6 Infinite-Volume Gibbs Measures

In this chapter, we give an introduction to the theory of Gibbs measures, which describes the properties of infinite systems at equilibrium. We will not cover all the aspects of the theory, but instead present the most important ideas and results in the simplest possible setting, the Ising model being a guiding example throughout the chapter.

Remark 6.1. Due to the rather abstract nature of this theory, it will be necessary to resort to some notions from measure theory that were not necessary in the previous chapters. From the probabilistic point of view, we will use extensively the fundamental notion of *conditional expectation*, central in the description of Gibbs measures. The reader familiar with these subjects (some parts of which are briefly presented in Appendix B, Sections B.5 and B.8) will certainly feel more comfortable. Certain topological notions will also be used, but will be presented from scratch along the chapter. Nevertheless, we emphasize that although of great importance in the understanding of the mathematical framework of statistical mechanics, *a detailed understanding of this chapter is not required for the rest of the book.* ⋄

Some models to which the theory applies. The theory of Gibbs measures presented in this chapter is general and applies to a wide range of models. Although the description of the equilibrium properties of these models will always follow the standard prescription of Equilibrium Statistical Mechanics, what distinguishes them is their *microscopic* specificities. That is, in our context: (i) the possible values of a spin at a given vertex of $\mathbb{Z}^d$, and (ii) the interactions between spins contained in a finite region $\Lambda \Subset \mathbb{Z}^d$.

A model is thus defined by first considering the set Ω_0, called the **single-spin space**, which describes all the possible states of one spin. The **spin configurations** on a (possible infinite) subset $S \subset \mathbb{Z}^d$ are defined as in Chapter 3:

$$\Omega_S \stackrel{\text{def}}{=} \Omega_0^S = \{(\omega_i)_{i\in S} : \omega_i \in \Omega_0 \, \forall i \in S\} .$$

When $S = \mathbb{Z}^d$, we simply write $\Omega \equiv \Omega_{\mathbb{Z}^d}$. Then, for each finite subset $\Lambda \Subset \mathbb{Z}^d$, the energy of a configuration in Λ is determined by a **Hamiltonian**

$$\mathscr{H}_\Lambda : \Omega \to \mathbb{R}.$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

We list some of the examples that will be used as illustrations throughout the chapter.

- For the **Ising model**,

$$\Omega_0 = \{+1, -1\}.$$

  The nearest-neighbor version studied in Chapter 3 corresponds to

$$\mathscr{H}_\Lambda(\omega) = -\beta \sum_{\{i,j\} \in \mathscr{E}_\Lambda^{\mathrm{b}}} \omega_i \omega_j - h \sum_{i \in \Lambda} \omega_i ,$$

  where we remind the reader that $\mathscr{E}_\Lambda^{\mathrm{b}}$ is the set of nearest-neighbor edges of $\mathbb{Z}^d$ with at least one endpoint in Λ, see (3.2). We will also consider a long-range version of this model:

$$\mathscr{H}_\Lambda(\omega) = -\sum_{\{i,j\} \cap \Lambda \neq \varnothing} J_{ij} \omega_i \omega_j - h \sum_{i \in \Lambda} \omega_i ,$$

  where $J_{ij} \to 0$ (sufficiently fast) when $\|j - i\|_1 \to \infty$.

- For the ***q*-state Potts model**, where $q \geq 2$ is an integer, we set

$$\Omega_0 = \{0, 1, 2, \ldots, q-1\},$$

$$\mathscr{H}_\Lambda(\omega) = -\beta \sum_{\{i,j\} \in \mathscr{E}_\Lambda^{\mathrm{b}}} \delta_{\omega_i, \omega_j} .$$

- For the **Blume–Capel model**,

$$\Omega_0 = \{+1, 0, -1\},$$

$$\mathscr{H}_\Lambda(\omega) = -\beta \sum_{\{i,j\} \in \mathscr{E}_\Lambda^{\mathrm{b}}} (\omega_i - \omega_j)^2 - h \sum_{i \in \Lambda} \omega_i - \lambda \sum_{i \in \Lambda} \omega_i^2 .$$

- The ***XY* model** is an example with an uncountable single-spin space,

$$\Omega_0 = \left\{ x \in \mathbb{R}^2 : \|x\|_2 = 1 \right\},$$

  and Hamiltonian

$$\mathscr{H}_\Lambda(\omega) = -\beta \sum_{\{i,j\} \in \mathscr{E}_\Lambda^{\mathrm{b}}} \omega_i \cdot \omega_j ,$$

  where $\omega_i \cdot \omega_j$ denotes the scalar product.

