Ruhr-Universität Bochum
Fakultät für Mathematik

Stochastic Processes on Trees

Christof Külske

September 28, 2017

Foreword

The following notes are lecture notes for a course held in Essen in the summer term 2016, in the context of the research training group RTG 2131. The material presented evolves around the concept of Gibbs measures on infinite graph-theoretic trees and the related concept of a tree-indexed Markov chain. The presentation is self-contained, when proofs are omitted, references are given.

In Section 1 we begin with a motivating elementary example of broadcasting on trees, and the associated reconstruction problem. It is elementary, as it is formulated without resorting to the infinite-volume formalism which will be introduced later, but the relation to the Ising model, and two types of non-trivial transitions at different critical temperatures already appear, with proofs to be given later.

Section 2 contains relevant concepts and results of the infinite-volume Gibbsian theory on general countable graphs, including trees, but allowing also for more general graphs than trees: definition of Gibbs measures, extremality, tail-triviality, Dobrushin uniqueness, and uniqueness in one dimension. Much of the presentation of this general material follows the exposition which can be found in the excellent book of Georgii on Gibbs measures.

In Section 3 we narrow down the supporting graphs to trees and discuss the concept of a boundary law, which is a useful concept very particular to trees. We have tried to give a detailed and motivating presentation, including many pictures. Examples presented are the Ising and the Potts model, for which we discuss all branches of tree-invariant Markov chains.

Section 4 comes back to the reconstruction problem which was introduced when we discussed broadcasting on trees, now formulated in terms of extremality of a given Gibbs measure. Bounds ensuring non-extremality (the Kesten-Stigum bound in the language of multi-color branching processes), and bounds ensuring extremality are both discussed.

Section 5 deals with tree-indexed models having non-compact local spin space. In such a situation infinite-volume Gibbs measures may cease to exist, but the more general concept of a gradient Gibbs measures is still meaningful. We discuss without proof, with a somewhat algebraic flavor, the relation between boundary laws with periodicity and gradient Gibbs measures.

After reading Section 2 and Section 3, Sections 4 or Sections 5 can be read independently. I am much indebted to my coauthors Aernout van Enter, Marco Formentin (Section 4), Utkir Rozikov (Section 3), and Philipp Schriever (Section 5). Philipp Schriever's Ph.D. thesis evolved around the topics covered in the present note, and I would like to say a particular thank you for all of his work during his Ph.D. thesis and in the preparation of the present manuscript. I also thank Florian Henning for a critical reading of the manuscript.

Contents

1 Introduction

1.1 Motivating Example: Broadcasting on Trees

Definition 1.1.1. *A tree(-graph) is a graph (V, E), where V is a countable or finite vertex set and E is the edge set, which has no loops and is connected. Notation: $x \sim y \Leftrightarrow x$ is adapted to $y :\Leftrightarrow \{x, y\} \in E \Leftrightarrow \{x, y\}$ is an edge.*

Note that for all vertices $x, y \in V$ there exists a unique self-avoiding path $x = x_0 \sim x_1 \sim \cdots \sim x_n = y$ of neighboring vertices $x_1, ..., x_n \in V$. The path length $n =: d(x, y)$ defines the tree metric.

Let $T_N = (V_N, E_N)$ be the binary tree of depth N, rooted at 0. Define the finite-volume state of the model to be $\Omega_N := \{-1, +1\}^{V_N}$. Together with the product sigma-algebra (which is simply the power set on our finite set) this will be our probability space.

Denote by $\sigma \in \Omega_N$ a configuration. We define a probability distribution μ on Ω_N in two steps. First, the distribution of the spin at the origin is chosen to be symmetric Bernoulli, i.e.

$$\mu(\sigma_0 = 1) = \mu(\sigma_0 = -1) = \frac{1}{2}.$$

Next we pass on information from the root to the outside of the tree by putting, for all pairs of neighboring vertices $v \to w$ (meaning that v is the parent of w, i.e. v is closer to the root than w),

$$\begin{aligned} \mu(\sigma_w = -1 \mid \sigma_v = 1) &= \varepsilon \\ \mu(\sigma_w = 1 \mid \sigma_v = -1) &= \varepsilon. \end{aligned} \tag{1.1.1}$$

Here $\varepsilon \in [0, \frac{1}{2}]$ is an error parameter, and the full probability distribution is obtained by applying this rule from the root to the outside of the tree. In this way we have the following probability distribution on our finite-volume state space.

Definition 1.1.2. *The probability measure defined by*

$$\begin{aligned} \mu(\sigma) &= \frac{1}{2} \prod_{v,w:v \to w} (1-\varepsilon)^{\mathbf{1}_{\sigma_v = \sigma_w}} \varepsilon^{\mathbf{1}_{\sigma_v \neq \sigma_w}} \\ &= \frac{1}{2} \prod_{v,w:v \to w} P_{\sigma_v, \sigma_w} \end{aligned} \tag{1.1.2}$$

with

$$P = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}$$

is called the symmetric channel *on the binary tree.*

This is a specific example of a *tree-indexed Markov chain.* This is an important concept which will be defined in more generality and investigated in the next chapters. We can imagine to replace P by another transition matrix to obtain a different distribution, and we can generalize the local state space. Note that $\varepsilon = 1/2 \Leftrightarrow \sigma_v$'s are independent.

Using simple calculations with ± 1-valued variables we can put our probability measure in the exponential form

$$\begin{aligned} \mu(\sigma) &= \frac{1}{2} \prod_{v,w: v \mapsto w} (1-\varepsilon)^{\mathbf{1}_{\sigma_v = \sigma_w}} \varepsilon^{\mathbf{1}_{\sigma_v \neq \sigma_w}} \\ &= \frac{\exp[\beta \sum_{v,w:v\to w} \sigma_v \sigma_w]}{Z_N(\beta)} \end{aligned} \tag{1.1.3}$$

with $\beta := \frac{1}{2} \log \frac{1-\varepsilon}{\varepsilon}$, called the inverse temperature in statistical mechanics, or equivalently $\varepsilon = \frac{1}{e^{2\beta}+1}$. Here

$$Z_N(\beta) = \sum_{\sigma \in \Omega_N} \exp\left[\beta \sum_{v,w \in V, v \sim w} \sigma_v \sigma_w \right] \tag{1.1.4}$$

is a normalizing constant, called partition function in statistical mechanics, that makes (1.1.3) a probability measure. We have recovered here the *(finite-volume) Gibbs measure* for the *Ising model on a tree* (with open boundary conditions).

We would like to understand this measure. In which way is possibly information preserved over long distances? Such questions will set the tone for subsequent investigations.

We write, for a vertex $v \in V$, $|v|$ for the distance to the origin. For $|w| = N$ we define

$$\mu(\sigma_0 \sigma_w = -1) =: F(N).$$

This is a meaningful quantity for all N, so we may take a limit. We start with the following simple one-dimensional observation.

Proposition 1.1.3. *We have*

$$\lim_{N\to\infty} F(N) = \frac{1}{2},$$

i.e., an observation of a single spin at the boundary at distance N does not allow us to deduce anything about the state at the root when N tends to infinity.

Proof. The problem is reduced to the the study of a Markov chain along the path which connects the root 0 to the vertex w. Such a problem is elementary and can be treated by diagonalization: With the transition matrix

$$P = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}.$$

we get

$$\begin{aligned} F(N) &= \sum_{\sigma_1,\dots,\sigma_{N-1}} P_{1,\sigma_1} P_{\sigma_1,\sigma_2} \dots P_{\sigma_{N-1},-1} \\ &= P^N(1,-1). \end{aligned} \tag{1.1.5}$$

Now diagonalize: The matrix P has eigenvalue 1, with eigenvector $(1,1)^T$, and eigenvalue $1-2\varepsilon$, with eigenvector $(1,-1)^T$. Hence we have

$$O^T P O = \begin{pmatrix} 1 & 0 \\ 0 & 1-2\varepsilon \end{pmatrix}$$

with

$$O = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = O^T,$$

and we arrive at

$$\begin{aligned} P^N &= O \begin{pmatrix} 1 & 0 \\ 0 & (1-2\varepsilon)^N \end{pmatrix} O \\ &= \frac{1}{2} \begin{pmatrix} 1+(1-2\varepsilon)^N & 1-(1-2\varepsilon)^N \\ 1-(1-2\varepsilon)^N & 1+(1-2\varepsilon)^N \end{pmatrix}. \end{aligned} \tag{1.1.6}$$

Hence, $\lim_{N\to\infty} F(N) = \lim_{N\to\infty} \frac{1}{2}(1-(1-2\varepsilon)^N) = \frac{1}{2}$. □

In the proof we have seen that the rate of convergence is governed by the second largest eigenvalue (or the spectral gap) of the transition matrix M. For smaller second largest eigenvalue $1-2\varepsilon$ we have faster convergence.

That was not surprising: In general one dimensional models have no long range order, unless the interactions are long-range. Markov chains on finite state spaces loose their memory exponentially fast.

A more interesting question now is the following, and this is a typical tree-question. When does the information at *all of the boundary sites* allow us to deduce the state at the origin? The chances are much better now, as there are exponentially many sites in N, and the boundary sites constitute a non-vanishing fraction of all sites of a tree of depth N. Write $\partial T_N = \{v \in V : |v| = N\}$ for the boundary of the tree of depth N. Consider the conditional probability that the variable at the origin is 1, if we condition on any configuration at distance N from the origin, that is

$$\pi_N(\xi) = \mu(\sigma(0) = 1 | \sigma_{\partial T_N} = \xi) \tag{1.1.7}$$

Question 1.1.4. *Is it always true that a conditioning of the boundary spins to take their maximal value has no predictive power for the spin at the origin, for large volumes? That is, do we have*

$$\lim_{N\uparrow\infty} \pi_N(+1_{\partial T_N}) = \frac{1}{2} \ ? \tag{1.1.8}$$

Does it depend on the value of the error parameter ε?

Exercise 1.1.5. *Prove that taking the expectation over the boundary spins yields* $\mathrm{E}_\mu \, \pi_N(\xi) = \frac{1}{2}$ *for all N.*

Question 1.1.6. *Is it true that*

$$\lim_{N\uparrow\infty} \pi_N(\xi) = \frac{1}{2}, \tag{1.1.9}$$

for typical realizations *of the boundary spins ξ? For which values of ε?*

What do we mean by typicality? More precisely, let us consider the variance of the random variable π_N obtained feeding it random boundary spins ξ distributed according to the measure μ itself. The second question reformulated reads then: When do we have

$$\lim_{N\uparrow\infty} \mathrm{var}_N(\pi^N) = \lim_{N\uparrow\infty} \mathrm{E}_\mu \left[\left(\pi^N - \frac{1}{2} \right)^2 \right] = 0 \ ? \tag{1.1.10}$$

Exercise 1.1.7. *Is this statement weaker or stronger or equivalent compared to stochastic convergence?*

The above questions have answers as follows.

Theorem 1.1.8. *Let T be a regular tree where every vertex has precisely d children. Then* (1.1.8) *holds iff* $d \tanh\beta \leq 1$.

Theorem 1.1.9. *Let T be a regular tree where every vertex has precisely d children. Then* (1.1.10) *holds iff* $d(\tanh\beta)^2 \leq 1$.

Note that $\tanh \beta$ is the second largest eigenvalue of the transition matrix M. This is no accident. The second largest eigenvalue will also appear in analogous statements for more general models.

In both cases we see critial values for β (or equivalently the error probability ε) appearing. We see examples of threshold phenomena, or phase transitions at a finite critical value. For readers with a background of statistical mechanics of lattice models it may seem surprising that these critical values are different. This is particular to things happening on trees, as we shall see later. It is however in accordance with intuition that the value of the parameter β (which can be considered as a coupling strength) needs to be bigger to ensure propagation of a typical boundary condition.

The questions above have been formulated in a pedestrian way, in the sense that we made statements in terms of limits of finite-volume quantities. We did not need any measure theory yet. However, the appropriate setting to discuss them is the formalism of infinite-volume Gibbs measures to which will come now.

2 Gibbsian theory on countable vertex sets: Basic definitions and results

In this chapter we will give a short introduction to the general theory of Gibbs measures in the infinite volume and present some basic definitions and results that will be used throughout. We have to omit many of the proofs to keep this course reasonably short, but they can be found in the refences below. The theory presented incorporates lattice models and models on trees, and does not use any particular geometric properties of the index set. Gibbs measures in the infinite volume are commonly defined in terms of a consistency relation, the so-called DLR relation. It expresses the invariance of the Gibbs measure under a whole family of probability kernels. This family, often denoted by γ, is called a specification, as it specifies the Gibbs measure, which may however very well be non-unique. This possibility for non-uniqueness is very relevant for the theory, when non-uniqueness occurs, we speak of a phase transition.

Specifications are often taken from physics, and in such a case they come often in an exponential form, with a Hamiltonian which is given in terms of an interaction potential Φ. An example for this is the Ising model. The Gibbs measures for a given specification form a convex set, or more precisely a simplex (in which each element has a unique representation in terms of extremals). We will describe the connection between extremality, triviality on the tail-sigma algebra (which contains all events that do not depend on finitely many spins), and correlation decay. Tail-triviality will become relevant again later when we talk about the reconstruction problem (we touched upon in the introduction) for general finite-state space models on trees.

A more pedestrian approach to Gibbs measures which has been used historically is to look at limits of finite-volume Gibbs measures with fixed boundary conditions. In the general theory these are nothing but the specification kernels with fixed boundary condition. This approach is closely connected to the more sophisticated DLR-approach: Each extremal Gibbs measure can be obtained as a finite volume limit, for typical boundary condition.

For compact local state spaces, each quasilocal specification has at least one Gibbs measure. To decide whether it has only one, can be difficult. Uniqueness holds in situations of weak dependence, and a very useful method to ensure uniqueness is provided by the so-called Dobrushin-uniqueness theory. We give a full proof of the relevant uniqueness result. We also present and prove a useful comparision theorem which allows to control the difference between Gibbs measures for two specifications, viewed on local observables.

We conclude this section by explaining that uniqueness holds very generally for one-dimensional models under "finite energy condition", including all finite range models.

For more thorough and detailed exhibitions of the theory, see e.g., [3], [5], [6], [11].

2.1 Gibbsian specifications

Let V be a countably infinite set and let Ω_0 be a Polish space with sigma-algebra $\mathcal{F}_0$. We call Ω_0 the local state space, the simplest non-trivial example we have just seen in the introduction is $\Omega_0 = \{-1, 1\}$.

V could appear naturally as the vertex set of some graph, e.g., $V = \mathbb{Z}^d$ with $d \in \mathbb{N}_*$ (the positive integers). In subsequent chapters we will assume V to be the vertex set of an infinite tree. For any sub-volume $\Lambda \subset V$ (possibly infinite) define

$$\Omega_\Lambda = \Omega_0^\Lambda = \{(\omega_x)_{x\in\Lambda} : \omega_x \in \Omega_0 \ \forall x \in \Lambda\}.$$

When $\Lambda = V$, we write $\Omega = \Omega_V$. The measurable structure on Ω_Λ is given by the product sigma-algebra

$$\mathcal{B}_\Lambda = \bigotimes_{i\in\Lambda} \mathcal{F}_0 =: \mathcal{F}_0^\Lambda.$$

For any $x \in V$ the projection onto the x'th coordinate will be denoted by

$$\begin{aligned} \sigma_x : \Omega &\to \Omega_0, \\ \omega &\mapsto \omega_x. \end{aligned}$$

The restriction of a configuration in the infinite volume, $\omega \in \Omega$, to a sub-volume $\Lambda \subset V$, can be given by using the projection $\sigma_\Lambda : \Omega \to \Omega_\Lambda$ with $\sigma_\Lambda(\omega) = \omega_\Lambda$. Similarly, if $\Lambda \subset \Delta \subset V$ we will use the same notation σ_Λ for the projection from Ω_0^Δ to Ω_0^Λ. The concatenation of two configurations $\omega \in \Omega_0^\Lambda$ and $\rho \in \Omega_0^{\Delta\setminus\Lambda}$ is denoted as $\omega\rho \in \Omega_0^\Delta$ and is defined by having the properties $\sigma_\Lambda(\omega\rho) = \omega$ and $\sigma_{\Delta\setminus\Lambda}(\omega\rho) = \rho$.

Let us agree to write $\Lambda \Subset V$ if Λ is a finite subset of V. For any finite sub-volume $\Lambda \Subset V$ we define the sigma-algebra of *cylinders with base in* Λ as

$$\mathcal{C}(\Lambda) := \sigma_\Lambda^{-1}(\mathcal{B}_\Lambda).$$

For any (possibly infinite) $\Delta \subset V$, consider the algebra of cylinders with base in Δ, i.e.,

$$\mathcal{C}_\Delta := \bigcup_{\Lambda \Subset \Delta} \mathcal{C}(\Lambda).$$

For each $\Delta \subset V$ the sigma-algebra $\mathcal{F}_\Delta$ "of all events occurring in Δ" is then by definition generated by $\mathcal{C}(\Delta)$, i.e.,

$$\mathcal{F}_\Delta := \sigma(\mathcal{C}_\Delta).$$

When $\Delta = V$ we simply write $\mathcal{F} = \mathcal{F}_V$ and we have by our previous definition that $\mathcal{F}$ is the smallest sigma-algebra on Ω containing the cylinder events, i.e., $\mathcal{F} = \sigma(\mathcal{C}_V) = \bigotimes_{i \in V} \mathcal{F}_0$.

