbb{N}$, define the map $K_n : \Omega \to \operatorname{Lip}(\bar{D})$ as follows. First, for each $f \in \Omega$, define $\bar{f}$ to be the smallest Lipschitz extension of f to H. Define each map K_n by

$$K_n(f) : \bar{D} \to \mathbb{R},\, x \mapsto \frac{1}{n} \bar{f}(nx).$$

Finally, let λ denote the unique translation-invariant measure on H for which

$$\{a_1\mathbf{g}_1 + \cdots + a_d\mathbf{g}_d : a_1, \ldots, a_d \in [0, 1]\} \subset H$$

has measure one.

Theorem 4.8.5. *Let (D, b) denote a boundary profile and $((D_n, b_n))_{n\in\mathbb{N}}$ an approximation of (D, b) with associated measures $\gamma_n := \gamma_{D_n}(\cdot, b_n)$. Write γ_n^* for the pushforward of γ_n along K_n. Then the sequence of measures $(\gamma_n^*)_{n\in\mathbb{N}}$ satisfies a large deviations principle in the topological space $\operatorname{Lip}(\bar{D})$ with speed n^d and rate function*

$$I(f) := -P(D, b) + \begin{cases} \int_D \sigma(\nabla f) d\lambda & \text{if } f|_{\partial D} = b, \\ \infty & \text{otherwise,} \end{cases}$$

where $P(D, b)$ is called the pressure *of the boundary profile (D, b), defined to be the unique constant such that the minimum of I is 0, and equal to*

$$P(D, b) = \lim_{n\to\infty} -\frac{1}{n^d} \log |\Omega(D_n, b_n)|.$$

We shall continue using the definitions of I and $P(D, b)$ in the sequel. The large deviations principle was proven in a much more general setting by Sheffield in [54]. The large deviations principle with boundary conditions is stated in Subsection 7.3.2. The large deviations principle in [54] does not only address the macroscopic profile of each height functions but also its “local statistics” within macroscopic regions, something we are not concerned with here. In [54] it is required that the boundary profile is “not taut”. This requirement is however only necessary to understand the local statistics, and may be omitted when one is interested in the macroscopic profile only. For a more recent and elementary proof of the large deviations principle for the macroscopic profile only, we refer to Theorem 2.17 in [36].

4.8.3 The variational principle

The variational principle is a direct corollary of the large deviations principle. Note that the set of minimisers of I in Theorem 4.8.5 is exactly the set of minimisers of

$$\int_D \sigma(\nabla f) d\lambda$$

over all functions $f \in \operatorname{Lip}(\bar{D})$ which restrict to b on ∂D.

Theorem 4.8.6. *Assume the setting of Theorem 4.8.5. Let A denote an open neighbourhood of $\{I = 0\} \subset \operatorname{Lip}(\bar{D})$. Then $\gamma_n^*(A) \to 1$ as $n \to \infty$. In particular, if σ is strictly convex, then I has a unique minimiser $f^* \in \operatorname{Lip}(\bar{D})$, and in that case $\gamma_n^*(A) \to 1$ as $n \to \infty$ for any open neighbourhood A of f^*.*

4.9 Strict convexity of the surface tension

In [5], the authors find an explicit formula for σ for the case $d = 2$ by appealing to the integrable nature of the model. A direct corollary is that the surface tension is strictly convex. In [54], Sheffield proves that the surface tension related to *any* simply attractive model is strictly convex. In particular, this implies the following theorem.

Theorem 4.9.1. *For any $d \geq 2$, the surface tension is strictly convex on the interior of $\mathcal{S}$.*

The proof of Sheffield relies crucially on cluster swapping. The purpose of the section is to give an alternative proof of Theorem 4.9.1, which is simpler than the proof in [54] due to the special nature of the cluster swap in the particular setting of this thesis.

4.9.1 The specific entropy

First, we give an alternative characterisation of $\sigma(s)$ in terms of the shift-invariant gradient Gibbs measure of slope s whose existence is guaranteed by Theorem 4.7.2. For this, we require the notions of *entropy* and *specific entropy*. Let $(X, \mathcal{X})$ be an arbitrary measurable space endowed with a probability measure μ and a nonzero finite measure ν. Then the *relative entropy of μ with respect to ν*, denoted $\mathcal{H}(\mu, \nu)$, is defined by

$$\mathcal{H}(\mu, \nu) := \begin{cases} \nu(h \log h) = \mu(\log h) & \text{if } \mu \ll \nu \text{ and } h = d\mu/d\nu, \\ \infty & \text{if } \mu \not\ll \nu. \end{cases}$$

If $\mathcal{A}$ is a sub-σ-algebra of $\mathcal{X}$, then write $\mathcal{H}_\mathcal{A}(\mu, \nu)$ for $\mathcal{H}(\mu|_\mathcal{A}, \nu|_\mathcal{A})$. It is well-known that μ minimises $\mathcal{H}(\cdot, \nu)$ over all probability measures if and only if μ is the normalised version of ν, in which case $\mathcal{H}(\mu, \nu) = -\log \nu(X)$. Also, if ν is a counting measure, then $h \leq 1$, and in that case $\mathcal{H}(\mu, \nu) \leq 0$.

If R is a finite subset of X^d, then write $D_R : \Omega \to \mathbb{Z}^{R \times R}$ for the map satisfying

$$(D_R f)(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}) - f(\mathbf{x})$$

for every $f \in \Omega$, $\mathbf{x}, \mathbf{y} \in R$. Call D_R the *differences map.* Note that $\operatorname{Im} D_R$ is finite, and that $\mathcal{F}_R^\nabla = \sigma(D_R)$. Stronger: D_R may be seen as a bijection from $\mathcal{F}_R^\nabla$ to the powerset of $\operatorname{Im} D_R$. Write λ^R for the pullback of the counting measure on $\operatorname{Im} D_R$ along the map D_R—λ^R is a measure on $(\Omega, \mathcal{F}_R^\nabla)$ of size $\lambda^R(\Omega) = |\operatorname{Im} D_R| \in \mathbb{Z}_{>0}$.

Remark 4.9.2. If R is connected and $f \in \Omega$, then the values of $D_R f$ can be recovered from the values of ∇f on the edges of E^d which are contained in R, by integrating ∇f along the appropriate paths through R. Thus, for connected sets $R \subset X^d$, one should think of the map D_R as an alternative for the map

$$f \mapsto (\nabla f)|_{E^d \cap (R \times R)}$$

in the above construction. This also implies that $|\operatorname{Im} D_R| \leq 2^{|R|}$ whenever R is connected.