All the models above have a common property: their single-spin space is *compact* (see below). Models with non-compact single-spin spaces present additional interesting difficulties which will not be discussed in this chapter. One important case, the **Gaussian Free Field** for which $\Omega_0 = \mathbb{R}$, will be studied separately in Chapter 8.

About the point of view adopted in this chapter. Describing the above models in infinite volume will require a fair amount of mathematical tools. For simplicity, *we will only expose the details of the theory for models whose spins take their values in* $\{\pm 1\}$; the set of configurations is thus the same as in Chapter 3: $\Omega = \{\pm 1\}^{\mathbb{Z}^d}$.

Even with this simplification, we will face most of the mathematical difficulties that are unavoidable when attempting to describe infinite systems at equilibrium. It will however allow us to provide elementary proofs, in several cases, and to somewhat reduce the overall amount of abstraction (and notation) required.

Let us stress that the set $\{\pm 1\}$ has been chosen for convenience, but that it could be replaced by any finite set; our discussion (including the proofs) applies essentially verbatim also in that setting. In fact, all the results presented here remain valid, modulo some minor changes, for any model whose spins take their values in a compact set. At the end of the chapter, in Section 6.10, we will mention the few differences that appear in this more general situation.

So, from now on, and until the end of the chapter, unless explicitly stipulated otherwise, Ω_0 *will be* $\{\pm 1\}$*, and*

$$\Omega_\Lambda = \{\pm 1\}^\Lambda, \quad \Omega = \{\pm 1\}^{\mathbb{Z}^d}.$$

Outline of the chapter

The probabilistic framework used to describe infinite systems on the lattice will be presented in Section 6.2, together with a motivation for the notion of *specification*, central to the definition of infinite-volume Gibbs measures. After introducing the necessary topological notions, the existence of Gibbs measures will be proved in Section 6.4. Several uniqueness criteria, among which Dobrushin's condition of weak dependence, will be described in Section 6.5. Gibbs measures enjoying symmetries will be described rapidly in Section 6.6; translation invariance, which plays a special role, will be described in Section 6.7. In Section 6.8, the convex structure of the set of Gibbs measures will be described, as well as the decomposition of any Gibbs measure into a convex combination of extremal elements and the latter's remarkable properties. In Section 6.9, we will present the variational principle, which provides an alternative description of translation-invariant Gibbs measures, in more thermodynamical terms. In Section 6.10, we will sketch the changes necessary in order to describe infinite systems whose spins take infinitely many values, the latter being considered at several places in the rest of the book. In Section 6.11, we give a criterion for non-uniqueness involving the non-differentiability of the pressure, which will be used later in the book. The remaining sections are complements to the chapter.

6.1 The problem with infinite systems

Let us recall the approach used in Chapter 3. By considering for example the + boundary condition, we started in a finite volume $\Lambda \Subset \mathbb{Z}^d$, and defined the Gibbs distribution unambiguously by

$$\mu^+_{\Lambda;\beta,h}(\omega) = \frac{e^{-\mathscr{H}_{\Lambda;\beta,h}(\omega)}}{\mathbf{Z}^+_{\Lambda;\beta,h}}, \quad \omega \in \Omega^+_\Lambda.$$

Then, to describe the Ising model on the infinite lattice, we introduced the thermodynamic limit. We considered a sequence of subsets $\Lambda_n \uparrow \mathbb{Z}^d$ and showed, for each local function f, existence of the limit

$$\langle f \rangle^+_{\beta,h} = \lim_{n\to\infty} \langle f \rangle^+_{\Lambda_n;\beta,h}.$$

This defined a linear functional $\langle \cdot \rangle^+_{\beta,h}$ on local functions, which was called an *infinite-volume Gibbs state.*