A *spin model* is then simply a probability measure on the product space $(\Omega, \mathcal{F})$. We call Ω_0 the *local state space* and Ω the *configuration space* of the spin model. In the following we will work towards the introduction of equilibrium states of a spin model in the infinite volume. First, we need some more definitions and notations.

Definition 2.1.1. *Let $\Lambda \Subset V$. A* probability kernel *from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$ is a map $\pi_\Lambda : \mathcal{F} \times \Omega \to [0,1]$ with the properties*

(i) $\pi_\Lambda(\cdot \mid \omega)$ is a probability measure on $(\Omega, \mathcal{F})$ for any $\omega \in \Omega$,

(ii) $\pi_\Lambda(A \mid \cdot)$ is $\mathcal{F}_{\Lambda^c}$-measurable for each $A \in \mathcal{F}$.

If

$$\pi_\Lambda(A \mid \omega) = \mathbf{1}_A(\omega), \quad \forall A \in \mathcal{F}_{\Lambda^c}$$

for all $\omega \in \Omega$, π_Λ is called proper*.*

A probability kernel pulls functions back and pushes measures forward in the following sense: If μ is a probability measure on the measurable space $(\Omega, \mathcal{F}_{\Lambda^c})$ and π_Λ is a probability kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$ then

$$\mu\pi_\Lambda(A) = \int \pi_\Lambda(A \mid \omega)\mu(d\omega), \qquad \forall A \in \mathcal{F}$$

defines a probability measure on $(\Omega, \mathcal{F})$. Also, if $f : \Omega \to \mathbb{R}$ is measurable w.r.t. $\mathcal{F}$ then the function $\pi_\Lambda f : \Omega \to \mathbb{R}$ given by

$$\pi_\Lambda f(\omega) = \pi_\Lambda(f|\omega) = \int \pi_\Lambda(d\zeta \mid \omega) f(\zeta), \qquad \forall \omega \in \Omega$$

is measurable w.r.t. $\mathcal{F}_{\Lambda^c}$.

The *composition of kernels* π_Λ and π_Δ is defined by the formula

$$\pi_\Lambda\pi_\Delta(A \mid \omega) := \int \pi_\Delta(A \mid \rho)\pi_\Lambda(d\rho \mid \omega)$$

for any $A \in \mathcal{F}$ and all $\omega \in \Omega$ and is itself a kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$.

Let us assume that we have a proper probability kernel π_Λ from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$. Then the probability measure $\pi_\Lambda(\cdot \mid \omega)$ will be supported on the set $\Omega_\Lambda^\omega := \sigma_{\Lambda^c}^{-1}(\omega)$ for any $\omega \in \Omega$ as

$$\pi_\Lambda(\Omega_\Lambda^\omega \mid \omega) = \mathbf{1}_{\Omega_\Lambda^\omega}(\omega) = 1$$

since $\Omega_\Lambda^\omega \in \mathcal{F}_{\Lambda^c}$. Therefore one can interpret the configuration $\omega \in \Omega$ as the *boundary condition* of the measure $\pi_\Lambda(\cdot \mid \omega)$. In the following all kernels π_Λ to be considered will be proper and therefore they will be entirely determined by all the numbers $\pi_\Lambda(\eta_\Lambda\omega_{\Lambda^c}|\omega)$. We will sometimes use the shorter notation $\pi_\Lambda(\eta_\Lambda\omega_{\Lambda^c}|\omega) = \pi_\Lambda(\eta_\Lambda|\omega_{\Lambda^c})$.

As it turns out, it will be necessary to use an infinite family of probability kernels, $\{\pi_\Lambda\}_{\Lambda \Subset V}$, to describe Gibbs measures in the infinite volume directly. The key concept in that regard is that of a (local) specification:

Definition 2.1.2. *A* specification *is a family of proper probability kernels $\gamma = \{\gamma_\Lambda\}_{\Lambda \Subset V}$ from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$ which satisfies the consistency relation, i.e.,*

$$\gamma_\Delta \gamma_\Lambda = \gamma_\Delta$$

for any finite subsets $\Lambda, \Delta \Subset V$ with $\Lambda \subset \Delta$. A measure $\mu \in \mathcal{M}_1(\Omega)$ (the probability measures on Ω) is said to be compatible with *(or* specified by*) γ if*

$$\mu = \mu\gamma_\Lambda \quad \forall \Lambda \Subset V.$$

The set of measures which are compatible with γ will be denoted by $\mathcal{G}(\gamma)$.

A first natural question that arises with regard to this definition is if there is a way to construct specifications. Before we answer this question we want to discuss the following lemma [6, Remark 1.20].

Lemma 2.1.3. *Suppose π_Λ is a proper probability kernel from $\mathcal{F}_{\Lambda^c}$ to $\mathcal{F}$.*

(i) We have that

$$\pi_\Lambda(A \cap B \mid \cdot) = \pi_\Lambda(A \mid \cdot)\mathbf{1}_B(\cdot)$$

for all $A \in \mathcal{F}$ and all $B \in \mathcal{F}_{\Lambda^c}$.

(ii) Let $\mu \in \mathcal{M}_1(\Omega)$. Then $\mu\pi_\Lambda = \pi_\Lambda$ if and only if

$$\mu(A \mid \mathcal{F}_{\Lambda^c}) = \pi_\Lambda(A \mid \cdot) \quad \mu\text{-a.s.} \tag{2.1.1}$$

for all $A \in \mathcal{F}$.

Proof. (i): First assume that $\omega \notin B$. Then

$$\pi_\Lambda(A \cap B \mid \omega) \leq \pi_\Lambda(B \mid \omega) = \mathbf{1}_B(\omega) = 0.$$

Now suppose $\omega \in B$. We have

$$\pi_\Lambda(A \cap B \mid \omega) = \pi_\Lambda(A \mid \omega) - \pi_\Lambda(A \cap B^c \mid \omega) = \pi_\Lambda(A \mid \omega).$$

(ii): If (2.1.1) holds, then for all $\Lambda \Subset V$ and all $A \in \mathcal{F}$,

$$\mu\pi_\Lambda(A) = \int \pi_\Lambda(A \mid \omega)\mu(d\omega) = \int \mu(A \mid \mathcal{F}_{\Lambda^c})(\omega)\,\mu(d\omega) = \mu(A).$$

Now suppose that $\mu\pi_\Lambda = \mu$. Hence

$$\mu(A \cap B) = \mu\pi_\Lambda(A \cap B) = \int \pi_\Lambda(A \cap B \mid \omega)\mu(d\omega) = \int_B \pi_\Lambda(A \mid \omega)\mu(d\omega)$$

for all $A \in \mathcal{F}$ and all $B \in \mathcal{F}_{\Lambda^c}$. By the definition of the conditional probability we have

$$\mu(A \cap B) = \int_B \mu(A \mid \mathcal{F}_{\Lambda^c})(\omega)\mu(d\omega) \quad \forall B \in \mathcal{F}_{\Lambda^c}$$

and by the almost sure uniqueness of the conditional expectation we see that

$$\mu(A \mid \mathcal{F}_{\Lambda^c})(\cdot) = \pi_\Lambda(A \mid \cdot)$$

μ-almost surely for all $A \in \mathcal{F}$. □

The second part of Lemma 2.1.3 tells us that for a given specification $(\gamma_\Lambda)_{\Lambda \Subset V}$ the measures $\mu \in \mathcal{G}(\gamma)$ are characterized by having a regular conditional distribution provided by γ_Λ, when conditioning with respect to $\mathcal{F}_{\Lambda^c}$.

The most important class of specifications are the so-called *Gibbsian specifications* which we will introduce in the following definition.

Definition 2.1.4. *Let $\Phi = \{\Phi_\Lambda\}_{\Lambda \Subset V}$ be a family of real-valued functions on the configuration space Ω. We call Φ an* interaction potential *if it has the following properties:*

(i) The functions Φ_Λ are $\mathcal{F}_\Lambda$-measurable for any $\Lambda \Subset V$.

(ii) For all $\Lambda \Subset V$ and $\omega \in \Omega$, the series

$$H_\Lambda^\Phi(\omega) = \sum_{A \Subset V, A \cap \Lambda \neq \emptyset} \Phi_\Lambda(\omega) \tag{2.1.2}$$

exists.

We call H_Λ^Φ *the* Hamiltonian *in the finite sub-volume* Λ *associated to the potential* Φ.

By "existence of the series" we mean that for any increasing sequence of volumes Δ_n which converges to V we have that

$$\lim_{n\uparrow\infty} \sum_{A\subset\Delta_n, A\cap\Lambda\neq\emptyset} \Phi_\Lambda(\omega) \tag{2.1.3}$$

exists and does not depend on the volume sequence.

Since the sum (2.1.2) contains possibly infinitely many terms there is no guarantee that it converges. However, for an important class of interaction potentials this is not an issue: Let $d_G(\cdot,\cdot)$ denote the graph distance on V, which is the number of edges in the shortest path connecting two vertices. We define the *diameter* of a finite set Λ by $\mathrm{diam}(\Lambda) := \sup_{x,y\in\Lambda} d_G(x,y)$. Let

$$r(\Phi) := \inf\{R > 0 : \Phi_\Lambda = 0 \text{ for all } \Lambda \text{ with } \mathrm{diam}(\Lambda) > R\}.$$

If $r(\Phi) < \infty$, the interaction potential Φ is said to be of *finite range* and clearly the Hamiltonian H_Λ^Φ is well defined for any finite sub-volume Λ in this case. Later in this lecture notes we will only consider potentials that are of finite range.

In the following we will assume that the local state space is equipped with a so-called *a priori measure* $\lambda \in \mathcal{M}_1(\Omega_0)$ and denote for any $\Lambda \subset V$ the product measure on $(\Omega_0^\Lambda, \mathcal{F}_0^\Lambda)$ by λ^Λ. The *(conditional) partition function* is then defined by

$$Z_\Lambda^\Phi(\omega) = \int e^{-H_\Lambda^\Phi(\zeta_\Lambda\omega_{\Lambda^c})} \lambda^\Lambda(d\zeta_\Lambda).$$

A potential Φ is said to be λ-*admissible* if the partition function $Z_\Lambda^\Phi(\omega)$ is a finite number in the open interval $(0,\infty)$, for all $\Lambda \Subset V$ and all $\omega \in \Omega$.

Proposition 2.1.5. *Suppose that* Φ *is a* λ*-admissible interaction potential. Then the family of probability kernels* $\gamma^\Phi = \{\gamma_\Lambda^\Phi\}_{\Lambda\Subset V}$ *from* $\mathcal{F}_{\Lambda^c}$ *to* $\mathcal{F}$ *defined by*

$$\gamma_\Lambda^\Phi(A \mid \omega) = \frac{1}{Z_\Lambda^\Phi(\omega)} \int e^{-H_\Lambda^\Phi(\zeta_\Lambda\omega_{\Lambda^c})} \mathbf{1}_A(\zeta_\Lambda\omega_{\Lambda^c}) \lambda^\Lambda(d\zeta_\Lambda)$$

constitutes a specification and it is called the Gibbs specification *for* Φ. *A probability measure* $\mu \in \mathcal{G}(\gamma^\Phi)$ *is called an* infinite-volume Gibbs measure *(or simply a* Gibbs measure*) associated to the potential* Φ.

To verify the specifaction properties note that the measurability properties are evident, while the consistency is obtained by a rearrangement of sums, see [6, Proposition 2.5]. The measures $\gamma^\Phi(\cdot \mid \omega) \in \mathcal{M}_1(\Omega)$ are also called *finite-volume Gibbs measures under boundary condition* ω. The way we have defined them they actually are measures on the infinite volume. However, recall that they are supported on the set Ω_Λ^ω which consists only of configurations that are equal to ω outside the *finite* volume Λ.

2.2 Extremal Gibbs measures

One basic observation is that as the DLR equation is linear, $\mathcal{G}(\gamma)$ is a convex set: If $\mu_1, ..., \mu_N$ belong to $\mathcal{G}(\gamma)$, then so does any convex combination of them. This makes the extremal elements of this set, which we will denote by $\mathrm{ex}\,\mathcal{G}(\gamma)$, especially interesting. The following questions arise naturally:

1. What properties, if any, distinguish the elements of $\mathrm{ex}\,\mathcal{G}(\gamma)$ from the non-extremal ones?

2. What is the physical interpretation of these extremal points of $\mathcal{G}(\gamma)$?

Before we answer these questions we will give a condition under which the set of extremal Gibbs measures is non-empty. Let $\mathcal{C}_b(\Omega)$ denote the set of bounded real-valued functions on Ω which are continuous w.r.t. the product topology obtained from the topology on the Polish local state space Ω_0. A particular class of specifications is given in the following definition:

Definition 2.2.1. *A specification $\gamma = (\gamma_\Lambda)_{\Lambda \Subset V}$ is said to be* Feller-continuous *if, for each $\Lambda \Subset V$, $f \in \mathcal{C}_b(\Omega)$ implies $\gamma_\Lambda f \in \mathcal{C}_b(\Omega)$.*

An important example of Feller-continuous specifications is provided by the Gibbsian specifications γ^Φ where the interaction potential Φ is continuous and *uniformly convergent* (and λ-admissible) [3]. An interaction is by definition uniformly convergent if for every $\Lambda \Subset V$ the sum in (2.1.2) converges uniformly in ω. Note that this is always the case if the interaction is of finite range which will be the case for all models considered in these notes.

Let $(\Lambda_n)_{n\in\mathbb{N}}$ be any sequence of finite subsets of V. We say that $(\Lambda_n)_{n\in\mathbb{N}}$ *exhausts* V if for every $v \in V$ there exists a $N \in \mathbb{N}$ such that $v \in \Lambda_n$ for every $n \geq N$.

Proposition 2.2.2. *[3, Proposition 2.22] Suppose γ is a Feller-continuous specification and let $(\Lambda_n)_{n\in\mathbb{N}}$ be any sequence of finite subsets of V that* exhausts V. *Let $\nu_n \in \mathcal{M}_1(\Omega)$ be any sequence of measures. If $\nu_n\gamma_{\Lambda_n}$ converges weakly to some $\mu \in \mathcal{M}_1(\Omega)$, then $\mu \in \mathcal{G}(\gamma)$.*

We like to note the following fact: If Ω_0 is compact, so is $\Omega = \Omega_0^V$ w.r.t. the product topology. Also, Ω is Polish since it is the countable product of Polish spaces. Hence, $\mathcal{M}_1(\Omega)$ is weakly compact. Therefore, in the case of a Feller-continuous specification every sequence $\nu_n\gamma_{\Lambda_n}$ has a convergent subsequence and hence $\mathcal{G}(\gamma)$ is not empty.

In general this might not be true; the question of whether or not $|\mathcal{G}(\gamma)| = 0$ is a non-trivial one. There indeed exist physically reasonable models for which there are no infinite-volume Gibbs measures. Examples are the massless discrete Gaussian free field on the lattice $\mathbb{Z}^d$ in dimensions $d \leq 2$ and the solid-on-solid model in $d = 1$ [6]. In both cases the local state space equals the set of all integers.

One nice property of Feller-continuous specifications is that they allow the identification of Gibbs measures as weak limits, at least the extremals. To be more specific, we have the following statement [3, Proposition 2.23]:

Proposition 2.2.3. *Let Ω_0 be a compact metric space and let $(\gamma_\Lambda)_{\Lambda \Subset V}$ be a Feller-continuous specification. Furthermore let μ be an element of $\operatorname{ex}\mathcal{G}(\gamma)$. Then, for μ-a.e. ω*

$$\lim_{n\to\infty} \gamma_{\Lambda_n}(\cdot \mid \omega) = \mu$$

in the weak limit for any sequence of finite sub-volumes $(\Lambda_n)_{n\in\mathbb{N}}$ that exhausts V.

Let us assume a Feller-continuous specification is given. Then the previous two propositions show the connection between the DLR-approach to the Gibbs theory in infinite-volume and the classical approach using the thermodynamic limit of finite-volume Gibbs measures under boundary condition (see e.g., Chapter 3 of [5] for a detailed exhibition of this ansatz). Proposition 2.2.2 tells us that any weak limit of finite-volume Gibbs measures is in fact an infinite-volume Gibbs measure. Conversely, Proposition 2.2.3 states that if we have an extremal Gibbs measure μ and sample any typical configuration from μ and use it as a boundary condition, in the infinite-volume limit we will recover μ itself.