Let $\mu \in \mathcal{P}(\Omega, \mathcal{F}^\nabla)$ and $R \subset X^d$ finite. Then the *entropy of μ in R*, denoted $\mathcal{H}_R(\mu)$, is defined by

$$\mathcal{H}_R(\mu) := \mathcal{H}_{\mathcal{F}_R^\nabla}(\mu, \lambda^R) = \sum_{x \in \operatorname{Im} D_R} \mu(D_R f = x) \log \mu(D_R f = x) \in [-\log|\operatorname{Im} D_R|, 0].$$

The *specific entropy* of μ, denoted $\mathcal{H}(\mu)$, is defined to be the limit

$$\mathcal{H}(\mu) := \lim_{n \to \infty} \frac{1}{|\Pi_n|} \mathcal{H}_{\Pi_n}(\mu) = \lim_{n \to \infty} \frac{1}{|\Pi_n|} \mathcal{H}_{\mathcal{F}_{\Pi_n}^\nabla}(\mu, \lambda^{\Pi_n})$$

whenever the sequence is convergent. Otherwise simply replace the limit by the limit inferior to obtain a well-defined limit. It can in fact be shown that the sequence is always convergent, see for example [54, Chapter 2], but we shall not rely on this fact.

Theorem 4.9.3. *Let μ denote a measure of Theorem 4.7.2 of slope $s \in \mathcal{S}$. Then $\mathcal{H}(\mu) = \sigma(s)$.*

Proof. Let μ denote any gradient Gibbs measure for now. Write h^R for the Radon-Nikodym derivative

$$h^R := \frac{d\mu|_{\mathcal{F}_R^\nabla}}{d\lambda^R}$$

for any finite $R \subset X^d$. Fix $R, S \subset X^d$ finite with $S \cup \partial S \subset R$. As μ is Gibbs, we know that μ is uniformly random in $\Omega(S, f)$ whenever μ is conditioned on the values of f on S^c. This implies immediately that

$$h^R = \frac{1}{|\Omega(S, f)|} h^{R \smallsetminus S}.$$

The function $|\Omega(S, \cdot)|$ is $\mathcal{F}_{\partial S}^\nabla$-measurable as the model is Markov, and consequently h^R is $\mathcal{F}_{R \smallsetminus S}^\nabla$-measurable.

Let now μ be a measure of Theorem 4.7.2 of slope $s \in \mathcal{S}$, and pick $n \in \mathbb{N}$. Then

$$\begin{aligned}(2n)^{-d} \mathcal{H}_{\Pi_n}(\mu) &= (2n)^{-d} \mu(\log h^{\Pi_n}) = (2n)^{-d} \mu(\log h^{\Pi_n \smallsetminus \Pi_{n-1}} - \log|\Omega(\Pi_{n-1}, f)|) \\ &= (2n)^{-d} \mathcal{H}_{\Pi_n \smallsetminus \Pi_{n-1}}(\mu) + \mu(-(2n)^{-d} \log|\Omega(\Pi_{n-1}, f)|). \end{aligned} \quad (4.9.4)$$

The first term in (4.9.4) vanishes as $n \to \infty$ because $\Pi_n \smallsetminus \Pi_{n-1}$ is connected as a subset of (X^d, E^d):

$$|\mathcal{H}_{\Pi_n \smallsetminus \Pi_{n-1}}(\mu)| \leq \log|\operatorname{Im} D_{\Pi_n \smallsetminus \Pi_{n-1}}| \leq |\Pi_n \smallsetminus \Pi_{n-1}| \log 2 = O(n^{d-1}).$$

The term within the expectation in (4.9.4) converges to $\sigma(s)$ pointwise by Proposition 4.7.5 and Theorem 4.8.2. We may apply the dominated convergence theorem because the expression within the expectation is always absolutely bounded by $\log 2$, since $0 \leq \log|\Omega(\Pi_{n-1}, f)| \leq |\Pi_{n-1}| \log 2 \leq (2n)^d \log 2$. $\square$

Before proceeding, let us quote an important result from the literature.

Theorem 4.9.5. *Let $\mu \in \mathcal{P}_\Theta(\Omega, \mathcal{F}^\nabla)$ denote a measure which satisfies the concentration of* (4.7.6) *for some $s \in \mathcal{S}$, but which is not a Gibbs measure. Then $\mathcal{H}(\mu) > \sigma(s)$.*

Proof overview. Write $\mu^{n,g}$ for the measure μ conditioned on $f(\mathbf{x}) = g(\mathbf{x})$ for all $\mathbf{x} \in \Pi_n \smallsetminus \Pi_{n-1}$. Then

$$\mathcal{H}_{\Pi_n}(\mu) = \mathcal{H}_{\Pi_n \smallsetminus \Pi_{n-1}}(\mu) + \int \mathcal{H}_{\Pi_n}(\mu^{n,g}) d\mu(g).$$

This is the same decomposition as in the proof of the previous theorem. For fixed g, the integrand in this display is clearly minimised if μ is a Gibbs measure, because $\mu^{n,g}$ is then uniformly random in all extensions of $g|_{\Pi_n \smallsetminus \Pi_{n-1}}$ to Π_n. This proves that $\mathcal{H}(\mu) \geq \sigma(s)$ whenever μ is concentrated as in (4.7.6); the difficulty is in proving the strict inequality whenever μ is not Gibbs. If μ is not Gibbs, then the integral in the display will for some n be strictly larger than if μ were Gibbs, but it is nontrivial to demonstrate that this difference survives the normalisation by $|\Pi_n|$ in the definition of $\mathcal{H}(\mu)$. This follows from a standard superadditivity argument, see Lemma 2.4.1 in Chapter 2 or Theorem 15.37 in [20]. See Theorem 2.5.2 in [54] for a proof of the current theorem in full detail. □

4.9.2 The product setting

For the double dimer model, the cluster swap, and the level set decomposition developed in this thesis, it is essential to work in the product setting. We shall introduce some straightforward technical machinery before proceeding; essentially we must adapt the constructions and results from Section 4.7 and from the previous subsection to the product setting. Write

$$\begin{aligned} \mathcal{F}^{2\nabla} &:= \mathcal{F}^\nabla \times \mathcal{F}^\nabla, & \gamma_R^2 &:= \gamma_R \times \gamma_R, & \lambda_2^R &:= \lambda^R \times \lambda^R, \\ \mathcal{F}_R^{2\nabla} &:= \mathcal{F}_R^\nabla \times \mathcal{F}_R^\nabla, & \gamma_R^{2\nabla} &:= \gamma_R^\nabla \times \gamma_R^\nabla. \end{aligned}$$

Let $\mathcal{P}(\Omega^2, \mathcal{F}^{2\nabla})$ denote the collection of probability measures on $(\Omega^2, \mathcal{F}^{2\nabla})$; such measures are called *double gradient measures*. If $\mu \in \mathcal{P}(\Omega^2, \mathcal{F}^{2\nabla})$ then we shall by default write (f_1, f_2) for the pair of random height functions, and $g := f_1 - f_2$ for the random difference. Write $\mathcal{P}_\Theta(\Omega^2, \mathcal{F}^{2\nabla})$ for the collection of shift-invariant measures $\mu \in \mathcal{P}(\Omega^2, \mathcal{F}^{2\nabla})$; the measure μ is called *shift-invariant* if $\mu(\tilde{\theta}A \times \tilde{\theta}B) = \mu(A \times B)$ for every $\theta \in \Theta$ and $A, B \in \mathcal{F}^\nabla$. This is equivalent to requiring that $(\nabla f_1, \nabla f_2)$ and $(\theta \nabla f_1, \theta \nabla f_2)$ have the same law in μ for each shift $\theta \in \Theta$.