This procedure was sufficient for us to determine the phase diagram of the Ising model (Section 3.7), but leaves several natural questions open. For instance, we know that

$$\lim_{n\to\infty} \mu^+_{\Lambda_n;\beta,h}(\sigma_0 = -1) = \lim_{n\to\infty} \tfrac{1}{2}\bigl(1 - \langle \sigma_0 \rangle^+_{\Lambda_n;\beta,h}\bigr) = \tfrac{1}{2}\bigl(1 - \langle \sigma_0 \rangle^+_{\beta,h}\bigr)$$

exists. This raises the question whether this limit represents the probability that $\sigma_0 = -1$ under some *infinite-volume* probability measure $\mu^+_{\beta,h}$:

$$\mu^+_{\beta,h}(\sigma_0 = -1) = \tfrac{1}{2}\bigl(1 - \langle \sigma_0 \rangle^+_{\beta,h}\bigr). \tag{6.1}$$

In infinite volume, neither the Hamiltonian nor the partition function are well-defined. Moreover, it is easy to check that each individual configuration would have to have probability zero. Therefore, extending the definition of a Gibbs distribution to the uncountable set of configurations Ω requires a different approach, involving the methods of measure theory.

6.2 Events and probability measures on Ω

As we said above, it is easy to construct a probability distribution on a finite set such as Ω_Λ, since this can be done by specifying the probability of each configuration. Another convenient consequence of the finiteness of Ω_Λ is that the set of events associated to Ω_Λ is naturally identified with the collection $\mathscr{P}(\Omega_\Lambda)$ of all subsets of Ω_Λ. The set of probability distributions on the finite measurable space $(\Omega_\Lambda, \mathscr{P}(\Omega_\Lambda))$ is denoted simply $\mathscr{M}_1(\Omega_\Lambda)$.

Notation 6.2. *In this chapter, it will often be convenient to add a subscript to configurations to specify explicitly the domain in which they are defined. For example elements of Ω_Λ will usually be denoted $\omega_\Lambda, \eta_\Lambda$, etc.*

Given $S \subset \mathbb{Z}^d$ and a configuration ω defined on a set larger than S, we will also write ω_S to denote the restriction of ω to S, $(\omega_i)_{i\in S}$. We will also often decompose a configuration $\omega_S \in \Omega_S$ as a concatenation: $\omega_S = \omega_\Lambda \omega_{S\setminus\Lambda}$ (for some $\Lambda \subset S$).

These notations should not to be confused with the notation in Chapter 3, where σ_Λ was used to denote the product of all spins in Λ, while the restriction of ω to Λ was written $\omega|_\Lambda$.

We first define the natural collection of *events* on Ω, based on the notion of *cylinder*. The restriction of $\omega \in \Omega$ to $S \subset \mathbb{Z}^d$, ω_S, can be expressed using the **projection map** $\Pi_S : \Omega \to \Omega_S$:

$$\Pi_S(\omega) \stackrel{\text{def}}{=} \omega_S.$$

In particular, with this notation, given $A \in \mathscr{P}(\Omega_\Lambda)$, the event that "$A$ occurs in Λ" can be written $\Pi_\Lambda^{-1}(A) = \{\omega \in \Omega : \omega_\Lambda \in A\}$.

For each $\Lambda \Subset \mathbb{Z}^d$, consider the set

$$\mathscr{C}(\Lambda) \stackrel{\text{def}}{=} \bigl\{\Pi_\Lambda^{-1}(A) : A \in \mathscr{P}(\Omega_\Lambda)\bigr\}$$

of all events on Ω that depend only on the spins located inside Λ. Each event $C \in \mathscr{C}(\Lambda)$ is called a **cylinder (with base Λ)**. For example, $\{\omega_0 = -1\}$, the event containing all configurations ω for which $\omega_0 = -1$, is a cylinder with base $\Lambda = \{0\}$.