The following theorem follows immediately from Proposition 2.2.3 and gives a condition for which there is a unique Gibbs measure:

Theorem 2.2.4. *Let Ω_0 be a compact metric space and let $(\gamma_\Lambda)_{\Lambda \Subset V}$ be a Feller-continuous specification. Suppose that for all sequences of finite sub-volumes $(\Lambda_n)_{n\in\mathbb{N}}$ exhausting V and every $\omega \in \Omega$ all the possible weak limits of $\gamma_{\Lambda_n}(\cdot \mid \omega)$ are identical. Then there exists exactly one Gibbs measure.*

Recall that for any $\Lambda \Subset V$ we defined $\mathcal{F}_{\Lambda^c}$ as the sigma-algebra which consists of all the events that only depend on the spins outside the finite set Λ. Now the *tail sigma-algebra* (or *tail field*) $\mathcal{T}$ is defined as the sigma-algebra which only depends on the spins outside *any* finite region Λ, i.e.,

$$\mathcal{T} := \bigcap_{\Lambda \Subset V} \mathcal{F}_{\Lambda^c}.$$

The extremal elements of $\mathcal{G}(\gamma)$ are characterized by the following properties [3, Proposition 2.20]:

Proposition 2.2.5. *Let $\mu \in \mathcal{G}(\gamma)$. Then the following statements are equivalent:*

(i) The measure μ is an extremal element of $\mathcal{G}(\gamma)$.

(ii) The measure μ is trivial *on the tail sigma-algebra $\mathcal{T}$, i.e.,*

$$\mu(A) \in \{0,1\}$$

for every $A \in \mathcal{T}$.

(iii) The measure μ has short-range correlations*, i.e., for each $A \in \mathcal{F}$ we have*

$$\lim_{\Lambda \uparrow V,\, \Lambda \Subset V} \sup_{B \in \mathcal{F}_{\Lambda^c}} |\mu(A \cap B) - \mu(A)\mu(B)| = 0.$$

Let us comment on the statement of the last proposition. Physical systems can in general have one or more possible "macrostates", depending on the values of some internal free parameters of the system. For example water can be in a gaseous, liquid or solid "macrostate" depending on the temperature and pressure. While the microscopic quantities change rapidly, the macroscopic quantities remain constant. To turn this into a mathematical exact statement we define the macroscopic quantities or *macroscopic observables* as the functions on Ω that are measurable w.r.t. the tail field $\mathcal{T}$, i.e., the functions that do not depend on spins in any finite volume $\Lambda \Subset V$. The physical relevance of the preceding theory presented in this chapter lies in the assumption that the statistical mechanical information of the physical system can be obtained from a suitable specification γ, that is, the space of measures $\mathcal{G}(\gamma)$ describes the "macrostates" of the system. Proposition 2.2.5 tells us that these "macrostates" are given by the extremal elements of $\mathcal{G}(\gamma)$.

What is then the interpretation of the *non-extremal* elements of $\mathcal{G}(\gamma)$? Suppose that the local state space $(\Omega_0, \mathcal{F}_0)$ is Polish. Then every non-extremal measure $\mu \in \mathcal{G}(\gamma)$ is an (integral) convex combination of extremal ones. This decomposition is even unique, that is, $\mathcal{G}(\gamma)$ is a simplex [6, Theorem 7.26].

This means that a non-extremal Gibbs measure corresponds simply to the preparation of randomly chosen extremal Gibbs measures. The probabilities for this choice are given by the "coefficients" of the convex combination. This extra randomness can be interpreted as the uncertainty in the experiment regarding the true nature of the systems "macrostate" (for a more detailed discussion, see Chapter 6 of [5]).

Therefore the non-extremal Gibbs measures do not lead to new physics: Everything that we can observe under such a measure is typical for one of the extremal ones that appear in its (unique) decomposition. Hence the extremal Gibbs measures are the physically important ones, which is why they are also called the "pure" states. This is the reason why we say that a physical system exhibits a *phase transition* when there exist multiple extremal Gibbs measures for the model.

Finally, we want to point out that the extremal Gibbs measures are suitable to describe the different phases of the system as it is possible to distinguish those measures by looking at macroscopic observables only. This is important since we should be able to tell "macrostates" apart by looking at macroscopic measurements:

Theorem 2.2.6. *[6, Theorem 7.7] Let μ_1, μ_2 be two distinct extremal Gibbs measures w.r.t. a specification $(\gamma_\Lambda)_{\Lambda \Subset V}$. Then there exists some tail-measurable event $A \in \mathcal{T}$ such that $\mu_1(A) = 1$ and $\mu_2(A) = 0$, i.e., μ_1 and μ_2 are mutually singular.*

2.3 Uniqueness

2.3.1 Dobrushin's condition for uniqueness

In this section we will give a criterion for the uniqueness of Gibbs measures on any graph. To do this we will need to formulate a weak dependence requirement on the specification γ. First, let us recall the notion of *total variational distance* on the space of probability measures $\mathcal{M}_1(\Omega)$. (For the sake of this definition we allow Ω to be any measurable space.) The total variational distance is defined as

$$d_{TV}(\alpha_1, \alpha_2) := \sup_{A \in \mathcal{F}} |\alpha_1(A) - \alpha_2(A)|, \quad \alpha_1, \alpha_2 \in \mathcal{M}_1(\Omega).$$

If the state space Ω is discrete we have

$$d_{TV}(\alpha_1, \alpha_2) = \frac{1}{2} \sum_{\omega \in \Omega} |\alpha_1(\omega) - \alpha_2(\omega)|.$$

Look at the single-site kernels γ_i and define

$$\sup_{\omega_{\{i,j\}^c}, \bar{\omega}_j, \hat{\omega}_j} \|\gamma_i(d\omega_i \mid \omega_{\{i,j\}^c} \bar{\omega}_j) - \gamma_i(d\omega_i \mid \omega_{\{i,j\}^c} \hat{\omega}_j)\|_{TV} =: C_{ij} \in [0,1].$$

We call $(C_{ij})_{i,j \in V}$ the *Dobrushin matrix*. In this definition the kernels are compared in total variation only on the local state space at i (and not in the whole infinite-volume). Note that $C_{ii} = 0$ for all $i \in V$. The entry C_{ij} in a sense measures the dependence of the observation at site i when perturbing the spin at site j.

Definition 2.3.1. *A function $f : \Omega \to \mathbb{R}$ is called* local *if it is $\mathcal{F}_\Lambda$-measurable for some finite $\Lambda \subset V$, i.e., if f only depends on finitely many coordinates.*

A function $f : \Omega \to \mathbb{R}$ is called quasilocal *if there exists a sequence of local functions f_n s.t. $\sup_{\omega\in\Omega} |f_n(\omega) - f(\omega)| \to 0$ for $n \to \infty$.*

A specification γ is called quasilocal if $\omega \mapsto \gamma_\Lambda(f \mid \omega)$ is quasilocal for any finite $\Lambda \subset V$ and any quasilocal function f.

The set of local functions is denoted by $\mathcal{L}$ and the set of quasilocal functions is denoted by $\bar{\mathcal{L}}$.

For an example look at the Ising model on any graph V. If $W \Subset V$ is a finite subset $\omega \mapsto \sum_{i,j\in W: i\sim j} \omega_i\omega_j$ is $\mathcal{F}_W$-measurable.

Theorem 2.3.2. *Let $\gamma = (\gamma_\Lambda)_{\Lambda\Subset V}$ be a quasilocal specification on $(\Omega, \mathcal{F})$. Assume that γ satisfies the so-called* Dobrushin condition*:*

$$c := \sup_{i\in V} \sum_{j\in V\setminus\{i\}} C_{ij} < 1.$$

Then $|\mathcal{G}(\gamma)| \leq 1$.

Before we turn to the proof of this uniqueness theorem we need to introduce some more notation. The *variation* of $f \in \bar{\mathcal{L}}$ *at the site* i is defined by

$$\delta_i(f) := \sup_{\omega_{i^c},\omega_i,\bar{\omega}_i} |f(\omega_i\omega_{i^c}) - f(\bar{\omega}_i\omega_{i^c})|.$$

Let $\mu_1, \mu_2 \in \mathcal{M}_1(\Omega, \mathcal{F})$. An infinite vector $a = (a_i)_{i\in V} \in [0, \infty]^V$ is called an *estimate for* μ_1, μ_2 if

$$|\mu_1(f) - \mu_2(f)| \leq \sum_{i\in V} a_i\delta_i(f), \quad \forall f \in \bar{\mathcal{L}}.$$

Note: The constant vector $a = 1$ is always an estimate (small exercise).

Lemma 2.3.3 (Georgii 8.18)**.** *Suppose we have two specifications $\gamma, \tilde{\gamma}$ on $(\Omega, \mathcal{F})$ and $\mu \in \mathcal{G}(\gamma)$ and $\tilde{\mu} \in \mathcal{G}(\tilde{\gamma})$. Furthermore, suppose the "closeness-condition",*

$$\|\gamma_i(\cdot \mid \omega_{i^c}) - \tilde{\gamma}_i(\cdot \mid \omega_{i^c})\|_{TV} \leq b_i(\omega_{i^c}) < \infty,$$

holds for every $\omega \in \Omega$ and γ is a quasilocal specification with Dobrushin matrix $(C_{ij}(\gamma))_{i,j\in V}$. Then: If $a = (a_i)_{i\in V}$ is an estimate for $\mu, \tilde{\mu}$, then also $\bar{a} = (\bar{a}_i)_{i\in V}$ with

$$\bar{a}_i := \sum_{j\in V\setminus\{i\}} C_{ij}(\gamma)a_j + \tilde{\mu}(b_i), \quad i \in V,$$

is an estimate for $\mu, \tilde{\mu}$.

Proof. Take $\Lambda \Subset V$ and put

$$a_i^\Lambda = \begin{cases} \bar{a}_i \wedge a_i, & \text{if } i \in \Lambda \\ a_i, & \text{if } i \notin \Lambda \end{cases}.$$

We claim that a^Λ is an estimate for every $\Lambda \Subset V$.

Proof by induction: This is clear for $\Lambda = \emptyset$.
Induction step: $\Lambda \to \Lambda \cup \{i\}$ where $i \notin \Lambda$. Take $f \in \mathcal{L}$. Then we may add crossterms to get the upper bound

$$|\mu(f)-\tilde{\mu}(f)| = |\mu(\gamma_i(f))-\tilde{\mu}(\tilde{\gamma}_i(f))| \leq |\mu(\gamma_i(f))-\tilde{\mu}(\gamma_i(f))|+|\tilde{\mu}(\gamma_i(f))-\tilde{\mu}(\tilde{\gamma}_i(f))|$$

because $\mu \in \mathcal{G}(\gamma)$ and $\tilde{\mu} \in \mathcal{G}(\tilde{\gamma})$. We estimate each term on the right separately. For the second term we look at the integrand and observe, using the single-site closeness condition,

$$\begin{aligned} |\gamma_i(f)(\omega) - \tilde{\gamma}_i(f)(\omega)| &= |\int \gamma_i(d\tilde{\omega}_i \mid \omega_{i^c}) f(\tilde{\omega}_i\omega_{i^c}) - \int \tilde{\gamma}_i(d\tilde{\omega}_i \mid \omega_{i^c}) f(\tilde{\omega}_i\omega_{i^c})| \\ &\leq ||\gamma_i(\cdot \mid \omega_{i^c}) - \tilde{\gamma}_i(\cdot \mid \omega_{i^c})||_{TV}\delta_i(f) \\ &\leq b_i(\omega)\delta_i(f). \end{aligned} \tag{2.3.1}$$

Hence, after the $\tilde{\mu}$-integration the second term is at most $\tilde{\mu}(b_i)\delta_i(f)$.

To estimate the first term we use the hypotheses that $\gamma_i f \in \bar{\mathcal{L}}$ and a^Λ is an estimate. This gives us

$$|\mu(\gamma_i(f)) - \tilde{\mu}(\gamma_i f)| \leq \sum_{j\in V} a_j^\Lambda \delta_j(\gamma_i(f)).$$

As $\gamma_i(f)$ does not depend on the spin at site i we have $\delta_i(\gamma_i(f)) = 0$. Hence take any $j \neq i$. Then, adding cross-terms,

$$\begin{aligned} &|\gamma_i(f \mid \omega_{\{i,j\}^c}\omega_j) - \gamma_i(f \mid \omega_{\{i,j\}^c}\eta_j)| \\ \leq &|\gamma_i(f(\cdot_i\omega_{\{i,j\}^c}\omega_j) \mid \omega_{\{i,j\}^c}\omega_j) - \gamma_i(f(\cdot_i\omega_{\{i,j\}^c}\eta_j) \mid \omega_{\{i,j\}^c}\omega_j)| \\ &+ |\gamma_i(f(\cdot_i\omega_{\{i,j\}^c}\eta_j) \mid \omega_{\{i,j\}^c}\omega_j) - \gamma_i(f(\cdot_i\omega_{\{i,j\}^c}\eta_j) \mid \omega_{\{i,j\}^c}\eta_j)|. \end{aligned} \tag{2.3.2}$$

The first summand on the r.h.s. is bounded from above by $\delta_j(f)$ and the second summand by

$$\delta_i(f(\cdot_i\omega_{\{i,j\}^c}\eta_j))||\gamma_i(\cdot_i \mid \omega_{\{i,j\}^c}\omega_j)-\gamma_i(\cdot_i \mid \omega_{\{i,j\}^c}\eta_j)|| \leq \delta_i(f(\cdot_i\omega_{\{i,j\}^c}\eta_j))C_{ij}(\gamma).$$

This shows that

$$\delta_j(\gamma_i(f)) \leq \delta_j(f) + C_{ij}(\gamma)\delta_i(f).$$

Combining all preceding inequalities leads to

$$\begin{aligned}
|\mu(f) - \tilde{\mu}(f)| &\le \sum_{j(\neq i)} a_j^\Lambda [\delta_j(f) + C_{ij}(\gamma)\delta_i(f)] + \tilde{\mu}(b_i)\delta_i(f) \\
&= \sum_{j\neq i} a_j^\Lambda \delta_j(f) + \delta_i(f) \left(\sum_{j\neq i} a_j^\Lambda C_{ij}(\gamma) + \tilde{\mu}(b_i) \right) \\
&\le \sum_{j\neq i} a_j^\Lambda \delta_j(f) + \delta_i(f) \left(\sum_{j\neq i} a_j C_{ij}(\gamma) + \tilde{\mu}(b_i) \right) \\
&= \sum_{j\neq i} a_j^\Lambda \delta_j(f) + \bar{a}_i \delta_i(f).
\end{aligned} \tag{2.3.3}$$

In the second inequality we have used that $a_j^\Lambda \le a_j$ and in the third that $C_{ii}(\gamma) = 0$ for all $i \in V$ and the definition of $\bar{a}_i$. As a^Λ is an estimate the last inequality also holds when we replace $\bar{a}_i$ by a_i and so we may also replace it by $\bar{a}_i \wedge a_i$. This shows that $a^{\Lambda\cup\{i\}}$ is an estimate. This finishes the proof. $\square$

For a given specification γ and any integer $n \ge 0$ we let

$$C^n(\gamma) = (C_{ij}^n(\gamma))_{i,j\in V}$$

denote the n'th matrix power of the interaction matrix $C(\gamma)$. We put

$$D(\gamma) = (D_{ij}(\gamma))_{i,j\in V} = \sum_{n\ge 0} C^n(\gamma).$$

Theorem 2.3.4. *Let γ and $\tilde{\gamma}$ be two specifications. Suppose γ satisfies Dobrushin's condition. For each $i \in V$ we let b_i be a measurable function on Ω s.t.*

$$||\gamma_i(\cdot_i \mid \omega) - \tilde{\gamma}_i(\cdot_i \mid \omega)|| \le b_i(\omega)$$

for all $\omega \in \Omega$. If $\mu \in \mathcal{G}(\gamma)$ and $\tilde{\mu} \in \mathcal{G}(\tilde{\gamma})$ then

$$|\mu(f) - \tilde{\mu}(f)| \le \sum_{i,j\in V} \delta_i(f) D_{ij}(\gamma)\tilde{\mu}(b_j)$$

for all $f \in \bar{\mathcal{L}}$.

Proof. We write $C = C(\gamma), D = D(\gamma)$, and $\tilde{b} = (\tilde{\mu}(b_i))_{i\in V}$. Replacing b_i by $1 \wedge b_i$ if necessary we may assume that $\tilde{b}_i \le 1$ for all $i \in V$. We need to show that the vector $D\tilde{b}$ is an estimate for μ and $\tilde{\mu}$. Consider the estimate $a_i = 1$ for all i. A repeated application of Lemma 2.3.3 shows that

$$a^{(n)} = C^n a + \sum_{k=0}^{n-1} C^k \tilde{b}$$

is an estimate for every $n \in \mathbb{N}$. In the following we will show that $a^{(n)} \to D\tilde{b}$ coordinatewise as $n \to \infty$.