The kernel $\gamma_R^2 = \gamma_R \times \gamma_R$ is simply the kernel from $(\Omega^2, \mathcal{F}_{R^\mathrm{c}}^2)$ to $(\Omega^2, \mathcal{F}^2)$ with the property that the probability measure $(\gamma_R \times \gamma_R)(\cdot, (f_1, f_2))$ is uniformly random in the set $\Omega(R, f_1) \times \Omega(R, f_2)$, and it restricts naturally to the kernel $\gamma_R^{2\nabla} = \gamma_R^\nabla \times \gamma_R^\nabla$—this is a probability kernel from $(\Omega^2, \mathcal{F}_{R^\mathrm{c}}^{2\nabla})$ to $(\Omega^2, \mathcal{F}^{2\nabla})$. A double gradient measure μ is called a *double gradient Gibbs measure* if it satisfies, for every finite $R \subset X^d$, the DLR equation

$$\mu = \mu \gamma_R^{2\nabla}.$$

If μ is the product of two gradient Gibbs measures μ_1 and μ_2, then μ is also Gibbs as

$$\mu = \mu_1 \times \mu_2 = (\mu_1 \gamma_R^\nabla) \times (\mu_2 \gamma_R^\nabla) = (\mu_1 \times \mu_2)(\gamma_R^\nabla \times \gamma_R^\nabla) = \mu \gamma_R^{2\nabla}.$$

Now let $\mu \in \mathcal{P}(\Omega^2, \mathcal{F}^{2\nabla})$ and $R \subset X^d$ finite. The *entropy of μ in R*, denoted $\mathcal{H}^2_R(\mu)$, is defined by

$$\mathcal{H}^2_R(\mu) := \mathcal{H}_{\mathcal{F}^{2\nabla}_R}(\mu, \lambda_2^R).$$

The *specific entropy* of μ, denoted $\mathcal{H}^2(\mu)$, is defined to be the limit

$$\mathcal{H}^2(\mu) := \lim_{n\to\infty} \frac{1}{|\Pi_n|}\mathcal{H}^2_{\Pi_n}(\mu) = \lim_{n\to\infty} \frac{1}{|\Pi_n|}\mathcal{H}_{\mathcal{F}^{2\nabla}_{\Pi_n}}(\mu, \lambda_2^{\Pi_n})$$

whenever the limit is convergent, and the limit inferior otherwise.

The direct generalisation of Theorems 4.9.3 and 4.9.5 to the product setting reads as follows.

Theorem 4.9.6. *Let $\mu \in \mathcal{P}_\Theta(\Omega^2, \mathcal{F}^{2\nabla})$ denote a shift-invariant product measure such that*

$$\lim_{n\to\infty} \frac{1}{n}\|(f_i - f_i(\mathbf{0}))|_{\Pi_n} - s_i|_{\Pi_n}\|_\infty = 0 \tag{4.9.7}$$

almost surely for $i \in \{1,2\}$ and for some fixed slopes $s_1, s_2 \in \mathcal{S}$. Then $\mathcal{H}^2(\mu) \geq \sigma(s_1) + \sigma(s_2)$, with equality if and only if μ is a Gibbs measure.

4.9.3 Proof overview

Fix throughout this section two distinct Lipschitz slopes $s_1, s_2 \in \mathcal{S}$ such that their average $s_a := (s_1 + s_2)/2$ lies in the interior of $\mathcal{S}$. The ultimate goal of this section is to prove that $2\sigma(s_a) < \sigma(s_1) + \sigma(s_2)$, which implies Theorem 4.9.1: that σ is strictly convex on the interior of $\mathcal{S}$.

In the remainder of this section, let μ_i denote the shift-invariant gradient Gibbs measure of Theorem 4.7.2 of slope s_i for each $i \in \{1,2\}$, and fix $\mu := \mu_1 \times \mu_2$. Then μ is a shift-invariant double gradient Gibbs measure. Moreover, μ has the concentration of (4.9.7), and therefore Theorem 4.9.6 implies that $\mathcal{H}^2(\mu) = \sigma(s_1) + \sigma(s_2)$.

The sets $T(f_1)$ and $T(f_2)$, the graph $G_g = (V_g, E_g)$, the g-level sets, the g-boundaries, and the directed graph $(\mathrm{LSD}(g), \nabla g)$ are all invariant under adding constants to f_1 and f_2, as each of them is characterised entirely by the gradients ∇f_1, ∇f_2, and $\nabla g := \nabla f_1 - \nabla f_2$. The gradient ∇g also determines $X_g^\pm(E)$ for any g-boundary E.

Lemma 4.9.8. *It is μ-almost certain that $\mathrm{LSD}(g)$ contains a subgraph that is graph isomorphic to $\mathbb{Z}$. Moreover, every g-level set and every g-boundary involved in such a subgraph of $\mathrm{LSD}(g)$ is unbounded.*

This lemma is essential in understanding the geometry of $\mathrm{LSD}(g)$. It is expected that the difference function $g = f_1 - f_2$ of a typical sample from μ looks somewhat like the leftmost subfigure of Figure 4.4.

Proof of Lemma 4.9.8. As $s_g := s_1 - s_2 \neq 0$, there exists an index $1 \leq i \leq d+1$ such that $s_g(\mathbf{g}_i) \neq 0$. Fix such an i, and write $\mathbf{p}$ for the $\mathbb{Z}$-indexed path $\mathbf{p} := (\mathbf{p}_k)_{k\in\mathbb{Z}} := (k\mathbf{g}_i)_{k\in\mathbb{Z}}$ through (X^d, E^d). Write $\mathbf{q}_k$ for the g-level set containing $\mathbf{p}_k$ for each $k \in \mathbb{Z}$. For each $k \in \mathbb{Z}$ the vertices $\mathbf{p}_k$ and $\mathbf{p}_{k+1}$ are either contained in the same g-level set, or in two distinct neighbouring g-level sets. We consider $\mathbf{q} := (\mathbf{q}_k)_{k\in\mathbb{Z}}$ a walk through $\mathrm{LSD}(g)$ although $\mathbf{q}$ is not a walk in the strict sense: it may visit the same g-level set multiple times in a row. Proposition 4.7.5 says that μ-almost surely