Exercise 6.1. *Show that $\mathscr{C}(\Lambda)$ has the structure of an **algebra**: (i) $\varnothing \in \mathscr{C}(\Lambda)$, (ii) $A \in \mathscr{C}(\Lambda)$ implies $A^{\mathrm{c}} \in \mathscr{C}(\Lambda)$, and (iii) $A, B \in \mathscr{C}(\Lambda)$ implies $A \cup B \in \mathscr{C}(\Lambda)$.*

For any $S \subset \mathbb{Z}^d$ (possibly infinite), consider the collection

$$\mathscr{C}_S \stackrel{\text{def}}{=} \bigcup_{\Lambda \Subset S} \mathscr{C}(\Lambda)$$

of all **local events** in S, that is, all events that depend on finitely many spins, all located in S.

Exercise 6.2. *Check that, for all $S \subset \mathbb{Z}^d$, $\mathscr{C}_S$ contains at most countably many events and that it has the structure of an algebra.* Hint: *first, show that $\mathscr{C}(\Lambda) \subset \mathscr{C}(\Lambda')$ whenever $\Lambda \subset \Lambda'$.*

The **σ-algebra generated by cylinders with base contained in S** is denoted by

$$\mathscr{F}_S \stackrel{\text{def}}{=} \sigma(\mathscr{C}_S)$$

and consists of all the events that depend only on the spins inside S. When $S = \mathbb{Z}^d$, we simply write

$$\mathscr{C} \equiv \mathscr{C}_{\mathbb{Z}^d}\,, \quad \mathscr{F} \equiv \sigma(\mathscr{C})\,.$$

The cylinders $\mathscr{C}$ should be considered as the algebra of *local* events. Although generated from these local events, the σ-algebra $\mathscr{F}$ automatically contains *macroscopic* events, that is, events that depend on the system as a whole (a precise definition of macroscopic events will be given in Section 6.8.1). For example, the event

$$\Bigl\{\omega \in \Omega : \limsup_{n\to\infty} \frac{1}{|\mathrm{B}(n)|} \sum_{i \in \mathrm{B}(n)} \omega_i > 0\Bigr\} = \bigcup_{k\geq 1} \bigcap_{n \geq 1} \bigcup_{m \geq n} \Bigl\{\frac{1}{|\mathrm{B}(m)|} \sum_{i \in \mathrm{B}(m)} \omega_i \geq \tfrac{1}{k}\Bigr\}$$

belongs to $\mathscr{F}$ (and is obviously not local). The importance of macroscopic events will be emphasized in Section 6.8.

The reader might wonder whether there are interesting events that do not belong to $\mathscr{F}$. As a matter of fact, all events which we will need can be described explicitly *in terms of the individual spins in S, using (possibly infinite) unions and intersections. Those are all in $\mathscr{F}$.* ⋄

The set of probability measures on $(\Omega, \mathscr{F})$ will be denoted $\mathscr{M}_1(\Omega, \mathscr{F})$, or simply $\mathscr{M}_1(\Omega)$ when no ambiguity is possible. The elements of $\mathscr{M}_1(\Omega)$ will usually be denoted μ or ν.

A function $g : \Omega \to \mathbb{R}$ is **measurable with respect to** $\mathscr{F}_S$ (or simply $\mathscr{F}_S$**-measurable**) if $g^{-1}(I) \in \mathscr{F}_S$ for all Borel sets $I \subset \mathbb{R}$. Intuitively, such a function should be a *function* of the spins living in S:

Lemma 6.3. *A function $g : \Omega \to \mathbb{R}$ is $\mathscr{F}_S$-measurable if and only if there exists $\varphi : \Omega_S \to \mathbb{R}$ such that*

$$g(\omega) = \varphi(\omega_S)\,.$$

Proof. Let g be $\mathscr{F}_S$-measurable. On Ω_S, consider the set of cylinder events $\mathscr{C}'_S$, and $\mathscr{F}'_S = \sigma(\mathscr{C}'_S)$. If $\Pi_S : \Omega \to \Omega_S$ denotes the projection map, we have $\Pi_S^{-1}(C') \in \mathscr{F}_S$ for all $C' \in \mathscr{C}'_S$. This implies that $\mathscr{F}_S$ is generated by Π_S: $\mathscr{F}_S = \sigma(\Pi_S)$ (see Section B.5.2). Therefore, by Lemma B.38, there exists $\varphi : \Omega \to \Omega_S$ such that $g = \varphi \circ \Pi_S$. Conversely, if g is of this form, then clearly $g^{-1}(I) \in \mathscr{F}_S$ for each Borel set $I \subset \mathbb{R}$ so that g is $\mathscr{F}_S$-measurable. □