Dobrushin's condition for γ implies that, with the Dobrushin constant denoted by $c = \sup_i \sum_j C_{ij}$, we have

$$||C(\gamma)^n||_\infty \leq ||C(\gamma)||_\infty^n = c^n \to 0, \text{ with the matrix norm } ||A||_\infty = \sup_{i \in V} |\sum_{j \in V} A_{ij}|$$

and therefore

$$||\sum_{k=0}^{n-1} C(\gamma)^k||_\infty \leq \sum_{k=0}^{n-1} c^k \leq \frac{1}{1-c}.$$

Hence, the row sums of D are at most $1/(1-c)$. In particular, D has finite entries and $D\tilde{b}$ exists. Also,

$$(C^n a)_i = \sum_{j \in V} (C^n)_{ij} \to 0$$

coordinatewise as $n \to \infty$. Thus $a^{(n)} \to D\tilde{b}$. As the coordinatewise limit of estimates is also an estimate (see [6, Remark 8.17 (3)]), this finishes the proof. □

The proof of Theorem 2.3.2 is a consequence of Theorem 2.3.4, since we can put $\gamma = \tilde{\gamma}$ and $b_i = 0$ for all $i \in V$. This gives us that $\mu(f) = \tilde{\mu}(f)$ for all $f \in \mathcal{L}$ and thus $\mu = \tilde{\mu}$ whenever $\mu, \tilde{\mu} \in \mathcal{G}(\gamma)$.

2.3.2 Uniqueness in one dimension

Proposition 2.3.5. *Let γ be a specification on $(\Omega, \mathcal{F})$. Suppose there exists a constant $c > 0$ s.t. for all cylinder sets A there exists a finite $\Lambda \Subset V$ s.t.*

$$\gamma_\Lambda(A \mid \omega_{\Lambda^c}) \geq c\gamma_\Lambda(A \mid \bar{\omega}_{\Lambda^c})$$

for all $\omega, \bar{\omega} \in \Omega$. Then $|\mathcal{G}(\gamma)| \leq 1$.

Note that we ask the constant to be uniform in A. The constant should be the same for arbitrarily large but finite number of coordinates the event A may depend on. This is not to be expected for spatial models in more than one dimension, but it is very appropriate in one dimension, as the following example illustrates.

Example 2.3.6 (Ising model on $\mathbb{Z}$)**.** *The state space is given as* $\Omega = \{-1, +1\}^{\mathbb{Z}}$*. Take any cylinder* $A \in \mathcal{F}_{\{-k,-k+1,...,l\}}$*. We choose* $\Lambda = \{-k, -k+1, ..., l\}$*. Then*

$$\begin{aligned}\gamma_\Lambda(A \mid \omega_{\Lambda^c}) &= \frac{\sum_{\sigma_\Lambda \in A} \exp[\beta \sum_{i,j\in\Lambda: i\sim j} \sigma_i\sigma_j + \beta\sigma_{-k}\omega_{-k-1} + \beta\sigma_l\omega_{l+1}]}{\sum_{\sigma_\Lambda \in \Omega_0^\Lambda} \exp[\beta \sum_{i,j\in\Lambda: i\sim j} \sigma_i\sigma_j + \beta\sigma_{-k}\omega_{-k-1} + \beta\sigma_l\omega_{l+1}]} \\ &\geq e^{-4\beta} \frac{\sum_{\sigma_\Lambda \in A} \exp[\beta \sum_{i,j\in\Lambda: i\sim j} \sigma_i\sigma_j]}{\sum_{\sigma_\Lambda \in \Omega_0^\Lambda} \exp[\beta \sum_{i,j\in\Lambda: i\sim j} \sigma_i\sigma_j]}\end{aligned} \tag{2.3.4}$$

and we get a similar upper bound. Hence,

$$\gamma_\Lambda(A \mid \omega_{\Lambda^c}) \geq e^{-8\beta} \gamma_\Lambda(A \mid \bar{\omega}_{\Lambda^c})$$

for all $\omega, \bar{\omega} \in \Omega$ *and the proposition applies. Similar computations can be done to show uniqueness for many one-dimensional models.*

Proof of Proposition 2.3.5. Let the Gibbs measure $\mu \in \mathcal{G}(\gamma)$ be arbitrary and $B \in \mathcal{T} = \bigcap_{\Lambda \Subset V} \mathcal{F}_{\Lambda^c}$ a tail event. We want to show that under μ the event B has either probability 0 or 1. This means that μ is trivial on $\mathcal{T}$ which is equivalent to μ being extremal (Proposition 2.2.5).

Assume that $\mu(B) > 0$ and define $\nu := \mu(\cdot \mid B) = \frac{\mu(\cdot \cap B)}{\mu(B)}$. Observe that $\nu \in \mathcal{G}(\gamma)$ as we have for any event A that

$$\nu(\gamma_\Lambda(\mathbf{1}_A)) = \frac{\mu(\mathbf{1}_B \gamma_\Lambda(\mathbf{1}_A))}{\mu(\mathbf{1}_B)} = \frac{\mu(\gamma_\Lambda(\mathbf{1}_A \mathbf{1}_B))}{\mu(\mathbf{1}_B)} = \frac{\mu(\mathbf{1}_A \mathbf{1}_B)}{\mu(\mathbf{1}_B)} = \nu(\mathbf{1}_A).$$

Here we have used the properness of the specification in the second step and the DLR equation in the third.

Further, take in particular a local event $A \in \mathcal{F}_\Lambda$. Then

$$\begin{aligned}\nu(A) = \int \nu(d\omega)\gamma_\Lambda(A \mid \omega) &= \int \nu(d\omega) \int \mu(d\tilde{\omega})\gamma_\Lambda(A \mid \omega) \\ &\geq c \int \nu(d\omega) \int \mu(d\tilde{\omega})\gamma_\Lambda(A \mid \tilde{\omega}) \geq c\mu(A).\end{aligned} \tag{2.3.5}$$

Hence, $\nu(A) \geq c\mu(A)$ also holds for all $A \in \mathcal{F}$ by the monotone class theorem. This implies $0 = \nu(B^c) \geq c\mu(B^c)$ and therefore $\mu(B) = 1$. □

3 Gibbs measures on trees

We specialize the index set to be the vertex set of a countably infinite tree. We discuss several Markov properties. There is the notion of a (spatially) Markov specification which means that the finite-volume Gibbs measures depend on their boundary condition only via a boundary layer of thickness one. This notion is meaningful on any graph. Similarly, an infinite-volume measure is called a (spatially) Markov field if its finite-volume conditional probabilities depend on the boundary condition only via the boundary layer of thickness one.

To be distinguished from the above notion, there is the notion of a tree-indexed Markov chain. This is meaningful only on trees. It relies on the definition of past and future vertices relative to a given oriented edge. While each tree-indexed Markov chain is a (spatially) Markov field, the converse statement is ensured only for extremal Gibbs measures. Indeed, the non-trivial part is that any extremal Gibbs measure for a Markov specification is a tree-indexed Markov chain. We will explain in detail why this is true, using conditioning arguments involving "future-tail triviality".

Then we come to describe the one-to-one correspondence between boundary laws and tree-indexed Markov chains. Boundary laws are families of positive measures on the local state space, indexed by the set of oriented edges which satisfy a consistency equation (tree-recursion). There is also a one-to-one correspondence between boundary laws and transition matrices of the tree-indexed Markov chain, given the specification.

We conclude with a discussion of all homogeneous boundary laws on the Cayley tree for concrete examples of the Ising model and the Potts model in zero magnetic field.

3.1 Construction of Gibbs measures via boundary laws

One of the most important classes of stochastic processes are Markov chains. A Markov chain in its most elementary form is a sequence of random variables indexed by $\mathbb{N}$ (which is usually interpreted as time) which has the property that future events are independent of the past given the information about its present state, i.e.,

$$\begin{aligned}&\mu(X_{n+1} = x_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1}, ..., X_0 = x_0)\\ =&\mu(X_{n+1} = x_{n+1} \mid X_n = x_n) \quad \text{for all } x_{n+1}, ..., x_0 \in \Omega_0.\end{aligned}$$

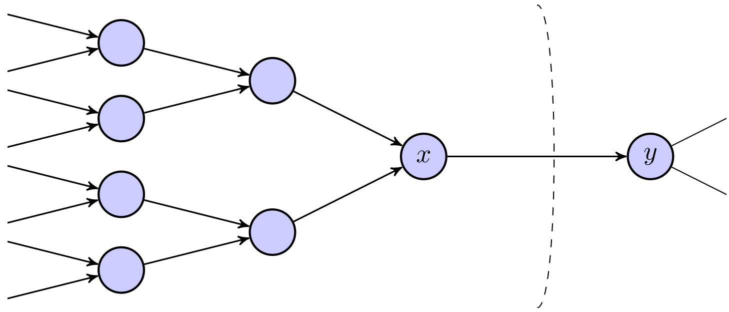


Figure 3.1: The set of oriented edges $(-\infty, xy)$.

There is a natural way to generalize this definition to the situation where the stochastic process is no longer indexed by $\mathbb{N}$ but by the vertices V of a tree. To formulate this we need some more notation. For any vertex $w \in V$ the set of the directed edges pointing away from w is given by

$$\vec{E}_w = \{\langle x, y\rangle \in E : d(w, y) = d(w, x) + 1\}.$$

This is an orientation of the set of edges induced by the vertex w. Furthermore we define the "past" of any oriented edge $\langle x, y\rangle \in \vec{E}$ by

$$(-\infty, xy) = \{w \in V \mid \langle x, y\rangle \in \vec{E}_w\}.$$

This is the set of sites w from which the oriented edge $\langle x, y\rangle$ is pointing away. The definition of the future of an oriented edge is analogous. Note that the tree property, i.e., the absence of loops, is clearly needed to give a meaningful definition of the "past" and "future" of an oriented edge.

In the following we will always restrict ourselves to the case where the local state space Ω_0 is finite. This simplifies the analysis but still allows the occurrence of phase transitions on trees.

Definition 3.1.1. *Let Ω_0 be the local state space and $\Omega = \Omega_0^V$. A measure $\mu \in \mathcal{M}_1(\Omega)$ is called a* tree-indexed Markov chain *if*

$$\mu(\sigma_y = \omega_y \mid \mathcal{F}_{(-\infty,xy)}) = \mu(\sigma_y = \omega_y \mid \mathcal{F}_{\{x\}}) \tag{3.1.1}$$

μ-a.s. for any $\langle x, y\rangle \in \vec{E}$ and any $\omega_y \in \Omega_0$. Any stochastic matrix P_{ij} on Ω_0 with

$$\mu(\sigma_j = y \mid \mathcal{F}_{\{i\}}) = P_{ij}(\sigma_i, y) \quad \mu\text{-a.s.}$$

for all $y \in \Omega_0$ will then be called a transition matrix *from i to j for μ.*

A Markov chain is said to be completely homogeneous *with transition matrix P if*

$$\mu(\sigma_j = y \mid \mathcal{F}_{\{i\}}) = P(\sigma_i, y) \quad \mu\text{-a.s.}$$

for all $y \in \Omega_0$ and all $\langle i,j\rangle \in \vec{E}$.

Remark 3.1.2. *Every tree-indexed Markov chain μ with transition matrices $(P_{ij})_{\langle i,j\rangle \in \vec{E}}$ and marginal distribution α_k at vertex $k \in V$ has the following representation*

$$\mu(\sigma_\Lambda = \zeta) = \alpha_k(\zeta_k) \prod_{\langle i,j\rangle \in \vec{E}_k : i,j \in \Lambda} P_{ij}(\zeta_i, \zeta_j) \tag{3.1.2}$$

for alle finite connected sets $\Lambda \Subset V$ and all $\zeta \in \Omega_0^\Lambda$ and $k \in \Lambda$. This can be seen by induction on the number of vertices in Λ. If μ is completely homogeneous it follows from equation (3.1.2) *that μ is invariant under the group $I(E)$ of all graph automorphisms of V. $I(E)$ is the set of all transformations $\tau : \Omega \to \Omega$ s.t. $\tau\omega = (\omega_{\tau_*^{-1} i})_{i \in V}$ where $\tau_* : V \to V$ is a bijection with the property that $\{\tau_* i, \tau_* j\} \in E$ if and only if $\{i,j\} \in E$.*

Besides the one-sided Markov property there is also the notion of a spatial Markov property:

Definition 3.1.3. *Let γ be a specification for Ω_0 and V. The specification is said to be a* Markov specification *if $\gamma_\Lambda(\sigma_\Lambda = \zeta \mid \cdot)$ is $\mathcal{F}_{\partial\Lambda}$-measurable for all $\zeta \in \Omega_0^\Lambda$ and every $\Lambda \Subset V$.*

Here $\partial\Lambda = \{i \in V : d(i,\Lambda) = 1\}$ is the outer boundary layer of thickness one. If γ is a Markov specification, then every $\mu \in \mathcal{G}(\gamma)$ is a *Markov field*, in that μ satisfies the spatial Markov property

$$\mu(\sigma_\Lambda = \zeta \mid \mathcal{F}_{\Lambda^c}) = \mu(\sigma_\Lambda = \zeta \mid \mathcal{F}_{\partial\Lambda}) \quad \mu\text{-a.s.}$$

for all $\zeta \in \Omega_0^\Lambda$ and every $\Lambda \Subset V$. Note that every Gibbsian specification which is defined by a nearest-neighbor potential is Markov.

Remark 3.1.4. *Every tree-indexed Markov chain is a Markov field.*

Proof. Assume that μ is a Markov chain. For $\Lambda \Subset V$ let $\Delta \Subset V$ be some finite connected set with $\Lambda \cup \partial\Lambda \subset \Delta$. The explicit form of the finite volume marginals of equation (3.1.2), applied in the bigger volume Δ, shows that

$$\mu(\sigma_\Delta = \zeta\omega\eta)\mu(\zeta'\omega\eta') = \mu(\sigma_\Delta = \zeta'\omega\eta)\mu(\sigma_\Delta = \zeta\omega\eta')$$

for all $\zeta, \zeta' \in \Omega_0^\Lambda, \omega \in \Omega_0^{\partial\Lambda}$, and $\eta, \eta' \in \Omega_0^{\Delta\setminus(\Lambda\cup\partial\Lambda)}$. Summing over ζ' and η' we obtain

$$\mu(\sigma_\Delta = \zeta\omega\eta)\mu(\sigma_{\partial\Lambda} = \omega) = \mu(\sigma_{\Delta\setminus\Lambda} = \omega\eta)\mu(\sigma_{\Lambda\cup\partial\Lambda} = \zeta\omega).$$

If $\mu(\sigma_{\Delta\setminus\Lambda} = \omega\eta) > 0$ we have

$$\mu(\sigma_\Lambda = \zeta \mid \sigma_{\Delta\setminus\Lambda} = \omega\eta) = \mu(\sigma_\Lambda = \zeta \mid \sigma_{\partial\Lambda} = \omega)$$

which means that

$$\mu(\sigma_\Lambda = \zeta \mid \mathcal{F}_{\Delta\setminus\Lambda}) = \mu(\sigma_\Lambda = \zeta \mid \sigma_{\partial\Lambda} = \omega) \quad \mu\text{-a.s.}$$

Since $\mathcal{F}_{\Lambda^c}$ is generated by the union of all these $\mathcal{F}_{\Delta\setminus\Lambda}$'s, we conclude that

$$\mu(\sigma_\Lambda = \zeta \mid \mathcal{F}_{\Lambda^c}) = \mu(\sigma_\Lambda = \zeta \mid \mathcal{F}_{\partial\Lambda}) \quad \mu\text{-a.s.}$$

Hence μ is a Markov field. □

Theorem 3.1.5. *Let γ be a Markov specification. Then each $\mu \in ex\,\mathcal{G}(\gamma)$ is a tree-indexed Markov chain.*

Proof. Take an oriented bond $\langle i, j\rangle \in \vec{E}$ and let $\Delta(n)$ be the ball of radius n around j and $\Lambda(n) = \Delta(n) \cap (ij, \infty)$ be the future in this ball relative to the oriented edge. As μ is assumed to be extremal, we know that μ is trivial on the tail-sigma-algebra $\mathcal{T} = \bigcap_{n\in\mathbb{N}} \mathcal{F}_{\Delta(n)^c}$ (see Proposition 2.2.5). Hence, μ is also trivial on the smaller sigma-algebra

$$\bigcap_{n\in\mathbb{N}} \mathcal{F}_{(ij,\infty)\setminus\Lambda(n)}.$$

This is the future tail sigma-algebra relative to the oriented bond.

This "future-tail triviality" implies that

$$\mathcal{F}_{\{i\}} = \bigcap_{n\geq 1} \mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)} \quad \mu\text{-a.s.}$$

Indeed, the sigma-algebra on the left is clearly contained in that on the right. Conversely, if $f : \Omega \to \mathbb{R}$ is bounded and measurable with respect the latter sigma-algebra then $f(x\sigma_{V\setminus\{i\}})$ is measurable w.r.t. $\bigcap_{n\geq 1} \mathcal{F}_{(ij,\infty)\setminus\Lambda(n)}$ for all $x \in \Omega_0$ and hence

$$f(x\sigma_{V\setminus\{i\}}) = \int f(x\omega_{V\setminus\{i\}})\mu(d\omega) \quad \mu\text{-a.s.}$$

Therefore f is also measurable w.r.t. $\mathcal{F}_{\{i\}}$.