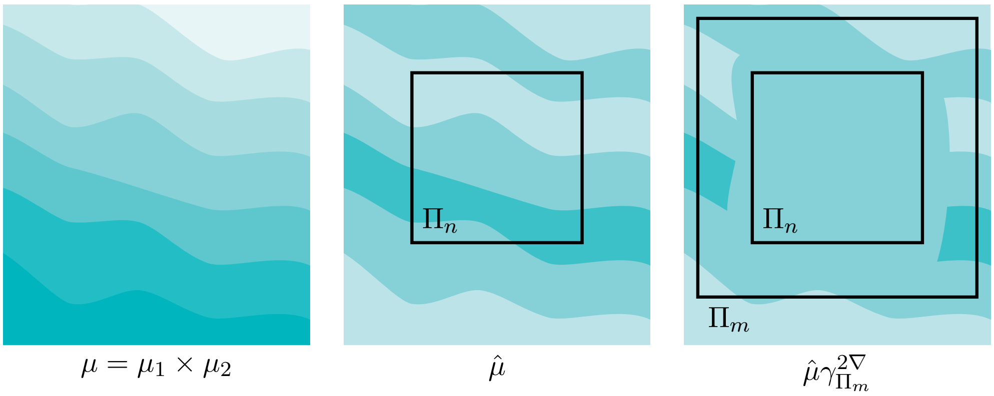


Figure 4.4: Difference samples from different measures; Π_m is a trifurcation box.

$g(\mathbf{p}_k) - g(\mathbf{p}_0) = ks_g(\mathbf{g}_i) + o(k)$ as $k \to \infty$ or $k \to -\infty$. This implies that there is a well-defined and unique loop-erased bi-infinite version of the path $\mathbf{q}$ up to indexation, which is the desired $\mathbb{Z}$-isomorphic subgraph of $\mathrm{LSD}(g)$. This proves the first part of the lemma.

Focus on the second statement, which is deterministic in nature. Fix a g-boundary $E \subset E^d$ that is an edge of a subgraph of $\mathrm{LSD}(g)$ that is isomorphic $\mathbb{Z}$. Then removing E from $\mathrm{LSD}(g)$ disconnects $\mathrm{LSD}(g)$ and separates the graph into two infinite components. In particular, this implies that the graph $(X^d, E^d \smallsetminus E)$ consists of two infinite connected components. If E were finite, then one of the two connected components of $(X^d, E^d \smallsetminus E)$ had to be finite, and therefore we conclude that E is infinite. The g-boundary E connects the two g-level sets $X_g^-(E)$ and $X_g^+(E)$ when considered an $\mathrm{LSD}(g)$-edge, and these must also be infinite as one of them contains the infinite set $\mathbf{x}_g^-(E)$ and the other $\mathbf{x}_g^+(E)$. This proves the second statement of the lemma. $\square$

We now give an overview of the remainder of the proof. The key idea is to construct a new shift-invariant double gradient measure $\hat{\mu} \in \mathcal{P}_\Theta(\Omega^2, \mathcal{F}^{2\nabla})$. Write $\hat{f}_1$, $\hat{f}_2$ and $\hat{g} := \hat{f}_1 - \hat{f}_2$ for the random functions in $\hat{\mu}$. To sample from $\hat{\mu}$, first draw a pair (f_1, f_2) from the original measure $\mu = \mu_1 \times \mu_2$. Then obtain $(\hat{f}_1, \hat{f}_2)$ from (f_1, f_2) by flipping a fair coin for every g-boundary in order to determine whether or not to reverse the orientation of that g-boundary. In other words, we rerandomise the orientation of each g-boundary. In the measure $\hat{\mu}$, the orientations of the edges in the graph $(\mathrm{LSD}(\hat{g}), \nabla\hat{g})$ are thus uniformly random and independent of all other structure that is present. First, we show that the resampling operation does not affect the specific entropy, that is,

$$\mathcal{H}^2(\hat{\mu}) = \mathcal{H}^2(\mu) = \sigma(s_1) + \sigma(s_2).$$

Second, we prove that for $i \in \{1, 2\}$ we $\hat{\mu}$-almost surely have

$$\lim_{n\to\infty} \frac{1}{n} \|(\hat{f}_i - \hat{f}_i(\mathbf{0}))|_{\Pi_n} - s_a|_{\Pi_n}\|_\infty = 0. \tag{4.9.9}$$

Note that the concentration of the gradient of either function is around the average slope s_a. Third, we prove that $\hat{\mu}$ is not Gibbs. Theorem 4.9.6 therefore implies that $\mathcal{H}^2(\hat{\mu}) > 2\sigma(s_a)$, the desired result.

Let us elaborate on the third step, before proceeding. Suppose that $\hat{\mu}$ is Gibbs, in order to derive a contradiction. A *trifurcation box* of $\hat{g}$ is a finite subset of X^d of the form $R = \theta\Pi_n$ such that, for some infinite $\hat{g}$-level set $X \subset X^d$, removing R from X means breaking X into at least three infinite components. We show that $\hat{g}$ has a trifurcation box with positive probability in the measure $\hat{\mu}\gamma^{2\nabla}_{\Pi_m}$ for m sufficiently large: see the middle and rightmost subfigures in Figure 4.4. If $\hat{\mu}$ were Gibbs then $\hat{\mu} = \hat{\mu}\gamma^{2\nabla}_{\Pi_m}$, and therefore a sample $\hat{g}$ from $\hat{\mu}$ has a trifurcation box with positive probability. Trifurcation boxes do almost surely not occur in shift-invariant measures, by a simple geometrical argument described by Burton and Keane in their celebrated paper [4]. This proves that $\hat{\mu}$ is not Gibbs.

4.9.4 Detailed proof

Lemma 4.9.10. *The specific entropy of $\hat{\mu}$ equals the specific entropy of μ.*

Proof. We prove the stronger statement that $\mathcal{H}^2_{\Pi_n}(\hat{\mu}) = \mathcal{H}^2_{\Pi_n}(\mu) + O(n^{d-1})$ as $n \to \infty$. Fix $n \in \mathbb{N}$ large. The measure μ is Gibbs and therefore satisfies the DLR equation

$$\mu = \mu\gamma^{2\nabla}_{\Pi_n}.$$

This implies in particular that the distribution of a sample (f_1, f_2) from μ is invariant under subsequently rerandomising the orientation of each g-boundary that is contained in $E^d(\Pi_n)$. As this is true for all n, we derive that the distribution of the sample is in fact invariant under rerandomising the orientation of any finite g-boundary. Thus, to sample $(\hat{f}_1, \hat{f}_2)$ from $\hat{\mu}$, one may first sample a pair (f_1, f_2) from μ, then rerandomise the orientation of only the g-boundaries which are infinite. Write $\tilde{\mu}$ for the corresponding coupling of μ and $\hat{\mu}$, and write $(f_1, f_2, \hat{f}_1, \hat{f}_2)$ for the random 4-tuple in $\tilde{\mu}$.