Remember that $f : \Omega \to \mathbb{R}$ is **local** if it only depends on a finite number of spins: there exists $\Lambda \Subset \mathbb{Z}^d$ such that $f(\omega) = f(\omega')$ as soon as $\omega_\Lambda = \omega'_\Lambda$. By Lemma 6.3, this is equivalent to saying that f is $\mathscr{F}_\Lambda$-measurable. In fact, since the spins take finitely many values, a local function can only take finitely many values and can therefore be expressed as a finite linear combination of indicators of cylinders. Since, for each of the latter, $f^{-1}(I) \in \mathscr{C} \subset \mathscr{F}$, local functions are always measurable. In the sequel, all the functions $f : \Omega \to \mathbb{R}$ which we will consider will be assumed to be measurable.

Notation 6.4. *In Chapter 3, we denoted the expectation of a function f under a probability measure μ by $\langle f \rangle_\mu$. For the rest of this chapter, it will be convenient to also use the following equivalent notations: $\int f \,\mathrm{d}\mu$, or $\mu(f)$.*

States vs. probability measures

Remember from Section 3.4 that a *state* is a normalized positive linear functional $f \mapsto \langle f \rangle$ acting on local functions. Observe that a state can be associated to each probability measure $\mu \in \mathscr{M}_1(\Omega)$ by setting, for all local functions f,

$$\langle f \rangle \stackrel{\text{def}}{=} \mu(f)\,.$$

It turns out that all states are of this form:

Theorem 6.5. *For every state $\langle\cdot\rangle$, there exists a unique probability measure $\mu \in \mathscr{M}_1(\Omega)$ such that $\langle f \rangle = \mu(f)$ for every local function $f : \Omega \to \mathbb{R}$.*

This result is a particular case of the Riesz–Markov–Kakutani Representation Theorem. Its proof requires a few tools that will be presented later, and can be found in Section 6.12.

Two infinite-volume measures for the Ising model

Using Theorem 6.5, we can associate a probability measure to each Gibbs state of the Ising model. In particular, let us denote by $\mu^+_{\beta,h}$ (resp. $\mu^-_{\beta,h}$) the measure associated to $\langle\cdot\rangle^+_{\beta,h}$ (resp. $\langle\cdot\rangle^-_{\beta,h}$). For these measures, relations such as (6.1) hold. A lot will be learned about these measures throughout the chapter.

For the time being, one should remember that the construction of $\mu^+_{\beta,h}$ and $\mu^-_{\beta,h}$ was based on the thermodynamic limit, which was used to define the states $\langle\cdot\rangle^+_{\beta,h}$ and $\langle\cdot\rangle^-_{\beta,h}$. Our aim, in the following sections, is to present a way of defining measures *directly* on the infinite lattice, without involving any limiting procedure. As we will see, this alternative approach presents a number of substantial advantages.

Why not simply use Kolmogorov's Extension Theorem?

In probability theory, the standard approach to construct infinite collections of dependent random variables relies on Kolmogorov's Extension Theorem, in which the strategy is to define a measure by requiring it to satisfy a set of local conditions. In our case, these conditions should depend on the microscopic description of the system under consideration, which is encoded in its Hamiltonian. We briefly outline this approach and explain why it does not solve the problem we are interested in.

Given $\mu \in \mathscr{M}_1(\Omega)$ and $\Lambda \Subset \mathbb{Z}^d$, the **marginal distribution of $\boldsymbol{\mu}$ on $\boldsymbol{\Lambda}$** is the probability distribution $\mu|_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$ defined by

$$\mu|_\Lambda \stackrel{\text{def}}{=} \mu \circ \Pi_\Lambda^{-1}. \tag{6.2}$$

In other words, $\mu|_\Lambda$ is the only distribution in $\mathscr{M}_1(\Omega_\Lambda)$ such that, for all $A \in \mathscr{P}(\Omega_\Lambda)$, $\mu|_\Lambda(A) = \mu(\{\omega \in \Omega : \omega_\Lambda \in A\})$. By construction, the marginals satisfy:

$$\mu|_\Delta = \mu|_\Lambda \circ (\Pi_\Delta^\Lambda)^{-1}, \quad \forall \Delta \subset \Lambda \Subset \mathbb{Z}^d, \tag{6.3}$$

where $\Pi_\Delta^\Lambda : \Omega_\Lambda \to \Omega_\Delta$ is the canonical projection defined by $\Pi_\Delta^\Lambda \stackrel{\text{def}}{=} \Pi_\Delta \circ \Pi_\Lambda^{-1}$.