As $\mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)}$ is a decreasing sequence of sigma-algebras we can apply the backwards martingale convergence theorem and arrive at

$$\mu(\sigma_j = y \mid \mathcal{F}_{\{i\}}) = \lim_{n\to\infty} \mu(\sigma_j = y \mid \mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)}) \quad \mu\text{-a.s.}$$

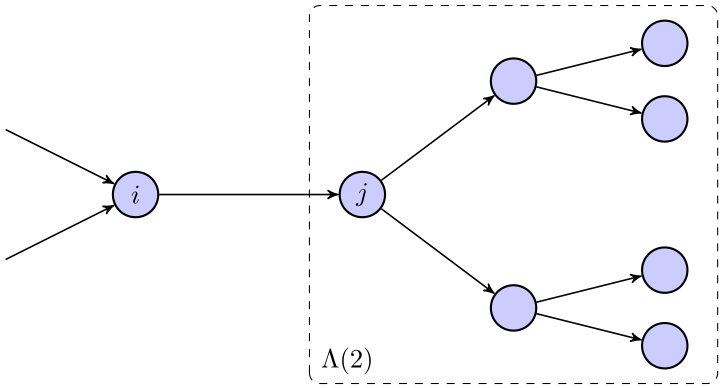


Figure 3.2: The set $\Lambda(2)$ in the binary tree.

Note that the term under the limit on the r.h.s. equals

$$\mu(\sigma_j = y \mid \mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)}) = \mu\left(\mu(\sigma_j = y \mid \mathcal{F}_{\Lambda(n)^c}) \mid \mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)}\right) \tag{3.1.3}$$

by the tower property of conditional expectation, as the "inner" sigma-algebra on the r.h.s. is finer that the "outer" sigma-algebra.

Since μ is a Gibbs measure, we have inside the conditional expectation on the r.h.s.

$$\mu(\sigma_j = y \mid \mathcal{F}_{\Lambda(n)^c}) = \gamma_{\Lambda(n)}(\sigma_j = y \mid \cdot) \tag{3.1.4}$$

which holds μ-a.s. Note that $\{i\} \cup (ij,\infty) \setminus \Lambda(n) \supset \partial\Lambda(n)$. Hence, by the Markov specification property for γ we may pull this out of the conditional expectation and arrive at the μ-a.s. equality

$$\mu(\sigma_j = y \mid \mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)}) = \gamma_{\Lambda(n)}(\sigma_j = y \mid \cdot) \tag{3.1.5}$$

Taking now the n-limit, we have

$$\begin{aligned}\lim_{n\to\infty} \mu\left(\sigma_j = y \mid \mathcal{F}_{\{i\}\cup(ij,\infty)\setminus\Lambda(n)}\right) &= \lim_{n\to\infty} \gamma_{\Lambda(n)}(\sigma_j = y \mid \cdot) \\ &= \lim_{n\to\infty} \mu(\sigma_j = y \mid \mathcal{F}_{\Lambda(n)^c}) \\ &= \mu\left(\sigma_j = y \mid \bigcap_{n\geq 1} \mathcal{F}_{\Lambda(n)^c}\right)\end{aligned} \tag{3.1.6}$$

where the second equality follows from the DLR-equation and where the last equation follows again from the backward martingale convergence theorem.

So we find that

$$\mu(\sigma_j = y \mid \mathcal{F}_{\{i\}}) = \mu\left(\sigma_j = y \mid \bigcap_{n\geq 1} \mathcal{F}_{\Lambda(n)^c}\right)$$

and by the inclusion $\bigcap_{n\geq 1} \mathcal{F}_{\Lambda(n)^c} \supset \mathcal{F}_{(-\infty,ij)} \supset \mathcal{F}_{\{i\}}$ follows from the tower property that

$$\mu(\sigma_j = y \mid \mathcal{F}_{(-\infty,ij)}) = \mu(\sigma_j = y \mid \mathcal{F}_{\{i\}}) \quad \mu\text{-a.s.}$$

Hence μ is a Markov chain.

□

Let Φ be some nearest-neighbor interaction potential which may contain also single-site terms. Recall that the corresponding Gibbsian specification γ^Φ is then given by

$$\begin{aligned}\gamma^\Phi_\Lambda(\sigma_\Lambda = \omega_\Lambda \mid \omega) &= Z_\Lambda(\omega)^{-1} \exp[-H_\Lambda(\omega)] \\ &= Z_\Lambda(\omega)^{-1} \exp[-\sum_{b\cap\Lambda\neq\emptyset} \Phi_b(\omega_b)].\end{aligned} \tag{3.1.7}$$

where the sum runs over all non-oriented edges b touching the finite volume Λ.

When we define *transfer operators* (or *transfer matrices*) by

$$Q_b(\omega_b) = \exp[-\Phi_b(\omega_b) - |\partial i|^{-1}\Phi_{\{i\}}(\omega_i) - |\partial j|^{-1}\Phi_{\{j\}}(\omega_j)] \tag{3.1.8}$$

where $b = \{i,j\} \in E$ and $\omega_b \in \Omega_0^b$, we can rewrite the specification kernel (3.1.7) as

$$\gamma^\Phi_\Lambda(\sigma_\Lambda = \omega_\Lambda \mid \omega) = Z_\Lambda(\omega)^{-1} \prod_{b\cap\Lambda\neq\emptyset} Q_b(\omega_b). \tag{3.1.9}$$

Note that by definition the transfer matrices are symmetric, i.e.,

$$Q_{ij}(x,y) = Q_{ji}(y,x)$$

whenever $\{i,j\} \in E$ and $y, x \in \Omega_0$.

In the following we will work towards a representation of tree-indexed Gibbs measures via the notion of so-called boundary laws [1], [6], [14]:

Definition 3.1.6. *A family of vectors $\{l_{ij}\}_{\langle i,j\rangle\in\vec{E}}$ with $l_{ij} \in (0,\infty)^{\Omega_0}$ is called a* boundary law *for the transfer operators $\{Q_b\}_{b\in E}$ if for each $\langle i,j\rangle \in \vec{E}$ there exists a constant $c_{ij} > 0$ such that the consistency equation*

$$l_{ij}(\omega_i) = c_{ij} \prod_{k\in\partial i\setminus\{j\}} \sum_{\omega_k\in\Omega_0} Q_{ki}(\omega_i,\omega_k) l_{ki}(\omega_k) \tag{3.1.10}$$

holds for every $\omega_i \in \Omega_0$.

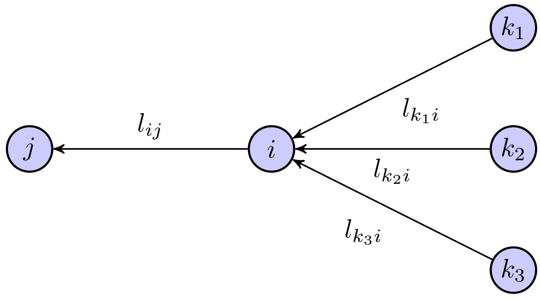


Figure 3.3: For the boundary law l the value of $l_{ij}(\omega_i)$ along the oriented edge $\langle i, j\rangle$ can be recursively determined by the values of $\{l_{k_m i}\}_{m=1,2,3}$ along the oriented edges pointing towards i via equation (3.1.10).

Boundary laws are maps from the oriented edges $\langle k, i\rangle$ to the positive measures on the single-site spin space at the site k.

For any boundary law, the family $\{\alpha_{ij} l_{ij}\}_{\langle i,j\rangle \in \vec{E}}$ for any fixed choice of strictly positive numbers α_{ij} is also a boundary law (simple exercise).

We will now give the main theorem of this chapter which shows the equivalence of boundary laws and tree-indexed Markov chains which are Gibbs measures for the given set of transfer operators.

Theorem 3.1.7. *Let a Markov specification of the form* (3.1.9) *be given, and let $\{Q_b, b \in E\}$ be its associated family of transfer matrices.*

(a) Each boundary law $\{l_{ij}\}_{\langle i,j\rangle \in \vec{E}}$ for $\{Q_b, b \in E\}$ defines a unique tree-indexed Markov chain $\mu \in \mathcal{G}(\gamma)$ via the equation

$$\mu(\sigma_{\Lambda\cup\partial\Lambda} = \omega_{\Lambda\cup\partial\Lambda}) = Z_\Lambda^{-1} \prod_{y\in\partial\Lambda} l_{y y_\Lambda}(\omega_y) \prod_{b\cap\Lambda\neq\emptyset} Q_b(\omega_b) \qquad (3.1.11)$$

where $\Lambda \subset V$ is any finite connected set, $\omega_{\Lambda\cup\partial\Lambda} \in \Omega_0^{\Lambda\cup\partial\Lambda}$, and Z_Λ is a suitable normalizing constant. y_Λ is the unique nearest neighbor of y which lies inside Λ.

(b) Conversely, every tree-indexed Markov chain $\mu \in \mathcal{G}(\gamma)$ admits a representation of the form (3.1.11) *in terms of a boundary law. This representation is unique in the sense that every boundary law is unique up to a positive factor.*

Proof. (a) In the first step we will use Kolmogorov's extension theorem to show that the expressions on the r.h.s. of (3.1.11) describe the marginals of a unique

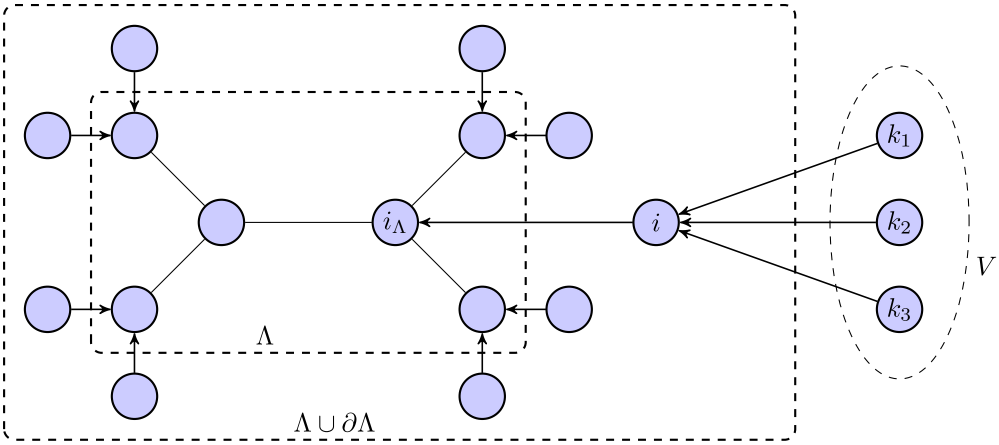


Figure 3.4: The boundary law property guarantees the consistency of the family of marginals given by equation (3.1.11).

measure $\mu \in \mathcal{M}(\Omega)$. This is true if the expressions are consistent, i.e.,

$$\sum_{\omega_V \in \Omega_0^V} Z_\Delta^{-1} \prod_{k \in \partial\Delta} l_{kk_\Delta}(\omega_k) \prod_{b \cap \Delta \neq \emptyset} Q_b(\omega_b) = Z_\Lambda^{-1} \prod_{k \in \partial\Lambda} l_{kk_\Lambda}(\omega_k) \prod_{b \cap \Lambda \neq \emptyset} Q_b(\omega_b) \tag{3.1.12}$$

whenever $\Lambda, \Delta \Subset V$ are any finite connected sets with $\Lambda \subset \Delta, V = (\Delta \cup \partial\Delta) \setminus (\Lambda \cup \partial\Lambda)$, and $\omega_{\Lambda \cup \partial\Lambda} \in \Omega_0^{\Lambda \cup \partial\Lambda}$. (Here we deviate inside the proof from the previous use of the symbol V to denote the infinite vertex set of the tree graph.) By induction it is enough to check this when $\Delta = \Lambda \cup \{i\}$ for any $i \in \partial\Lambda$.

In this case we have that $V = \partial i \setminus \{i_\Lambda\}$ and we find for the l.h.s. of (3.1.11)

$$\begin{aligned}
&\sum_{\omega_V \in \Omega_0^V} Z_\Delta^{-1} \prod_{k \in \partial\Delta} l_{kk_\Delta}(\omega_k) \prod_{b \cap \Delta \neq \emptyset} Q_b(\omega_b) \\
=& Z_\Delta^{-1} \prod_{k \in \partial\Lambda \setminus \{i\}} l_{kk_\Lambda}(\omega_k) \prod_{b \cap \Lambda \neq \emptyset} Q_b(\omega_b) \times \underbrace{\sum_{\omega_V \in \Omega_0^V} \left(\prod_{k \in V} l_{ki}(\omega_k) Q_{ki}(\omega_k, \omega_i) \right)}_{= \prod_{k \in V} \left(\sum_{\omega_k \in \Omega_0} l_{ki}(\omega_k) Q_{ki}(\omega_k, \omega_i) \right)} \\
=& \frac{1}{Z_\Delta c_{ii_\Lambda}} \prod_{k \in \partial\Lambda} l_{kk_\Lambda}(\omega_k) \prod_{b \cap \Lambda \neq \emptyset} Q_b(\omega_b)
\end{aligned} \tag{3.1.13}$$

where we have used the boundary law property in the last line. Summing over

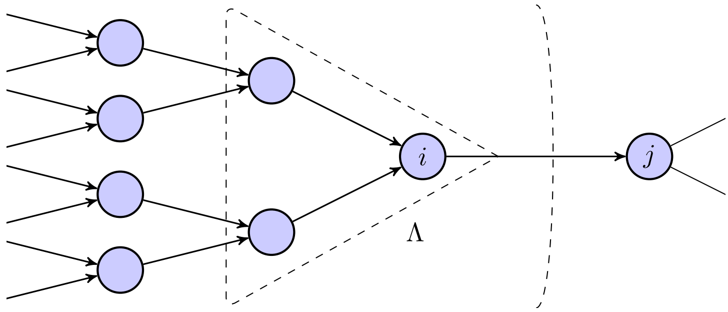


Figure 3.5: The finite connected set Λ lies in the past of the directed edge $\langle i, j\rangle$.

$\omega_{\Lambda\cup\partial\Lambda}$ shows that $Z_\Delta c_{ii_\Lambda} = Z_\Lambda$. This establishes the consistency.

In the next step we will see that every measure constructed in this way is a tree-indexed Markov chain. Let $\langle i, j\rangle \in \vec{E}$ be any directed edge and $\Lambda \Subset (-\infty, ij)$ be any finite connected set in the past of this edge with $j \in \partial\Lambda$. Furthermore, let $x, y \in \Omega_0$ and $\omega_{(\Lambda\cup\partial\Lambda)\setminus\{j\}} \in \Omega_0^{(\Lambda\cup\partial\Lambda)\setminus\{j\}}$. Substituting the finite-volume representation formula in terms of the boundary law we find

$$\frac{\mu(\sigma_j = y \mid \sigma_{(\Lambda\cup\partial\Lambda)\setminus\{j\}} = \omega_{(\Lambda\cup\partial\Lambda)\setminus\{j\}})}{\mu(\sigma_j = x \mid \sigma_{(\Lambda\cup\partial\Lambda)\setminus\{j\}} = \omega_{(\Lambda\cup\partial\Lambda)\setminus\{j\}})} = \frac{l_{ji}(y)Q_{ji}(y,\omega_i)}{l_{ji}(x)Q_{ji}(x,\omega_i)}. \tag{3.1.14}$$

Summing over $y \in \Omega_0$ gives us

$$\mu(\sigma_j = x \mid \sigma_{(\Lambda\cup\partial\Lambda)\setminus\{j\}} = \omega_{(\Lambda\cup\partial\Lambda)\setminus\{j\}}) = \frac{l_{ji}(x)Q_{ji}(x,\omega_i)}{\sum_y l_{ji}(y)Q_{ji}(y,\omega_i)}. \tag{3.1.15}$$

The expression on the r.h.s. of (3.1.15) depends on ω via ω_i only. Taking a limit of a sequence of finite sets $\Lambda_n \uparrow V$ gives us that

$$\mu(\sigma_j = x \mid \mathcal{F}_{(-\infty,ij)}) = \mu(\sigma_j = x \mid \mathcal{F}_{\{i\}}) \quad \mu\text{-a.s.}$$

and therefore μ is indeed a Markov chain.

In the third step we show that μ is a Gibbs measure. Let $\Lambda \Subset V$ be any finite subset of the infinite-volume vertex set V and take any two configurations $\zeta, \omega \in \Omega$ with $\zeta_{V\setminus\Lambda} = \omega_{V\setminus\Lambda}$. Let $\Delta \Subset V$ be any connected set with $\Lambda \subset \Delta$.

Then

$$\begin{aligned}\frac{\mu(\sigma_\Lambda = \zeta_\Lambda \mid \sigma_{(\Delta\cup\partial\Delta)\setminus\Lambda} = \omega_{(\Delta\cup\partial\Delta)\setminus\Lambda})}{\mu(\sigma_\Lambda = \omega_\Lambda \mid \sigma_{(\Delta\cup\partial\Delta)\setminus\Lambda} = \omega_{(\Delta\cup\partial\Delta)\setminus\Lambda})} &= \frac{\mu(\sigma_{\Delta\cup\partial\Delta} = \zeta_{\Delta\cup\partial\Delta})}{\mu(\sigma_{\Delta\cup\partial\Delta} = \omega_{\Delta\cup\partial\Delta})} \\ &= \prod_{b\cap\Delta\neq\emptyset} \frac{Q_b(\zeta_b)}{Q_b(\omega_b)} \\ &= \prod_{b\cap\Lambda\neq\emptyset} \frac{Q_b(\zeta_b)}{Q_b(\omega_b)} \\ &= \frac{\gamma_\Lambda(\sigma_\Lambda = \zeta_\Lambda \mid \omega)}{\gamma_\Lambda(\sigma_\Lambda = \omega_\Lambda \mid \omega)}.\end{aligned} \tag{3.1.16}$$

Finally we can sum over $\zeta_\Lambda \in \Omega_0^\Lambda$ and take the limit $\Delta \uparrow V$. This way we arrive at $\mu \in \mathcal{G}(\gamma)$.