Write $\mathcal{H}^2_{A,B}(\tilde{\mu})$ for $\mathcal{H}_{\mathcal{F}^{2\nabla}_A \times \mathcal{F}^{2\nabla}_B}(\tilde{\mu}, \lambda_2^A \times \lambda_2^B)$ for $A, B \subset X^d$ finite and claim that

1. $\mathcal{H}^2_{\Pi_n}(\mu) = \mathcal{H}^2_{\Pi_n, \Pi_n \smallsetminus \Pi_{n-1}}(\tilde{\mu}) + O(n^{d-1})$,
2. $\mathcal{H}^2_{\Pi_n}(\hat{\mu}) = \mathcal{H}^2_{\Pi_n \smallsetminus \Pi_{n-1}, \Pi_n}(\tilde{\mu}) + O(n^{d-1})$,
3. $\mathcal{H}^2_{\Pi_n, \Pi_n \smallsetminus \Pi_{n-1}}(\tilde{\mu}) = \mathcal{H}^2_{\Pi_n, \Pi_n}(\tilde{\mu}) = \mathcal{H}^2_{\Pi_n \smallsetminus \Pi_{n-1}, \Pi_n}(\tilde{\mu})$.

It is clear that these three claims jointly imply the lemma.

Focus on the first claim, and consider thus the measures

$$\mu|_{\mathcal{F}^{2\nabla}_{\Pi_n}}, \qquad \tilde{\mu}|_{\mathcal{F}^{2\nabla}_{\Pi_n} \times \mathcal{F}^{2\nabla}_{\Pi_n \smallsetminus \Pi_{n-1}}}.$$

The first claim is intuitive: the restriction of μ records the values of $D_{\Pi_n}(f_1)$ and $D_{\Pi_n}(f_2)$, and the restriction of $\tilde{\mu}$ records also the values of $D_{\Pi_n \smallsetminus \Pi_{n-1}}(\hat{f}_1)$ and $D_{\Pi_n \smallsetminus \Pi_{n-1}}(\hat{f}_2)$. Informally, the extra information that the restriction of $\tilde{\mu}$ records is of order n^{d-1}, because $\log|\operatorname{Im} D_{\Pi_n \smallsetminus \Pi_{n-1}}|^2 = O(n^{d-1})$. We now formalise this idea. For $x \in (\operatorname{Im} D_{\Pi_n})^2$, we write $\hat{\mu}^x$ for the measure $\tilde{\mu}$ conditioned on the event $\{(D_{\Pi_n} f_1, D_{\Pi_n} f_2) = x\}$ and projected onto the product of the third and fourth component of the product measurable space $(\Omega, \mathcal{F}^\nabla)^4$. Then a standard entropy calculation implies that

$$\mathcal{H}^2_{\Pi_n, \Pi_n \smallsetminus \Pi_{n-1}}(\tilde{\mu}) = \mathcal{H}^2_{\Pi_n}(\mu) + \int \mathcal{H}^2_{\Pi_n \smallsetminus \Pi_{n-1}}(\hat{\mu}^x) d\mu((D_{\Pi_n} f_1, D_{\Pi_n} f_2) = x),$$

see Theorems D.3 and D.13 of [11] or Lemma 2.1.3 of [54]. As in the proof of Theorem 4.9.3, we have

$$|\mathcal{H}^2_{\Pi_n \smallsetminus \Pi_{n-1}}(\hat{\mu}^x)| \leq \log |\operatorname{Im} D_{\Pi_n \smallsetminus \Pi_{n-1}}|^2 = O(n^{d-1}).$$

This proves the first claim. The second claim follows by identical reasoning.

Focus on the third claim, in particular on the equality on the left—the equality on the right shall follow by the same arguments. For the equality on the left it suffices to demonstrate that, with $\tilde{\mu}$-probability one, the tuple

$$(D_{\Pi_n} f_1, D_{\Pi_n} f_2, D_{\Pi_n} \hat{f}_1, D_{\Pi_n} \hat{f}_2) \tag{4.9.11}$$

can be reconstructed almost surely from

$$(D_{\Pi_n} f_1, D_{\Pi_n} f_2, D_{\Pi_n \smallsetminus \Pi_{n-1}} \hat{f}_1, D_{\Pi_n \smallsetminus \Pi_{n-1}} \hat{f}_2). \tag{4.9.12}$$

We know that (f_1, f_2) and $(\hat{f}_1, \hat{f}_2)$ differ by cluster boundary swaps, where all boundaries which are swapped, are infinite. To recover (4.9.11) from (4.9.12), we must therefore understand whether or not each infinite boundary which intersects Π_{n-1}, should be swapped or not. However, as each such boundary is infinite, it must intersect $\Pi_n \smallsetminus \Pi_{n-1}$, and its orientation for $(\hat{f}_1, \hat{f}_2)$ can therefore be read off from (4.9.12). $\square$

Lemma 4.9.13. *Equation 4.9.9 holds true $\hat{\mu}$-almost surely for $i \in \{1, 2\}$.*

Proof. It suffices to prove that $\hat{\mu}$-almost surely

$$\lim_{n\to\infty} \frac{1}{n} \|(\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n} + (\hat{f}_2 - \hat{f}_2(\mathbf{0}))|_{\Pi_n} - 2s_a|_{\Pi_n}\|_\infty = 0, \tag{4.9.14}$$

$$\lim_{n\to\infty} \frac{1}{n} \|(\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n} - (\hat{f}_2 - \hat{f}_2(\mathbf{0}))|_{\Pi_n}\|_\infty = 0. \tag{4.9.15}$$

First focus on (4.9.14). Recall the definition of μ; Proposition 4.7.5 implies that almost surely

$$\lim_{n\to\infty} \frac{1}{n} \|(f_1 - f_1(\mathbf{0}))|_{\Pi_n} + (f_2 - f_2(\mathbf{0}))|_{\Pi_n} - (s_1 + s_2)|_{\Pi_n}\|_\infty = 0.$$

Recall that the sum of two height functions is invariant under a cluster boundary swap, and that $2s_a = s_1 + s_2$. The previous display therefore implies (4.9.14). For (4.9.15) we must show that $\hat{\mu}$-almost surely

$$\lim_{n\to\infty} \frac{1}{n} \|(\hat{g} - \hat{g}(\mathbf{0}))|_{\Pi_n}\|_\infty = 0 \tag{4.9.16}$$