It turns out that a measure $\mu \in \mathscr{M}_1(\Omega)$ is entirely characterized by its marginals $\mu|_\Lambda$, $\Lambda \Subset \mathbb{Z}^d$, but more is true: given any collection of probability distributions $\{\mu_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$, with $\mu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$ for all Λ, which satisfies a compatibility condition of the type (6.3), there exists a unique probability measure $\mu \in \mathscr{M}_1(\Omega)$ admitting them as marginals. This is the content of the following famous

Theorem 6.6. *[Kolmogorov's Extension Theorem] Let $\{\mu_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$, $\mu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$, be* ***consistent*** *in the sense that*

$$\text{for all } \Lambda \Subset \mathbb{Z}^d : \quad \mu_\Delta = \mu_\Lambda \circ (\Pi_\Delta^\Lambda)^{-1}, \quad \forall \Delta \subset \Lambda. \tag{6.4}$$

Then there exists a unique $\mu \in \mathscr{M}_1(\Omega)$ such that $\mu|_\Lambda = \mu_\Lambda$ for all $\Lambda \Subset \mathbb{Z}^d$.

Proof. See Section 6.12. □

Theorem 6.6 yields an efficient way of constructing a measure in $\mathscr{M}_1(\Omega)$, provided that one can define the desired collection $\{\mu_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ of candidates for its marginals. An important such application is the construction of the product measure, that is, of an **independent field**; in our setting, this covers for example the case of the Ising model at infinite temperature, $\beta = 0$.

Exercise 6.3. *(Construction of a product measure on $(\Omega, \mathscr{F})$) For each $i \in \mathbb{Z}^d$, let ρ_i be a probability distribution on $\{\pm 1\}$ and let, for all $\Lambda \Subset \mathbb{Z}^d$,*

$$\mu_\Lambda(\omega_\Lambda) \stackrel{\text{def}}{=} \prod_{j \in \Lambda} \rho_j(\omega_j), \qquad \omega_\Lambda \in \Omega_\Lambda.$$

Check that $\{\mu_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is consistent. The resulting measure on $(\Omega, \mathscr{F})$ whose existence is guaranteed by Theorem 6.6, is denoted $\rho^{\mathbb{Z}^d}$.

If one tries to use Theorem 6.6 to construct infinite-volume measures for the Ising model on $\mathbb{Z}^d$, we face a difficulty. Namely, the Boltzmann weight allows one to define finite-volume Gibbs distributions in terms of the underlying Hamiltonian. However, as we will explain now, *in general, there is no way to express the marginals associated to an infinite-volume Gibbs measure without making explicit reference to the latter.*

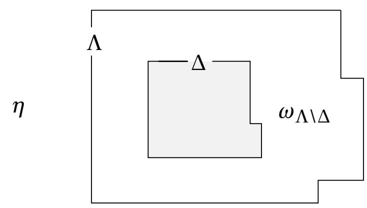
Indeed, let us consider the simplest case of the marginal distribution of the spin at the origin, σ_0, and let us assume that $d \geq 2$ and $h = 0$. Of course, σ_0 follows a Bernoulli distribution (with values in $\{\pm 1\}$) for some parameter $p \in [0,1]$. The only thing that needs to be determined is the value of p. However, we already know from the results in Chapter 3 that, for all large enough values of β, the average value of σ_0, and thus the relevant value of p, depends on the chosen Gibbs state. However, *all these states correspond to the same Hamiltonian and the same values of the parameters β and h.* This means that it is impossible to determine p from a knowledge of the Hamiltonian and the parameters β and h: one needs to know the macroscopic state the system is in, which is precisely what we are trying to construct. This shows that Kolmogorov's Extension Theorem is doomed to fail for the construction of the Ising model in infinite volume. [1]