(b) Now we assume that some Markov chain $\mu \in \mathcal{G}(\gamma)$ is given. On the one hand, we can condition from the inside to the outside using the Markov property. On the other hand we can also condition from the outside to the inside using the Gibbs property.

For any $\langle i, j\rangle \in \vec{E}$ we define transition probabilities in the usual way by $P_{ij}(x, y) = \mu(\sigma_j = y \mid \sigma_i = x)$. Let $\Lambda \Subset V$ be any finite connected set, $\zeta \in \Omega$, and $a \in \Omega_0$ some fixed reference state. Then

$$\mu(\sigma_{\Lambda\cup\partial\Lambda} = \zeta_{\Lambda\cup\partial\Lambda}) = \mu(A)\mu(B \mid A)\mu(C \mid B)/\mu(A \mid B)$$

where

$$A = \{\sigma_\Lambda = a\}, \quad B = \{\sigma_{\partial\Lambda} = \zeta_{\partial\Lambda}\}, \quad C = \{\sigma_\Lambda = \zeta_\Lambda\}.$$

Using the Markov property we find that

$$\mu(B \mid A) = \prod_{k\in\partial\Lambda} P_{k_\Lambda k}(a, \zeta_k).$$

On the other hand we get from the Gibbs property that

$$\frac{\mu(C \mid B)}{\mu(A \mid B)} = \frac{\gamma_\Lambda(C \mid \zeta)}{\gamma_\Lambda(A \mid \zeta)} = \frac{\prod_{b\cap\Lambda\neq\emptyset} Q_b(\zeta_b)}{\prod_{b\subset\Lambda} Q_b(a, a) \prod_{k\in\partial\Lambda} Q_{k_\Lambda k}(a, \zeta_k)}.$$

Hence

$$\mu(\sigma_{\Lambda\cup\partial\Lambda} = \zeta_{\Lambda\cup\partial\Lambda}) = \frac{\mu(\sigma_\Lambda = a)}{\prod_{b\subset\Lambda} Q_b(a, a)} \prod_{k\in\partial\Lambda} \frac{P_{k_\Lambda k}(a, \zeta_k)}{Q_{k_\Lambda k}(a, \zeta_k)} \prod_{b\cap\Lambda\neq\emptyset} Q_b(\zeta_b). \tag{3.1.17}$$

Therefore the finite-volume representation of the marginals as in equation (3.1.11) holds with

$$Z_\Lambda^{-1} = \frac{\mu(\sigma_\Lambda = a)}{\prod_{b \subset \Lambda} Q_b(a,a)}$$

and the candidate for a boundary law

$$l_{ij}(x) = \frac{P_{ji}(a,x)}{Q_{ji}(a,x)}$$

for all $\langle i,j \rangle \in \vec{E}$ and $x \in \Omega_0$. If we choose $\Delta = \Lambda \cup \{i\}$ with $i \in \partial\Lambda$ we can see the defining equation of the boundary by the consistency of μ, if we reread the steps of the proof for part (a) in the opposite direction.

To prove the uniqueness of the boundary law we assume that μ admits a second representation for the form (3.1.11) with a boundary law $\{l'_{ij} : \langle i,j \rangle \in \vec{E}\}$ and normalizing constants $z'_\Lambda > 0$. We apply (3.1.11) to a singleton $\Lambda = \{i\}$ and a configuration ω with $\omega_j = x$ for some $j \in \partial i$ and $\omega_k = a$ for all $k \in i \cup \partial i \setminus \{j\}$. We obtain

$$\frac{Z'_i}{Z_i} = \frac{l'_{ji}(x)}{l_{ji}(x)} \prod_{k \in \partial i \setminus \{j\}} \frac{l'_{ki}(a)}{l_{ki}(a)}$$

and hence l' is equal to l up to a positive pre-factor (in general depending on the directed edge). This completes the proof of the theorem. □

Remark 3.1.8. *If $l = 1$ is a solution to the boundary law equation we find that* (3.1.9) *is the marginal distribution of a Markov chain $\mu^{free} \in \mathcal{G}(\gamma)$. We call μ^{free} the* free Gibbs measure. *For a b.l. $l \neq 1$ we get a Gibbs measure that is different from this free solution.*

As every extremal Gibbs measure is a Markov chain, Theorem 3.1.7 gives us that $|\mathcal{G}(\gamma)| = 1$ if and only if there is exactly one solution to the b.l. equation.

3.2 Completely homogeneous tree-indexed Markov chains on the Cayley tree: Ising and Potts models

In the following we will take a closer look at completely homogeneous Markov chains $\mu \in \mathcal{G}(\gamma)$ on *Cayley trees*. Let $k \in \mathbb{N}_*$. A Cayley tree of order k is an infinite tree where each vertex has $k+1$ nearest neighbors.

It will be denoted by $\mathcal{CT}(k)$. The same object is equivalently called a $k+1$-regular tree. In the case $k = 2$ one speaks of a binary tree.

A Markov specification γ on $\mathcal{CT}(k)$ is said to be *completely homogeneous* with *transfer matrix Q* if γ satisfies (3.1.9) with $Q_b = Q$ for every $b \in E$. Recall that

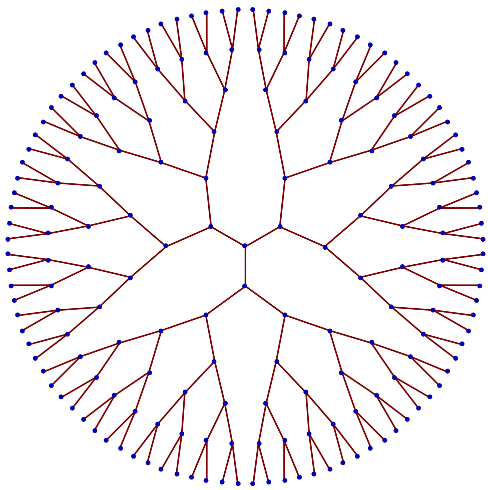


Figure 3.6: An embedding of the binary Cayley tree into the plane.

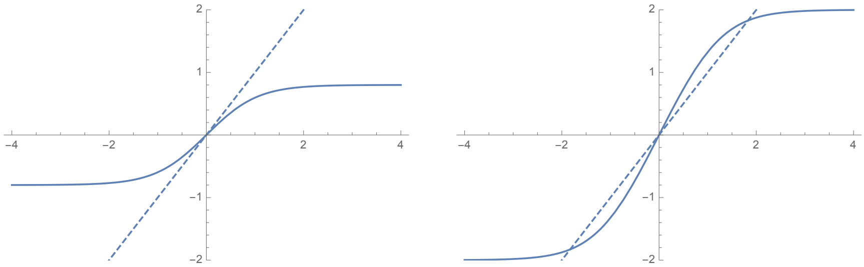
in the proof of part (b) of Theorem 3.1.7 we have not only shown that every $\mu \in \mathcal{G}(\gamma)$ admits a representation in the form of (3.1.11) but that the boundary law therein is given by $l_{ij}(x) = \frac{P_{ji}(a,x)}{Q_{ji}(a,x)}$ (up to an $\langle i,j \rangle$-dependent constant). Therefore μ is completely homogeneous if and only if $l_{ij} = l$ (up to edge-dependent multiplicative constants) for all $\langle i,j \rangle \in \vec{E}$ and some $l \in (0,\infty)^{\Omega_0}$.

As every boundary law is only unique up to a factor we may normalize at a reference state $a \in \Omega_0$. We say that a boundary law $\{l_{ij} : \langle i,j \rangle \in \vec{E}\}$ is *normalized at a* if $l_{ij}(a) = 1$ for all $\langle i,j \rangle \in \vec{E}$. If l corresponds to a completely homogeneous Markov chain it has to meet

$$l(x) = \left(\frac{\sum_{y \in \Omega_0} l(y) Q(x,y)}{\sum_{y \in \Omega_0} l(y) Q(a,y)} \right)^k . \tag{3.2.1}$$

3.2.1 The Ising model in zero magnetic field

In the Ising model the local state space is $\Omega_0 = \{-1, +1\}$. We have some interaction strength $J > 0$ which is fixed and a nearest-neighbor interaction potential Φ with $\Phi_{i,j}(\omega_i, \omega_j) = -J\omega_i\omega_j$. The corresponding Markov specifica-



(a) $J = 0.4$. (b) $J = 1$.

Figure 3.7: The function f_J is odd, concave on the positive half-line and convex on the negative half-line. The derivative $f_J'(0)$ is increasing in J and becomes bigger than 1 for large enough values of J.

tion γ is completely homogeneous and the transfer matrix Q is given by

$$Q(-,-) = Q(+,+) = \exp(J), \qquad Q(-,+) = Q(+,-) = \exp(-J).$$

According to our previous discussion there is a one-to-one correspondence between the completely homogeneous Markov chains $\mu \in \mathcal{G}(\gamma)$ and the positive solutions $s > 0$ of

$$s = \left(\frac{Q(-,+) + sQ(+,+)}{Q(-,-) + sQ(+,-)} \right)^k = \left(\frac{se^J + e^{-J}}{e^J + se^{-J}} \right)^k . \tag{3.2.2}$$

Here, we normalize at $a = -1$ and hence may look for boundary laws of the form $l = (1, s)$. Introducing the new variable $t = \frac{1}{2} \log s$ equation (3.2.2) is equivalent to

$$t = \frac{k}{2} \log \frac{\cosh(J + t)}{\cosh(J - t)} =: f_J(t).$$

The r.h.s. of this equation is an odd function in t which is concave for $t > 0$ and convex for $t < 0$. Hence, equation (3.2.2) has only the trivial solution $s = 1$ if and only if $f_J'(0) = k \tanh(J) \le 1$. If $f_J'(0) > 1$ we find two additional solutions $s_*, -s_*$ to the trivial one. Hence, there is a phase transition in this case as every solution s corresponds to a completely homogeneous Markov chain $\mu_s \in \mathcal{G}(\gamma)$. We only know that there only is one *completely homogeneous* Markov chain $\mu \in \mathcal{G}(\gamma)$ if $f_J'(0) = k \tanh(J) \le 1$. It can be shown that there actually is only one Gibbs measure overall in this case which means that $J = \operatorname{atanh}(1/k)$ is the sharp threshold for phase transition in this model [6, Theorem 12.31].

This provides a proof of Theorem 1.1.8.

3.2.2 The Potts model

In the Potts model the local state space is given by $\Omega_0 = \{1, ..., q\} \simeq \mathbb{Z}_q$ and the nearest-neighbor potential is $\Phi_{i,j}(\omega_i, \omega_j) = \beta \mathbf{1}_{\omega_i = \omega_j}$ which gives us for the transfer matrix $Q(\omega_i, \omega_j) = e^{\beta \mathbf{1}_{\omega_i = \omega_j}} = \theta^{\mathbf{1}_{\omega_i = \omega_j}}$ with $\theta = e^{\beta}$. The homogeneous b.l. equation is

$$l(s) = c \left(\sum_{\tilde{s}} l(\tilde{s}) Q(\tilde{s}, s) \right)^k$$

and hence

$$\frac{l(s)}{l(q)} = \left(\frac{l(s)(\theta - 1) + \sum_{\tilde{s}=1}^{q-1} l(\tilde{s}) + l(q)}{l(q)\theta + \sum_{\tilde{s}=1}^{q-1} l(\tilde{s})} \right)^k, \quad s \in \{1, ..., q-1\}.$$

For $z_s := \frac{l(s)}{l(q)} \in (0, \infty)$ this leads to

$$z_s = \left(\frac{(\theta - 1)z_s + \sum_{\tilde{s}=1}^{q-1} z_{\tilde{s}} + 1}{\theta + \sum_{\tilde{s}=1}^{q-1} z_{\tilde{s}}} \right)^k. \tag{3.2.3}$$

The solutions to this $(q-1)$-dimensional fixed-point equation are in a one-to-one correspondence with the completely homogeneous tree-indexed Markov chains $\mu \in \mathcal{G}(\gamma)$.

Proposition 3.2.1. *For any solution $z = (z_1, ..., z_{q-1})$ of (3.2.3) there exists a set $M \subset \{1, ..., q-1\}$ and some $z^* > 0$ s.t.*

$$z_s = \begin{cases} 1 & ,\text{if } s \notin M \\ z^* & ,\text{if } s \in M \end{cases}.$$

Proof. Assume we have a solution of the boundary law equation. Define the set M to be the set of indices for which the entry of the boundary law is different from 1. We will show the boundary law entries will have to be the same for all indices in M.

Indeed, take $\theta \neq 1$ and assume w.l.o.g. that $|M| = m$ with $M = \{1, ..., m\}$. Define $x_s := z_s^{1/k}$. Then

$$x_s = \frac{(\theta - 1)x_s^k + \left(\sum_{j=1}^m x_j^k + q - m \right)}{\sum_{j=1}^m x_j^k + (q - m - 1) + \theta}$$

where $z_s = 1$ for $s \notin M$ and $z_s \neq 1$ if $s \in M$. When we set $R := \sum_{j=1}^m x_j^k + q - m$ we get

$$\begin{aligned} x_s = \frac{(\theta - 1)x_s^k + R}{R + \theta - 1} &\Leftrightarrow (x_s^k - x_s)(\theta - 1) = (x_s - 1)R \\ &\Leftrightarrow x_s(x_s^{k-2} + x_s^{k-3} + ... + 1)(\theta - 1) = R. \end{aligned} \tag{3.2.4}$$

The polynomial on the l.h.s. has positive coefficients and is monotone increasing in x_s, hence injective. Therefore $x_s = x_{\tilde{s}}$ for all $s, \tilde{s} \in \{1, ..., m\}$. □

Corollary 3.2.2. *Any completely homogeneous tree-indexed Markov chain $\mu \in \mathcal{G}(\gamma)$ corresponds to a solution of*

$$z = f_m(z) := \left(\frac{(\theta + m - 1)z + q - m}{mz + q - m - 1 + \theta} \right)^k$$

for some $m \in \{1, ..., q-1\}$.

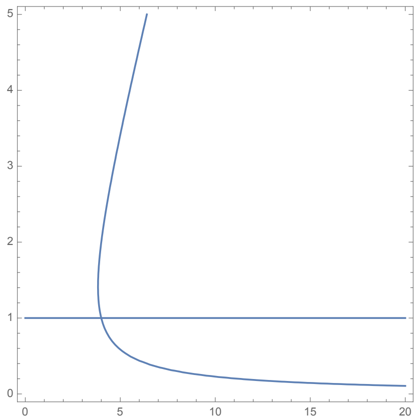
In the following let us specialize to the binary tree. For $x = \sqrt{z}$ we have

$$x = \frac{(\theta + m - 1)x^2 + q - m}{mx^2 + q - m - 1 + \theta}. \tag{3.2.5}$$

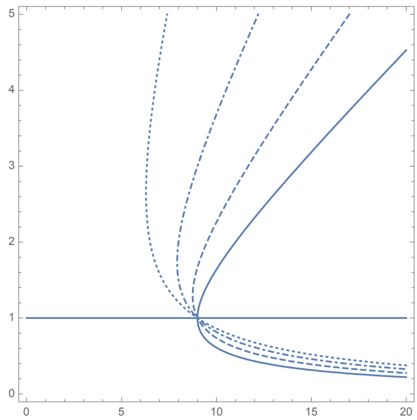
Now divide out the root $x = 1$ and solve the resulting quadratic equation. Put $\theta_m = 1 + 2\sqrt{m(q-m)}$, $m = 1, ..., q-1$ and note that $\theta_m = \theta_{q-m}$. We have $\theta_1 < \theta_2 < ... < \theta_{\lfloor q/2 \rfloor - 1} < \theta_{\lfloor q/2 \rfloor} \leq q + 1$ and the boundary law solutions are given by

$$x_{1,2}(m, \theta) = \frac{\theta - 1 \pm \sqrt{(\theta - 1)^2 - 4m(q-m)}}{2m}$$

and they exist for $\theta \geq \theta_m$.



(a) $q = 3$.



(b) $q = 8$.

Figure 3.8: The branches appearing in this figure correspond to all the possible solutions to the homogeneous b.l. equation (3.2.5). Here the value of θ is plotted on the horizontal axis against the corresponding solutions $x(\theta)$. For $q = 3$ there only is the trivial solution and the branch for $m = 1$. For $q = 8$ there are four branches for the cases $m = 1, 2, 3, 4$. The dotted branch corresponds to $m = 1$, the dotted and dashed one to $m = 2$, the dashed one to $m = 3$ and the full branch to $m = 4$.