Fix some $\mathbf{x} \in X^d$; we are interested in the distribution of $\hat{g}(\mathbf{x}) - \hat{g}(\mathbf{0})$. The distribution of $\hat{g}(\mathbf{x}) - \hat{g}(\mathbf{0})$ conditional on $\operatorname{LSD}(\hat{g})$ is given by summing $d_{\operatorname{LSD}(\hat{g})}(\mathbf{0}, \mathbf{x}) \leq d_{(X^d, E^d)}(\mathbf{0}, \mathbf{x})$ fair coins each valued $\pm(d+1)$. This follows immediately from the definition of the measure $\hat{\mu}$. Application of the Azuma-Hoeffding inequality yields, for $a \geq 0$ and $\mathbf{x} \neq \mathbf{0}$,

$$\hat{\mu}(|\hat{g}(\mathbf{x}) - \hat{g}(\mathbf{0})| \geq (d+1)a) \leq 2 \exp -\frac{a^2}{2d_{(X^d, E^d)}(\mathbf{0}, \mathbf{x})}.$$

A union bound now implies (4.9.16). $\square$

Lemma 4.9.17. *The double gradient measure $\hat{\mu}$ is not Gibbs.*

Proof. Define the $\mathcal{F}^{2\nabla}$-measurable event

$$I(n) := \left\{ \begin{array}{l} \hat{g} \text{ takes the same value on three distinct infinite } \hat{g}\text{-level} \\ \text{sets that all intersect } \Pi_n \text{, one of which contains } \mathbf{0} \end{array} \right\} \subset \Omega^2,$$

and claim that $\hat{\mu}(I(n)) > 0$ for $n \in \mathbb{N}$ sufficiently large. First, a cluster boundary swap leaves $\mathrm{LSD}(g)$ invariant, and therefore Lemma 4.9.8 holds true for g replaced with $\hat{g}$ and μ with $\hat{\mu}$. Therefore it is $\hat{\mu}$-almost certain that $\mathrm{LSD}(\hat{g})$ contains a $\mathbb{Z}$-indexed self-avoiding walk $\mathbf{p} = (\mathbf{p}_k)_{k\in\mathbb{Z}}$. Let $\mathbf{p}$ be chosen deterministically in terms of $\mathrm{LSD}(\hat{g})$, so that $(\hat{g}(\mathbf{p}_{k+1}) - \hat{g}(\mathbf{p}_k))_{k\in\mathbb{Z}}$ is a sequence of i.i.d. random variables each distributed uniformly in $\pm(d+1)$, independent of $\mathrm{LSD}(\hat{g})$. In particular, the event $\{\hat{g}(\mathbf{p}_{k\pm 2}) - \hat{g}(\mathbf{p}_k) = 0\}$ has probability $\frac{1}{4}$ for each fixed k. As $\hat{\mu}$ is shift-invariant, we may choose $\mathbf{p}$ such that $\hat{\mu}(\mathbf{0} \in \mathbf{p}_0) > 0$. Choose $n \in \mathbb{N}$ sufficiently large such that, conditional on $\{\mathbf{0} \in \mathbf{p}_0\}$, the set Π_n intersects $\mathbf{p}_{\pm 2}$ with positive probability. Note that $\hat{g}(\mathbf{p}_{\pm 2}) - \hat{g}(\mathbf{p}_0) = 0$ with probability $\frac{1}{4}$ independently of the occurrence of both previous events, and therefore the original event $I(n)$ has positive probability. This is the claim. Fix $n \in \mathbb{N}$ such that $\varepsilon := \frac{1}{2}\hat{\mu}(I(n)) > 0$. See the middle display in Figure 4.4 for the level set decomposition of the difference function $\hat{g}$ corresponding to a sample from the event $I(n)$.

If h and h' are real-valued functions defined on two disjoint subsets A and A' of X^d respectively, then write hh' for the unique function on $A \cup A'$ which equals h on A and h' on A'. Next, define for $m \geq n$ the $\mathcal{F}^{2\nabla}$-measurable event

$$\begin{aligned} L(m) &:= \\ &\left\{\text{the function } (\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n}(\hat{f}_2 - \hat{f}_2(\mathbf{0}))|_{\Pi_m^c} \text{ extends to a height function}\right\} \\ &= \left\{\text{the function } (\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n}(\hat{f}_2 - \hat{f}_2(\mathbf{0}))|_{\Pi_m^c} \text{ is Lipschitz}\right\} \subset \Omega^2, \end{aligned}$$

and claim that $\hat{\mu}(L(m)) \to 1$ as $m \to \infty$. Recall that (4.9.9) holds true for $i = 2$ with $\hat{\mu}$-probability one. Therefore it suffices to show that the function

$$(\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n}(\hat{f}_2 - \hat{f}_2(\mathbf{0}))|_{\Pi_m^c}$$

is Lipschitz for m sufficiently large whenever $\hat{f}_2$ satisfies (4.9.9). To see this, write $h_m^\pm$ for the largest and smallest functions in $\Omega(\Pi_m, \hat{f}_2 - \hat{f}_2(\mathbf{0}))$ respectively. In other words, the functions $h_m^\pm$ are the largest and smallest extensions of $(\hat{f}_2 - \hat{f}_2(\mathbf{0}))|_{\Pi_m^c}$ to X^d which are height functions. Note that there exist constants $\alpha^\pm > 0$ such that $h_m^\pm(\mathbf{0}) = \pm\alpha^\pm m + o(m)$ as $m \to \infty$ due to (4.9.9) and because s_a is in the *interior* of the set of Lipschitz slopes. In particular,

$$h_m^+|_{\Pi_n} \geq (\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n}, \qquad h_m^-|_{\Pi_n} \leq (\hat{f}_1 - \hat{f}_1(\mathbf{0}))|_{\Pi_n}$$

for m sufficiently large. This proves the claim. Fix m so large that $\hat{\mu}(L(m)) \geq 1 - \varepsilon$, which implies that $\hat{\mu}(I(n) \cap L(m)) \geq \varepsilon > 0$.

Define for $\mathbf{x} \in X^d$ the $\mathcal{F}^{2\nabla}$-measurable event

$$\begin{aligned} T(\mathbf{x}) &:= \left\{ \begin{array}{l} (X^d, E^d \smallsetminus (V_{\hat{g}} \cup E^d(\Pi_m + \mathbf{x}))) \text{ has three infinite connected} \\ \text{components that are contained in a single } \hat{g}\text{-level set} \end{array} \right\} \\ &= \left\{ \begin{array}{l} (X^d, E^d \smallsetminus (V_{\hat{g}} \cup E^d(\Pi_m + \mathbf{x}))) \text{ has three infinite connected} \\ \text{components that are contained in one connected compo-} \\ \text{nent of } (X^d, E^d \smallsetminus V_{\hat{g}}) \end{array} \right\} \subset \Omega^2; \end{aligned}$$