Exercise 6.4. *Consider $\{\mu_\Lambda^\varnothing\}_{\Lambda \Subset \mathbb{Z}^d}$, where $\mu_\Lambda^\varnothing$ is the Gibbs distribution associated to the two-dimensional Ising model in Λ, with free boundary condition, at parameters $\beta > 0$ and $h = 0$. Show that the family obtained is not consistent.*

6.2.1 The DLR approach

A key observation, made by Dobrushin, Lanford and Ruelle is that if one considers *conditional probabilities* rather than marginals, then one is led to a different consistency condition, much better suited to our needs. Before stating this condition precisely (see Lemma 6.7 below), we explain it at an elementary level, using the Ising model and the notations of Chapter 3.

Consider $\Delta \subset \Lambda \Subset \mathbb{Z}^d$ and a boundary condition $\eta \in \Omega$:



The Ising model in Λ with boundary condition η is described by $\mu^\eta_{\Lambda;\beta,h}$. Let f be a local function depending only on the variables ω_j, $j \in \Delta$, and consider the expectation of f under $\mu^\eta_{\Lambda;\beta,h}$. Since f only depends on the spins located inside Δ, this expectation can be computed by first fixing the values of the spins in $\Lambda \setminus \Delta$. As we already saw in Exercise 3.11, $\mu^\eta_{\Lambda;\beta,h}$, conditioned on $\omega_{\Lambda\setminus\Delta}$, is equivalent to the Gibbs distribution on Δ with boundary condition $\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}$ outside Δ. Therefore,

$$\begin{aligned}\langle f\rangle^\eta_{\Lambda;\beta,h} &= \sum_{\omega_{\Lambda\setminus\Delta}} \langle f\,\mathbf{1}_{\{\omega_{\Lambda\setminus\Delta}\text{ outside }\Delta\}}\rangle^\eta_{\Lambda;\beta,h} \\ &= \sum_{\omega_{\Lambda\setminus\Delta}} \langle f\rangle^{\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}}_{\Delta;\beta,h}\, \mu^\eta_{\Lambda;\beta,h}(\omega_{\Lambda\setminus\Delta}\text{ outside }\Delta)\,. \qquad (6.5)\end{aligned}$$

(Notice a slight abuse of notation in the last line.) A particular instance of (6.5) is when f is the indicator of some event A occurring in Δ, in which case

$$\mu^{\eta}_{\Lambda;\beta,h}(A) = \sum_{\omega_{\Lambda\setminus\Delta}} \mu^{\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}}_{\Delta;\beta,h}(A)\,\mu^{\eta}_{\Lambda;\beta,h}(\omega_{\Lambda\setminus\Delta} \text{ outside } \Delta)\,. \tag{6.6}$$

The above discussion expresses the idea of Dobrushin, Lanford and Ruelle: the relation (6.5), or its second equivalent version (6.6), can be interpreted as a consistency relation between the Gibbs distributions in Λ and Δ. We can formulate (6.5) in a more precise way:

Lemma 6.7. *For all $\Delta \subset \Lambda \Subset \mathbb{Z}^d$ and all bounded measurable $f : \Omega \to \mathbb{R}$,*

$$\langle f\rangle^{\eta}_{\Lambda;\beta,h} = \big\langle \langle f\rangle^{\cdot}_{\Delta;\beta,h}\big\rangle^{\eta}_{\Lambda;\beta,h}\,, \qquad \forall \eta \in \Omega\,. \tag{6.7}$$

Proof of Lemma 6.7. To lighten the notations, we omit any mention of the dependence on β and h. Each $\omega \in \Omega^{\eta}_{\Lambda}$ is of the form $\omega = \omega_\Lambda \eta_{\Lambda^c}$, with $\omega_\Lambda \in \Omega_\Lambda$. Therefore,

$$\big\langle \langle f\rangle^{\cdot}_{\Delta}\big\rangle^{\eta}_{\Lambda} = \sum_{\omega_\Lambda} \langle f\rangle^{\omega_\Lambda\eta_{\Lambda^c}}_{\Delta} \frac{e^{-\mathscr{H}_\Lambda(\omega_\Lambda\eta_{\Lambda^c})}}{\mathbf{Z}^{\eta}_{\Lambda}}\,. \tag{6.8}$$