4 Reconstruction problem and extremality

4.1 Some results and open problems

We have seen how to identify tree-indexed Markov chains $\mu \in \mathcal{G}(\gamma)$ via boundary laws. The extremal elements of $\mathcal{G}(\gamma)$ are necessarily Markov chains. However the converse statement is not always true. For the Ising model in zero magnetic field the situation is well-understood. The situation is captured in Figure 4.1a. There are two different critical transition temperatures. First, there is the critical value where additional solutions to the trivial one appear. This is where a 2nd order phase transition occurs. Secondly, there is the point where the Gibbs measure which corresponds to the trivial b.l. solution is no longer extremal. This is a strange behavior that only happens on trees.

For the Potts model the picture is not complete but we have some partial information about the extremality and non-extremality of the b.l. solutions [9], see Figure 4.1b for an example. There are parameter regions for which extremality is proved to occur, regions where non-extremality is proved to occur, between which there remain some gaps.

The following theorem gives a sufficient criterion for non-extremality. In the following we will assume all measures, boundary laws etc. to be completely homogeneous.

Theorem 4.1.1 (Kesten-Stigum bound, 1966 [8]). *Let μ be a completely homogeneous tree-indexed Markov chain on the Cayley tree $\mathcal{CT}(d)$. Denote by λ_2 the second largest in modulus eigenvalue of the transition matrix P. Assume that $d\lambda_2^2 > 1$. Then μ is not tail-trivial.*

This theorem goes back to Kesten and Stigum who have formulated it in the context of multi-colored branching processes. A multi-colored branching process is a generalization of a Galton-Watson process to multiple types. In the following we will discuss these results and why they give the above mentioned criterion for non-extremality. We give no proof of the Kesten-Stigum theorem itself here. (It relies on martingale convergence arguments.)

Kesten and Stigum define a multi-type branching process with q types in the following way. Let U denote the set of all q-dimensional vectors whose components are non-negative integers. Further, let $(X_n)_{n\in\mathbb{N}}$ be a vector Markov process with temporally homogeneous transition probabilities, whose states are vectors in U. We interpret X_n^i, the i-th component of X_n, as the number of objects of type i in the n-th generation.

The vector X_{n+1} in the next time-step is obtained by independently replacing each individual of type i in the population described by X_n, by an offspring family described by $(r_0, ..., r_{q-1})$, which consists precisely of r_j children of type j for all $j \in \{0, ..., q-1\}$, which happens with probability $p^i(r_0, ..., r_{q-1})$.

In particular, let $X_0 = e_i$ for some $i \in \{0, ..., q-1\}$, meaning that there is initally precisely one organism of type i. Then X_1 will have the generating function

$$\mathrm{E}(\prod_{l=0}^{q-1} s_l^{X_1^l}) = \sum_{r_1,...,r_{q-1}=0}^{\infty} p^i(r_0, ..., r_{q-1}) s_1^{r_0} ... s_{q-1}^{r_{q-1}}, \quad |s_0|, ..., |s_{q-1}| \leq 1, \tag{4.1.1}$$

A process (X_n) of this kind is called *multitype branching process* [7].
For the Ising or the Potts model e.g. it is clear that all tree-indexed Markov chain Gibbs measures μ can be interpreted as multitype branching processes. In this picture the process X_n describes the histogramm of the vertices in a layer of distance n to a root, over the different types. Note that the genealogy is not contained.

Let M be the matrix whose ij-th entry gives the expected number of particles of type j produced by a single particle of type i, i.e. $M_{ij} = \sum_{r=0}^{\infty} r p_{ij}(r)$, where $p_{ij}(r)$ is the probability that a particle of type i produces r particles of type j.

For the Cayley tree $\mathcal{CT}(d)$, denoting by P the transition matrix of a tree-indexed Markov chain Gibbs measure μ, we clearly have $M = dP$.
Now, assume that M is strictly positive with largest eigenvalue $m > 1$. Then there exists a unique probability vector w such that $wM = mw$, by the Perron-Frobenius theorem. Let r be the eigenvalue of M with second largest modulus. Note that in our case we always have $m = d$ and $r = d\lambda_2$, where λ_2 is again the eigenvalue of P with second largest modulus. Let $Z^{(n)} = X_n \cdot A$, where A is any vector orthogonal to w. Kesten and Stigum showed that if $|r^2| > m$ then $Z^{(n)}/|r|^n$ converges with probability one to a random variable Z, whose distribution depends on the transition probabilities of the branching process, but also, in a non-trivial way on the initial particle type i, [8, Theorem 2.1].

Let us assume that the tree-indexed Markov chain Gibbs measure μ was extremal. Then it would also be tail-trivial, which would imply that the tail-measurable random variable $Z := \lim_{n\to\infty} Z^{(n)}/r^n$ would have to be μ-a.s. constant. For $|r^2| < m$, which is equivalent to $d\lambda_1^2 > 1$, this is not the case, as the measure on the Cayley tree is obtained by mixing the distribution obtained by propagating type i from an arbitrary root, over the (non-degenerate) invariant distribution of P. Hence μ can not be extremal under this condition.

The bound is known to be sharp for the symmetric Ising model (which can e.g. be seen by evaluating the constant from Theorem 4.2.1 in the next section)

This provides a proof of Theorem 1.1.9.

It is even sharp for the asymmetric Ising model with small enough asymmetry which is more difficult to see. In the Potts model with $q = 3$ the bound is sharp at least for sufficiently large degree of the Cayley tree.

On the other hand the KS-bound is not sharp for the Ising model with sufficiently large asymmetry and the symmetric Potts model with $q \geq 5$. It is still an open question if the KS-bound is always tight for the symmetric Potts channel when $q \leq 4$ [13].

4.2 A sufficient criterion for extremality

There are several approaches to ensure extremality. We present just one of them.

Theorem 4.2.1 (Formentin/Külske, 2009 [4])**.** *Let μ be a tree-indexed Markov chain on the Cayley tree $\mathcal{CT}(d)$. Let $\alpha \in \mathcal{M}_1(\{1, ..., q\})$ be an invariant and reversible probability vector for the transition matrix P, i.e.,*

$$\alpha P = \alpha.$$

Define the symmetrized relative entropy *by*

$$L(p) := S(p \mid \alpha) + S(\alpha \mid p) = \sum_{a=1}^{q} (p(a) - \alpha(a)) \log \frac{p(a)}{\alpha(a)} \geq 0$$

where $p \in \mathcal{M}_1(\{1, ..., q\})$. Denote

$$c(P) := \sup_{\substack{p \in \mathcal{M}_1(\{1,...,q\}) \\ p \neq \alpha}} \frac{L(pP)}{L(p)}.$$

Assume $dc(P) < 1$. Then μ is tail-trivial.

Remark 4.2.2. *In the Ising model in zero magnetic field α is the equidistribution and it can be shown by computations that $c(P) = \lambda_2^2(P)$. It turns out the* sup *is achieved in any arbitrarily small neighborhood of the invariant distribution, but this is a somewhat lucky case.*

In general, the function L is non-negative and $L(p) = 0$ if and only if $p = \alpha$. Also $p \mapsto L(p)$ is convex and can be thought of as a "Lyapunov-function" for the boundary law equation (tree recursion).

In the Potts model it turns out that the supremum can be attained at a non-trivial value away from the invariant distribution which makes the problem harder.

Proof. We will only discuss the main ideas. We denote by T^N the tree rooted at 0 of depth N. The notation T_v^N indicates the sub-tree of T^N rooted at v obtained from "looking to the outside" on the tree T^N. Let $(\xi_i)_{i\in\partial T_v^N}$ be a boundary configuration. We define

$$\pi_v^N(\xi) := \mu(\sigma_v = s \mid \sigma_{\partial T_v^N} = \xi)_{s=1,\dots,q}$$

which is the conditional probability that the spin variable at the vertex v takes the spin value s, conditioned on the boundary configuration ξ at distance N. In the case of the Ising model clearly it suffices to consider only one (say the first) component. Compare Chapter 1.

We want to show that

$$\int \mu(d\xi)\|\pi_v^N(\xi) - \alpha\|_{TV} \to 0 \tag{4.2.1}$$

for $N \to \infty$. Then it follows that μ is tail-trivial. Indeed, if (4.2.1) holds then

$$\lim_{N\to\infty} \int \mu(d\xi)|\mu(f \mid \mathcal{F}_{\Delta_N^c}) - \mu(f)| = 0 \tag{4.2.2}$$

for all local functions $f \in \mathcal{F}_\Lambda$ where $(\Delta_N)_{N\in\mathbb{N}}$ is any sequence of finite sets with $\Delta_N \uparrow V$. This can be shown by using tree properties. Then (4.2.2) also holds for any bounded $\mathcal{F}$-measurable function f. Take $f = \mathbf{1}_A, A \in \mathcal{T}_\infty$. Then

$$\mu(|\mathbf{1}_A - \mu(A)|) = 0$$

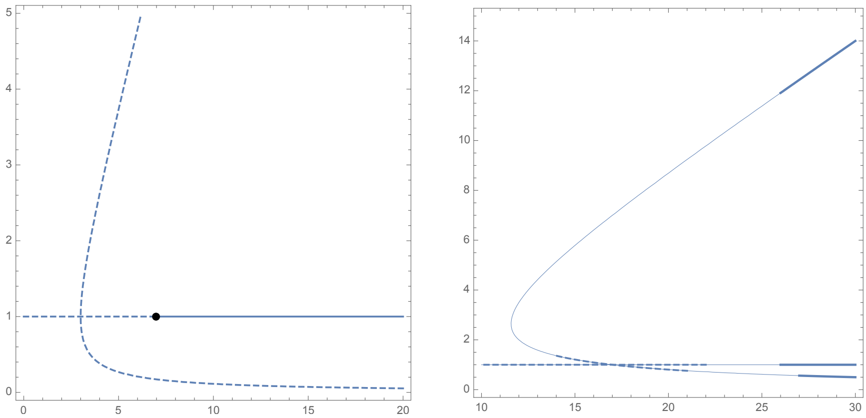
which implies $\mu(A) \in \{0,1\}$.

To prove (4.2.1) we make use of an invariance property of tree recursions. We only mention without proof (see [4, Proposition 3.3]) that we have the important (and non-obvious!) identity

$$\int \mu(d\xi)L(\pi_v^N(\xi)) = \sum_{w:v\to w} \int \mu(d\xi)L(\pi_w^N(\xi)P).$$

As a warning, note that this property fails pointwise, and in general, $L(\pi_v^N(\xi)) \neq \sum_{w:v\to w} L(\pi_w^N(\xi)P)$ for fixed boundary condition ξ. To control $L(\pi_w^N(\xi)P)$ we bring the constant $c(P)$ into play. For simplicity specialize to the Cayley tree $\mathcal{CT}(d)$. We find

$$\begin{aligned} \mu(L(\pi_v^N)) &= d\ \mu(L(\pi_w^N P)) \\ &\le d\ c(P)\mu(L(\pi_w^N)) \\ &= d\ c(P)\mu(L(\pi_v^{N-1})) \end{aligned} \tag{4.2.3}$$



(a) $q = 2$.

(b) $q = 16$.

Figure 4.1: Value of non-trivial boundary law component versus coupling parameter θ. In the Ising model there is only one branch of solutions ($m = 1$) besides the trivial one. These solutions are always extremal. The trivial solution is extremal only for small values of θ and becomes non-extremal for θ large. This change happens at the black dot.
In the Potts model with $q = 16$ the branch of solutions for $m = 2$ is shown. The dashed part is the region where these solutions correspond provably to an extremal Gibbs measure. The thick lines indicate provable non-extremality of the corresponding measure. We see that there are gaps between these two scenarios for which we don't know whether extremality or non-extremality occurs. The situation is similar for the trivial solution. For sufficiently small values of θ the free Gibbs measure is extremal and for sufficiently large values it becomes non-extremal.

and this implies

$$\lim_{N\to\infty} \mu(L(\pi_v^N)) = 0$$

□

There are also other methods ensuring extremality which in general give different parameter regions, see for instance [9].

5 Recent developments: Gradient Gibbs measures on trees

We have already seen that for finite local state spaces Ω_0 there is a one-to-one correspondence between boundary laws and tree-indexed Markov chains $\mu \in \mathcal{G}(\gamma)$. Let us mention without proof that a similar result still holds when Ω_0 is countable, e.g. $\Omega_0 = \mathbb{Z}$, for some but *not for all* boundary laws. The additional assumption to construct Gibbs measures via boundary laws that needs to be put on the boundary law is its *normalizability* which means in this context

$$\sum_{\omega_x \in \mathbb{Z}} \left(\prod_{z \in \partial x} \sum_{\omega_z \in \mathbb{Z}} Q_{zx}(\omega_x, \omega_z) l_{zx}(\omega_z) \right) < \infty \tag{5.0.1}$$

for each $x \in V$ [14]. What happens to *non-normalizable boundary law solutions* was systematically investigated in [10] and we will report on these finding here.

5.1 Definitions

Assume that the spin variables take value in the local state space $\Omega_0 = \mathbb{Z}$, so we are facing the new difficulty that the state space is unbounded and hence non-compact. This means that there is no general argument which would imply that the set of Gibbs measures is non-empty.

For a configuration $\omega = (\omega(x))_{x \in V}$ and $b = \langle v, w \rangle \in \vec{E}$ the *height difference* along the edge b is given by $\nabla \omega_b = \omega_w - \omega_v$. We call $\nabla \omega$ the *gradient field* of ω. The gradient spin variables are now defined by $\eta_{\langle x,y \rangle} = \sigma_y - \sigma_x$ for each $\langle x, y \rangle \in \vec{E}$, and we define the projection mappings similarly as before for the height variables. Let us denote the state space of the gradient configurations by $\Omega^\nabla = \mathbb{Z}^V / \mathbb{Z}$ which becomes a measurable space with the sigma-algebra $\mathcal{F}^\nabla = \sigma(\{\eta_b \mid b \in \vec{E}\}) = \mathcal{P}(\mathbb{Z})^{\vec{E}}$. This is the space of all the possible gradient fields that can be prescribed by some height configuration $\omega \in \mathbb{Z}^V$, and trivially every gradient field $\zeta \in \Omega^\nabla$ gives a height configuration ω^{ζ, ω_x} for a fixed value of $\omega_x, x \in V$ by

$$\omega_y^{\zeta, \omega_x} = \omega_x + \sum_{b \in \Gamma(x,y)} \zeta_b, \tag{5.1.1}$$

where $\Gamma(x, y)$ is the unique self-avoiding path from x to y.

Let some symmetric nearest-neighbor *gradient interaction potential* $U_b : \mathbb{Z} \to \mathbb{R}$ be given for every $b = \{x, y\} \in E$, i.e.

$$U_b(m) = U_b(-m)$$

for all $m \in \mathbb{Z}$. We put

$$\Phi_b(\omega_x, \omega_y) = U_b(|\omega_x - \omega_y|).$$

The local Gibbsian specification γ^Φ is then defined in the usual way as in (3.1.7) for the case of finite local state space (assuming finite partition functions).

Boundary laws will then be defined exactly in the same way as in Definition 3.1.6 where we now put $\Omega_0 = \mathbb{Z}$ with the transfer operator

$$Q_{xy}(\omega_x, \omega_y) = e^{-U_{\{x,y\}}(|\omega_x - \omega_y|)}.$$

As the transfer-operator depends only on height-differences it can be described by the functions

$$\mathbb{Z} \ni m \mapsto R_{xy}(m) = e^{-U_{\{x,y\}}(m)}.$$

In the following we will take a closer look at a special class of boundary laws.

Definition 5.1.1. *A boundary law is called to be* q-periodic *if* $l_{xy}(\omega_x + q) = l_{xy}(\omega_x)$ *for every oriented edge* $\langle x, y\rangle \in \vec{E}$ *and each* $\omega_x \in \mathbb{Z}$*.*

Note that these boundary laws are not normalizable. It is now an interesting question whether or not these objects still correspond to some well-defined equilibrium state in some sense. The short answer is yes, but only when we look at *gradient profiles*. Regular Gibbs measures will in general not exist as the local state space is unbounded. However, there is a common way to circumvent this problem: The gradient interaction potential is invariant under an overall height-shift by a constant in every spin variable and this allows us to divide out this degree of freedom. This naturally leads us to study Gibbs measures in the space of gradient configurations.

Definition 5.1.2. *The* gradient Gibbs specification *is defined as the family of probability kernels* $(\gamma'_\Lambda)_{\Lambda \Subset V}$ *from* $(\Omega^\nabla, \mathcal{T}^\nabla_\Lambda)$ *to* $(\Omega^\nabla, \mathcal{F}^\nabla)$ *such that*

$$\int F(\rho)\gamma'_\Lambda(d\rho \mid \zeta) = \int F(\nabla\varphi)\gamma_\Lambda(d\varphi \mid \omega) \tag{5.1.2}$$

for all bounded $\mathcal{F}^\nabla$*-measurable functions* F*, where* $\omega \in \Omega$ *is any height-configuration with* $\nabla\omega = \zeta$*.*

A measure $\nu \in \mathcal{M}_1(\Omega^\nabla)$ *is called a* gradient Gibbs measure (GGM) *if it satisfies the DLR equation*

$$\int \nu(d\zeta)F(\zeta) = \int \nu(d\zeta) \int \gamma'_\Lambda(d\tilde{\zeta} \mid \zeta)F(\tilde{\zeta}) \tag{5.1.3}$$

for every finite $\Lambda \subset V$ *and for all* $F \in \mathcal{C}_b(\Omega^\nabla)$. *The set of gradient Gibbs measures will be denoted by* $\mathcal{G}^\nabla(\gamma)$ *or* $\mathcal{G}^\nabla(R)$ *where* $(R_b)_{b\in E}$ *are the (functions describing the) transfer operators corresponding to* Φ.