we shall first focus on $T := T(\mathbf{0})$. See the rightmost subfigure in Figure 4.4 for the level set decomposition of the difference function $\hat{g}$ corresponding to a sample from T. We claim that $\Omega(\Pi_m, f_1) \times \Omega(\Pi_m, f_2)$ intersects T whenever $(f_1, f_2) \in I(n) \cap L(m)$. Fix such a pair (f_1, f_2) and assume, without loss of generality, that $f_1(\mathbf{0}) = f_2(\mathbf{0}) = 0$. Assert first that there exists a height function f'' which equals f_1 on Π_n and which equals f_2 on Π_m^{c}, and which at any vertex $\mathbf{x}$ takes values between $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$. The following construction works: write first f' for a height function which extends $f_1|_{\Pi_n} f_2|_{\Pi_m^{\mathsf{c}}}$ to X^d, then define f'' by

$$f'' := (f' \vee (f_1 \wedge f_2)) \wedge (f_1 \vee f_2).$$

This proves the assertion. Now $f'' \in \Omega(\Pi_m, f_2)$ as f'' equals f_2 on Π_m^{c}. Moreover, as at each vertex the value of f'' is in between the values of f_1 and f_2, we have $\{f_1 = f_2\} \subset \{f_1 = f''\} \subset X^d$, and we also know that $\Pi_n \subset \{f_1 = f''\}$. This implies that $(f_1, f'') \in T$, since the three connected components of $\{f_1 = f_2\}$ are contained in a single connected component of $\{f_1 = f''\}$, and because $\{f_1 = f_2\} \smallsetminus \Pi_m = \{f_1 = f''\} \smallsetminus \Pi_m$, that is, removing Π_m disconnects these three components again.

Suppose that $\hat{\mu}$ is a Gibbs measure, in order to derive a contradiction. Then the event T must occur with positive probability for $\hat{\mu}$, since $\hat{\mu}(I(n) \cap L(m)) > 0$, and since T has positive probability for

$$\gamma^{2\nabla}_{\Pi_m}(\cdot, (\hat{f}_1, \hat{f}_2))$$

whenever $(\hat{f}_1, \hat{f}_2) \in I(n) \cap L(m)$. The argument of Burton and Keane [4] dictates that trifurcation boxes do almost surely not occur in shift-invariant probability measures. In other words, $\hat{\mu}(T) = 0$. This is the desired contradiction. To see that $\hat{\mu}(T) = 0$, observe that $\hat{\mu}(T(\mathbf{x}))$ is independent of $\mathbf{x}$, implying that trifurcation boxes must occur with positive density whenever $\hat{\mu}(T) > 0$. This is impossible due to geometrical constraints of the amenable graph (X^d, E^d). $\square$

This establishes Theorem 4.9.1.

Bibliography

[1] Stefan Adams, Roman Kotecký, and Stefan Müller, *Strict convexity of the surface tension for non-convex potentials*, arXiv preprint arXiv:1606.09541v1 (2016).

[2] R. J. Baxter, *Exactly solved models in statistical mechanics*, Academic Press, 1982.

[3] J. Bricmont and A. Kupiainen, *Phase transition in the 3d random field ising model*, Comm. Math. Phys. **116** (1988), no. 4, 539–572.

[4] R. M. Burton and M. Keane, *Density and uniqueness in percolation*, Comm. Math. Phys. **121** (1989), no. 3, 501–505.

[5] Henry Cohn, Richard Kenyon, and James Propp, *A variational principle for domino tilings*, J. Amer. Math. Soc. **14** (2001), no. 2, 297–346.

[6] Codina Cotar and Jean-Dominique Deuschel, *Decay of covariances, uniqueness of ergodic component and scaling limit for a class of $\nabla\phi$ systems with non-convex potential*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), no. 3, 819–853.

[7] Codina Cotar, Jean-Dominique Deuschel, and Stefan Müller, *Strict convexity of the free energy for a class of non-convex gradient models*, Comm. Math. Phys. **286** (2009), no. 1, 359–376.

[8] Nilanjana Datta, *Min-and max-relative entropies and a new entanglement monotone*, IEEE Trans. Inform. Theory **55** (2009), no. 6, 2816–2826.

[9] Daniela De Silva and Ovidiu Savin, *Minimizers of convex functionals arising in random surfaces*, Duke Math. J. **151** (2010), no. 3, 487–532.

[10] A. Dembo and T. Funaki, *Lectures on probability theory and statistics*, Lecture Notes in Mathematics, vol. 1869, Springer-Verlag, 2005, Lectures from the 33rd Probability Summer School held in Saint-Flour, July 6–23, 2003, Edited by Jean Picard.

[11] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, 2010.

[12] Jean-Dominique Deuschel and Daniel W. Stroock, *Large deviations*, Pure and Applied Mathematics, vol. 137, Academic Press, 1989.

[13] Robert G. Edwards and Alan D. Sokal, *Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988), no. 6, 2009–2012.

[14] Aernout C. D. van Enter, Roberto Fernández, and Alan D. Sokal, *Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory*, J. Stat. Phys. **72** (1993), no. 5, 879–1167.

[15] Aernout van Enter, Christian Maes, Roberto H. Schonmann, and Senya Shlosman, *The Griffiths singularity random field*, American Mathematical Society Translations (R. A. Minlos, Senya Shlosman, and Yu. M. Suhov, eds.), 2, vol. 198, American Mathematical Society, 2000, pp. 51–58.

[16] Aernout van Enter and Evgeny Verbitskiy, *On the variational principle for generalized Gibbs measures*, arXiv preprint math-ph/0410052v1 (2004).

[17] Roberto Fernández, *Gibbsianness and non-Gibbsianness in lattice random fields*, Mathematical Statistical Physics: Lecture Notes of the Les Houches Summer School 2005 (Anton Bovier, François Dunlop, Aernout van Enter, Frank Den Hollander, and Jean Dalibard, eds.), Elsevier Science, 2006, pp. 731–798.

[18] Roberto Fernández, Arnaud Le Ny, and Frank Redig, *Variational principle and almost quasilocality for renormalized measures*, J. Statist. Phys. **111** (2003), no. 1-2, 465–478.

[19] T. Funaki and H. Spohn, *Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model*, Comm. Math. Phys. **185** (1997), no. 1, 1–36.

[20] Hans-Otto Georgii, *Gibbs measures and phase transitions*, De Gruyter Studies in Mathematics, vol. 9, Walter de Gruyter, 2 ed., 2011.

[21] Alessandro Giuliani, Vieri Mastropietro, and Fabio Lucio Toninelli, *Height fluctuations in interacting dimers*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 98–168.

[22] ———, *Non-integrable dimers: Universal fluctuations of tilted height profiles*, arXiv preprint arXiv:1904.07526v3 (2019).

[23] H. H. Goldstine, *A history of the calculus of variations from the 17th through the 19th century*, Studies in the History of Mathematics and Physical Sciences, vol. 5, Springer-Verlag, 1980.

[24] Geoffrey Grimmett, *The random-cluster model*, Grundlehren der mathematischen Wissenschaften, vol. 333, Springer-Verlag, 2006.

[25] ———, *Probability on graphs: random processes on graphs and lattices*, Institute of Mathematical Statistics Textbooks, vol. 8, Cambridge University Press, 2 ed., 2018.

[26] Geoffrey R. Grimmett and Zhongyang Li, *Locality of connective constants*, Discrete Math. **341** (2018), no. 12, 3483–3497.

[27] Piotr Hajłasz and Jan Malý, *Approximation in Sobolev spaces of nonlinear expressions involving the gradient*, Ark. Mat. **40** (2002), no. 2, 245–274.

[28] P. W. Kasteleyn, *The statistics of dimers on a lattice: I. the number of dimer arrangements on a quadratic lattice*, Physica **27** (1961), no. 12, 1209–1225.

[29] ———, *Dimer statistics and phase transitions*, J. Math. Phys. **4** (1963), no. 2, 287–293.

[30] ———, *Graph theory and crystal physics*, Graph Theory and Theoretical Physics (Frank Harary, ed.), Academic Press, London, 2004, pp. 43–110.

[31] Richard Kenyon, *Conformal invariance of loops in the double-dimer model*, Commun. Math. Phys. **326** (2014), no. 2, 477–497.

[32] ———, *Limit shapes and their analytic parameterizations*, Proc. Int. Cong. Math. (2018), no. 3, 3123–3138.

[33] Richard Kenyon and Andrei Okounkov, *Limit shapes and the complex Burgers equation*, Acta Math. **199** (2007), no. 2, 263–302.

[34] Richard Kenyon, Andrei Okounkov, and Scott Sheffield, *Dimers and amoebae*, Ann. Math. **163** (2006), no. 3, 1019–1056.

[35] Richard Kenyon and Robin Pemantle, *Double-dimers, the Ising model and the hexahedron recurrence*, J. Comb. Theory, Ser. A **137** (2016), 27–63.

[36] Andrew Krieger, Georg Menz, and Martin Tassy, *Deducing a variational principle with minimal a priori assumptions*, arXiv preprint arXiv:1910.01042v2 (2019).

[37] Christof Külske, Arnaud Le Ny, and Frank Redig, *Relative entropy and variational properties of generalized Gibbsian measures*, Ann. Probab. **32** (2004), no. 2, 1691–1726.

[38] R. Lefevere, *Variational principle for some renormalized measures*, J. Statist. Phys. **96** (1999), no. 1-2, 109–133.

[39] J. T. Lewis, C.-E. Pfister, and W. G. Sullivan, *Entropy, concentration of probability and conditional limit theorems*, Markov Process. Relat. **1** (1995), no. 3, 319–386.

[40] Elliott H. Lieb, *Residual entropy of square ice*, Phys. Rev. **162** (1967), no. 1, 162–172.

[41] Joakim Linde, Cristopher Moore, and Mats G. Nordahl, *An n-dimensional generalization of the rhombus tiling*, Discrete Math. Theor. Comput. Sci. Proc. **AA (DM-CCG)** (2001), 23–42.

[42] C. Maes, F. Redig, and A. Van Moffaert, *Almost Gibbsian versus weakly Gibbsian measures*, Stochastic Process. Appl. **79** (1999), no. 1, 1–15.

[43] ———, *The restriction of the Ising model to a layer*, J. Statist. Phys. **96** (1999), no. 1-2, 69–107.

[44] Georg Menz and Martin Tassy, *A variational principle for a non-integrable model*, Probab. Theory Related Fields (2020).

[45] Lars Onsager, *Crystal statistics. I. A two-dimensional model with an order-disorder transition*, Phys. Rev. (2) **65** (1944), no. 3–4, 117–149.

[46] Ron Peled and Yinon Spinka, *Lectures on the spin and loop $O(n)$ models*, arXiv preprint arXiv:1708.00058v1 (2017).

[47] Charles-Edouard Pfister, *Thermodynamical aspects of classical lattice systems*, In and Out of Equilibrium: Probability with a Physics Flavor (Vladas Sidoravicius, ed.), Birkhäuser Boston, 2002, pp. 393–472.

[48] Charles-Edouard Pfister and Koen Vande Velde, *Almost sure quasilocality in the random cluster model*, J. Stat. Phys. **79** (1995), no. 3, 765–774.

[49] Firas Rassoul-Agha and Timo Seppäläinen, *A course on large deviations with an introduction to Gibbs measures*, Graduate Studies in Mathematics, vol. 162, American Mathematical Society, 2015.

[50] Timo Seppäläinen, *Large deviations for lattice systems. I. Parametrized independent fields*, Probab. Theory Related Fields **96** (1993), no. 2, 241–260.

[51] ———, *Large deviations for lattice systems. II. Nonstationary independent fields*, Probab. Theory Related Fields **97** (1993), no. 1-2, 103–112.

[52] ———, *Entropy, limit theorems, and variational principles for disordered lattice systems*, Comm. Math. Phys. **171** (1995), no. 2, 233–277.

[53] ———, *Entropy for translation-invariant random-cluster measures*, Ann. Probab. **26** (1998), no. 3, 1139–1178.

[54] Scott Sheffield, *Random surfaces*, Astérisque **304** (2005).

[55] Ananth Sridhar, *Limit shapes in the six vertex model*, Ph.D. thesis, University of California, Berkeley, 2016.

[56] Daniel W. Stroock and Ofer Zeitouni, *Microcanonical distributions, Gibbs states, and the equivalence of ensembles*, Random Walks, Brownian Motion, and Interacting Particle Systems (Rick Durrett and Harry Kesten, eds.), Birkhäuser Boston, 1991, pp. 399–424.

[57] Bill Sutherland, *Exact solution of a two-dimensional model for hydrogen-bonded crystals*, Phys. Rev. Lett. **19** (1967), no. 3, 103–104.

[58] Robert H. Swendsen and Jian-Sheng Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett. **58** (1987), no. 2, 86–88.

[59] H. N. V. Temperley and Michael E. Fisher, *Dimer problem in statistical mechanics—An exact result*, Philos. Mag. **6** (1961), no. 68, 1061–1063.

[60] Evgeny Verbitskiy, *Variational principle for fuzzy Gibbs measures*, Mosc. Math. J. **10** (2010), no. 4, 811–829.

[61] C. N. Yang, *The spontaneous magnetization of a two-dimensional Ising model*, Phys. Rev. (2) **85** (1952), no. 5, 808–816.

[62] C. P. Yang, *Exact solution of a model of two-dimensional ferroelectrics in an arbitrary external electric field*, Phys. Rev. Lett. **19** (1967), no. 10, 586–588.

[63] Boguslav Zegarlinski, *Interactions and pressure functionals for disordered lattice systems*, Comm. Math. Phys. **139** (1991), no. 2, 305–339.