Here, $\mathcal{T}_\Lambda^\nabla$, the sigma-algebra of gradient configurations outside of the finite volume Λ is generated by all gradient variables outside of Λ *and* the relative height-difference on the boundary of Λ. For precise definitions and more discussion, see [10].

5.2 Construction of GGMs via q-periodic boundary laws

Definition 5.2.1. *Let the* mod-q fuzzy map $T_q : \mathbb{Z} \to \mathbb{Z}_q$ *be given by* $T_q(i) = i \bmod q$, *where* $\mathbb{Z}_q = \{0, ..., q-1\}$ *for* $n \in \mathbb{N}$.

In the first step we construct gradient measures which do not have the full Gibbs property. For this we define marginal measures in some analogy to the boundary law representation in the case of finite spaces (3.1.11), but supplemented with internal information about layers.

Theorem 5.2.2. *Let a vertex* $w \in \Lambda$, *where* $\Lambda \subset V$ *is any finite connected set, and a class label* $s \in \mathbb{Z}_q$ *be given. Then any* q*-periodic boundary law* $\{l_{xy}\}_{\langle x,y\rangle\in\vec{E}}$ *for* $\{R_b\}_{b\in E}$ *defines a consistent family of probability measures on the gradient space* Ω^∇ *by*

$$\nu_{w,s}(\eta_{\Lambda\cup\partial\Lambda} = \zeta_{\Lambda\cup\partial\Lambda}) = c_\Lambda(w,s) \prod_{y\in\partial\Lambda} l_{yy_\Lambda}\left(\varphi'_y(s,\zeta)\right) \prod_{b\cap\Lambda\neq\emptyset} R_b(\zeta_b), \tag{5.2.1}$$

where $\zeta_{\Lambda\cup\partial\Lambda} \in \mathbb{Z}^{\vec{E}(\Lambda\cup\partial\Lambda)}$. *Here*

$$\varphi'_y(s,\zeta) = T_q\Big(s + \sum_{b\in\Gamma(w,y)} \zeta_b\Big)$$

denotes the class in $\mathbb{Z}_q$ *obtained by walking from class* s *at the site* $w \in \Lambda$ *along the unique path* $\Gamma(w,y)$ *to the boundary site* y *whose class is determined by the gradient configuration* ζ. *Since the boundary law is a class function, expression* (5.2.1) *is well-defined, where* $c_\Lambda(w,s)$ *is a normalization factor that turns* $\nu_{w,s}$ *into a probability measure on* $\mathbb{Z}^{\vec{E}(\Lambda\cup\partial\Lambda)}$.

Let us agree to call the measures $\nu_{w,s}$ on the space of gradients Ω^{∇} we have just constructed: *pinned gradient measures*. To prove the theorem and verify Kolmogorov-compatibility (see [10]) is not much more difficult than the analogous but simpler statement of Theorem 3.1.7 we proved in this course for finite local state space models.

Given a boundary law $\{l_{xy}\}_{\langle x,y\rangle\in\vec{E}}$ we define an associated transition matrix by

$$P_{xy}(\omega_x,\omega_y)=\frac{R_{yx}(\omega_y-\omega_x)l_{yx}(\omega_y)}{\sum_{\omega_y\in\mathbb{Z}}R_{yx}(\omega_y-\omega_x)l_{yx}(\omega_y)}.$$

Such a transition matrix describes transition probabilities for a random walk on $\mathbb{Z}$. If the boundary law has period q, then, taking into account this periodicity we can introduce the associated transition matrices $\bar{P}_{xy}:\mathbb{Z}_q\times\mathbb{Z}\mapsto[0,1]$ in the following way:

$$P_{xy}(\omega_x,\omega_y)=:\bar{P}_{xy}(T_q(\omega_x),\omega_y-\omega_x).$$

In this notation $\bar{P}_{xy}(T_q(i),j-i)$ denotes the probability to see an height increase of $j-i$ along the edge $\langle x,y\rangle$ given the class $T_q(i)$ in the vertex x. This describes a random walk on $\mathbb{Z}$ in a periodic environment. If we restrict the pinned gradient measures from the tree to a path starting at w, they are random walk measures in q-periodic environment.

Let us additionally assume that $R_b=R$ for all $b\in E$. Until now T could have been any locally finite tree. From now on we will restrict ourselves to the case of the Cayley tree with $d+1$ nearest neighbors. We call a vector $l\in(0,\infty)^{\mathbb{Z}}$ a (spatially homogeneous) boundary law if there exists a constant $c>0$ such that the consistency equation

$$l(i)=c\left(\sum_{j\in\mathbb{Z}}R(i-j)l(j)\right)^d \tag{5.2.2}$$

is satisfied for every $i\in\mathbb{Z}$.

Note that by assumption $l(i)=1$ for every $i\in\mathbb{Z}$ is always a solution (assuming summability of R). Given such a homogeneous boundary law l we get for the associated transition matrix

$$P(i,j)=\frac{R(i-j)l(j)}{\sum_{k\in\mathbb{Z}}R(i-k)l(k)}.$$

The associated transition matrix $\bar{P}:\mathbb{Z}_q\times\mathbb{Z}\mapsto[0,1]$ is then given by

$$P(i,j)=:\bar{P}(T_q(i),j-i).$$

Furthermore, let the fuzzy transform $T_q P : \mathbb{Z}_q \times \mathbb{Z}_q \to [0,1]$ be defined by

$$T_q P(\bar{i}, \bar{j}) = \sum_{j : T_q(j) = \bar{j}} P(\bar{i}, j) =: P'(\bar{i}, \bar{j}) \tag{5.2.3}$$

for all layers $\bar{i}, \bar{j} \in \mathbb{Z}_q$.

The main result regarding GGMs on trees is stated in the following theorem. We have already seen how to construct pinned gradient measures. For this we have to single out a pinning site $w \in V$ and pin some fuzzy layer $s \in \mathbb{Z}_q$. Such a measure will now depend on the layer selected. We have to mix over all the layers in an appropriate way s.t. the measure becomes spatially homogeneous. The right mixing measure for this averaging procedure is the invariant measure of the associated fuzzy chain P'.

Theorem 5.2.3. *Let $\Lambda \subset V$ be any finite connected set and let $w \in \Lambda$ be any vertex. Let $\alpha^{(l)} \in \mathcal{M}_1(\mathbb{Z}_q)$ denote the unique invariant distribution for the fuzzy transform $T_q P^l$ of the transition matrix P^l corresponding to the q-periodic homogeneous boundary law l. Then the measure $\nu^l \in \mathcal{M}_1(\Omega^\nabla)$ with marginals given by*

$$\nu^l(\eta_\Lambda = \zeta_\Lambda) = \sum_{s \in \mathbb{Z}_q} \alpha^{(l)}(s)\, \nu_{w,s}(\eta_\Lambda = \zeta_\Lambda) \tag{5.2.4}$$

defines a (spatially) homogeneous GGM. Here $\nu_{w,s}$ are the pinned gradient measures constructed in Theorem 5.2.2.

We don't give a proof here, but point out that mixing is need to recover both properties, spatial homogeneity and Gibbs property.

5.3 Coupling measure $\bar{\nu}$, Gibbsian preservation under transform of measure

Let us make an additional comment on the structure which has unfolded. We have frequently used projections of the spins to different directions: On the one hand an infinite-volume spin-configuration $\omega \mapsto^{T_q} \omega'$ maps to a mod-q fuzzy spin ω', via our fuzzy map $T_q(i) = i \bmod q$. On the other hand, an infinite-volume spin-configuration $\omega = (\omega_w, \zeta) \mapsto^\nabla \zeta$ also maps to a gradient configuration. The additional information needed to recover the spin is provided by its value ω_w at a pinning site w.

Let us call an infinite-volume gradient configuration ζ and an infinite-volume fuzzy spin configuration ω' *compatible* iff there exists an infinite-volume spin-configuration ω for which $T_q\omega = \omega'$ and $\nabla\omega = \zeta$. This is to say that the two configurations have a joint lift to a proper spin configuration.

The defining function $R(m) = e^{-U(m)}$ of the gradient model also has a natural *mod-q-fuzzy image*, namely

$$R'(m) := T_q R(m) = \sum_{j\in\mathbb{Z}} R(qj+m)$$

Taking a logarithm R' describes a renormalized Hamiltonian for a clock model on $\mathbb{Z}_q$. (Commonly one speaks of a clock if it possesses a discrete rotational $\mathbb{Z}_q$-symmetry.)

Suppose now we are on a regular tree and have found an (in height-direction) q-periodic tree-automorphism invariant boundary law l. The definition of the GGMs ν which are mixed over the fuzzy chain (see formula (5.2.4)) extends in a natural way to a *joint measure (or coupling measure)* $\bar\nu(d\omega', d\zeta)$. This coupling measure has the following properties:

1. $\bar\nu\Big(\omega'\text{and } \zeta \text{ are compatible}\Big) = 1.$
2. The marginal on gradients $\bar\nu(d\zeta)$ is a tree-automorphism invariant GGM.
3. The marginal on fuzzy spins $\bar\nu(d\omega')$ is a tree-automorphism invariant finite-state Gibbs measure for R'.

The definition of $\bar\nu$ is given by spelling out its expectation $\bar\nu(F)$ on a bounded local observable $F(\zeta_\Lambda, \omega'_\Lambda)$ depending both on gradient variables and layer variables in a finite volume Λ. The formula says that we need to substitute the layer variables which are obtained by pinning the layer at one site, and the gradient information ζ and it reads

$$\begin{aligned}\bar\nu(F) = \sum_{\zeta_\Lambda}\sum_{s\in\mathbb{Z}_q} \alpha^{(l)}(s) \prod_{\langle x,y\rangle\in \vec{E}_w : x,y\in\Lambda} \bar P_{x,y}\Big(T_q\big(s + \sum_{b\in\Gamma(w,x)} \zeta_b\big), \zeta_{\langle x,y\rangle}\Big)\\ F\Big(\zeta_\Lambda, \big(T_q\big(s+\sum_{b\in\Gamma(w,u)}\zeta_b\big)\big)_{u\in\Lambda}\Big).\end{aligned} \tag{5.3.1}$$

Here we have assumed that the pinning site w is in the finite volume Λ.

In Figure 5.1 the main result of this chapter, Theorem 5.2.3, is visualized as the curved dashed arrow. Here we have denoted the set of q-periodic tree invariant b.l.'s by BL^{q}, the set of Gibbs measures on the fuzzy spins which are tree-indexed Markov chains by $\mathcal{G}'_{\mathrm{MC}}(T_q R)$, the set of tree-invariant measures on the fuzzy spins by $\mathcal{M}_1(\Omega')$, the set of coupling measures on (Ω^∇, Ω') which correspond to a b.l. via (5.3.1) by $\bar\nu^{\cdot}(\mathrm{BL}^q)$, the set of measures on the set of *compatible* gradient and fuzzy configurations which are tree-invariant by $\mathcal{M}_1^{\mathrm{cp}}(\Omega^\nabla, \Omega')$ and the set of tree-invariant measures on Ω^∇ by $\mathcal{M}_1(\Omega^\nabla)$.

$$
\begin{tikzcd}
 & \mathcal{G}'_{\mathrm{MC}}(T_q R) \arrow[dl, bend right] & \subset & \mathcal{M}_1(\Omega') & \text{fuzzy spins } \omega' \\
\mathrm{BL}^{\mathrm{q}} \arrow[r, "\bar{\nu}^{\cdot}\,:\, l \mapsto \bar{\nu}^{l}"] \arrow[dr, dashed, bend right, "\nu^{\cdot}\,:\, l \mapsto \nu^{l}"'] & \bar{\nu}^{\cdot}(\mathrm{BL}^{q}) \arrow[u, "(\omega',\eta) \mapsto \omega'"] \arrow[d, "(\omega',\eta) \mapsto \eta"'] & \subset & \mathcal{M}_1^{\mathrm{cp}}(\Omega', \Omega^{\nabla}) \arrow[u] \arrow[d] & \text{compatible } (\omega',\eta) \\
 & \mathcal{G}^{\nabla}(R) & \subset & \mathcal{M}_1(\Omega^{\nabla}) & \text{gradients } \eta \\
 & \cup & & \cup & \\
\mathrm{BL}^{\mathrm{Norm}} \arrow[r, leftrightarrow, "\nu^{\cdot}_{\mathrm{Zac}}\,:\, l \mapsto \nu^{l}"] & \mathcal{G}_{\mathrm{MC}}(R) & \subset & \mathcal{M}_1(\Omega) & \text{spins } \omega
\end{tikzcd}
$$

Figure 5.1: The relationship between q-periodic b.l.'s and the (gradient) Gibbs measures is displayed above the dashed line. The classical theory by Zachary for normalizable b.l.'s is visualized below the dashed line. Note that $\nu^{\cdot}_{\mathrm{Zac}} : l \mapsto \nu^{l}$ symbolizes the mapping which sends every normalizable b.l. to a unique tree-indexed Markov chain which is a Gibbs measure [14]. Here all objects appearing are assumed to be tree automorphism invariant.

Below the dashed line we have also given a visualization of the classical theory of Zachary [14] and its correspondence to the presented GGM theory. Every normalizable tree-invariant boundary law $l \in \mathrm{BL}^{\mathrm{Norm}}$ corresponds to a Gibbs measure which is a Markov chain. This set of measures is denoted by $\mathcal{G}_{\mathrm{MC}}(R)$. Conversely, every $\nu \in \mathcal{G}(R)$ which is also a Markov chain can be represented by a b.l. which is unique (up to a positive pre-factor). The set of measures $\mathcal{G}_{\mathrm{MC}}(R)$ can be thought of as a subset of the *gradient* Gibbs measures $\mathcal{G}^{\nabla}(R)$, as any Gibbs measure gives rise to a gradient measure, but not vice versa. Note that by the theory of Zachary applied to finite local state space, there is also a one-to-one correspondence between the elements of $\mathcal{G}'_{\mathrm{MC}}(T_q R)$ and BL^q. All the objects we construct above the dashed line are new. They are not contained in the theory of Zachary, as $\nu^{\cdot}(\mathrm{BL}^q) \subset \mathcal{G}^{\nabla}(R) \setminus \mathcal{G}_{\mathrm{MC}}(R)$, that is our gradient Gibbs measures live in the delocalized regime and can not be understood as projection of Gibbs measures to the gradient variables.

Bibliography

[1] J. T. Cox, *Entrance laws for Markov chains*, Ann. Probab., **5**, No. 3, 533-549, 1977.

[2] R. M. Dudley, *Real analysis and probability*, Revised reprint of the 1989 original, Cambridge Studies in Advanced Mathematics **74**, Cambridge University Press, Cambridge, 2002.

[3] A. C. D. van Enter, R. Fernández, A. D. Sokal, *Regularity properties and pathologies of position-space renormalization-group transformations: Scope and limitations of Gibbsian theory*, J. Stat. Phys., **7**2, No. 5, 879-1167, 1993.

[4] M. Formentin, C. Külske, *A Symmetric Entropy Bound on the Non-Reconstruction Regime of Markov Chains on Galton-Watson Trees*, Electron. Comm. Probab., **14**, 587-596, 2009.

[5] S. Friedli, Y. Velenik, *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*, unpublished book project, preprint available at `http://www.unige.ch/math/folks/velenik/smbook/`

[6] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, 2nd ed., de Gryuter Studies in Mathematics **9**, Walter de Gruyter, Berlin, 2011.

[7] T. E. Harris, *The Theory of Branching Processes,* Springer, Berlin, 1963.

[8] H. Kesten, B. P. Stigum, *Additional limit theorems for indecomposable multidimensional Galton-Watson processes*, Ann. Math. Stat., **37**, No. 6, 1463-1481, 1966.

[9] C. Külske, U. Rozikov, *Fuzzy transformations and extremality of Gibbs measures for the Potts model on a Cayley tree*, Random Structures and Algorithms **5**0 No. 4, 636-678, 2017.

[10] C. Külske, P. Schriever, *Gradient Gibbs measures and fuzzy transformations on trees*, Preprint, arXiv:1609.00159, 2016.

[11] F. Rassoul-Agha, T. Seppäläinen, *A Course on Large Deviations with an Introduction to Gibbs Measures*, Graduate Studies in Mathematics **162**, AMS, Providence, 2015.

[12] P. Schriever, *Gibbs measures and spin models on trees and random graphs*, PhD dissertation, Ruhr-Universität Bochum, 2016.

[13] A. Sly, *Reconstruction for the Potts Model*, Ann. Probab., **39**, No. 4, 1365-1406, 2011.

[14] S. Zachary, *Countable state space Markov random fields and Markov chains on trees*, Ann. Probab., **11**, No. 4, 894-903, 1983.