In the same way,

$$\langle f\rangle^{\omega_\Lambda\eta_{\Lambda^c}}_{\Delta} = \sum_{\omega'_\Delta} f(\omega'_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}) \frac{e^{-\mathscr{H}_\Delta(\omega'_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})}}{\mathbf{Z}^{\omega_\Lambda\eta_{\Lambda^c}}_{\Delta}}\,. \tag{6.9}$$

In (6.8), we decompose $\omega_\Lambda = \omega_\Delta\omega_{\Lambda\setminus\Delta}$, and sum separately over $\omega_{\Lambda\setminus\Delta}$ and ω_Δ. Observe that

$$\mathscr{H}_\Lambda(\omega_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}) - \mathscr{H}_\Delta(\omega_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}) = \mathscr{H}_\Lambda(\omega'_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}) - \mathscr{H}_\Delta(\omega'_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})\,. \tag{6.10}$$

Indeed, the difference on each side represents the interactions among the spins inside $\Lambda\setminus\Delta$, and between these spins and those outside Λ, and so does not depend on ω_Δ or ω'_Δ. Therefore, plugging (6.9) into (6.8), using (6.10), rearranging and calling $\omega'_\Delta\omega_{\Lambda\setminus\Delta} \equiv \omega'_\Lambda$, we get

$$\begin{aligned}
\big\langle \langle f\rangle^{\cdot}_{\Delta}\big\rangle^{\eta}_{\Lambda} &= \sum_{\omega_{\Lambda\setminus\Delta}} \sum_{\omega'_\Delta} f(\omega'_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c}) \frac{e^{-\mathscr{H}_\Lambda(\omega'_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})}}{\mathbf{Z}^{\eta}_{\Lambda}} \underbrace{\frac{\sum_{\omega_\Delta} e^{-\mathscr{H}_\Delta(\omega_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})}}{\mathbf{Z}^{\omega_\Lambda\eta_{\Lambda^c}}_{\Delta}}}_{=1} \\
&= \sum_{\omega'_\Lambda} f(\omega'_\Lambda\eta_{\Lambda^c}) \frac{e^{-\mathscr{H}_\Lambda(\omega'_\Lambda\eta_{\Lambda^c})}}{\mathbf{Z}^{\eta}_{\Lambda}} \\
&= \langle f\rangle^{\eta}_{\Lambda}\,.
\end{aligned}$$

□

Remark 6.8. The proof given above does not depend on the details of the Ising Hamiltonian, but rather on the property (6.10), which will be used again later. ⋄

We now explain why (6.7) leads to a natural characterization of *infinite-volume Gibbs states*, more general than the one introduced in Chapter 3.

First observe that, since we are considering the Ising model in which the interactions are only between nearest neighbors, the function $\omega \mapsto \langle f\rangle^{\omega}_{\Delta;\beta,h}$ is local (it depends only on those ω_i for which $i \in \partial^{\mathrm{ex}}\Delta$). So, *if* the distributions $\langle\cdot\rangle^{\eta}_{\Lambda;\beta,h}$ converge to a Gibbs state $\langle\cdot\rangle$ when $\Lambda \uparrow \mathbb{Z}^d$, in the sense of Definition 3.14, then we can take the thermodynamic limit on both sides of (6.7), obtaining

$$\langle f\rangle = \big\langle \langle f\rangle^{\cdot}_{\Delta;\beta,h}\big\rangle, \tag{6.11}$$

for all $\Delta \Subset \mathbb{Z}^d$ and all local functions f. We conclude that (6.11) must be satisfied by all states $\langle\cdot\rangle$ obtained as limits. But this can also be used to characterize states *without reference to limits*. Namely, we could extend the notion of infinite-volume Gibbs state by saying that *a state* $\langle\cdot\rangle$ *(not necessarily obtained as a limit) is an infinite-volume Gibbs state for the Ising model at* (β,h) *if* (6.11) *holds for every* $\Delta \Subset \mathbb{Z}^d$ *and all local functions* f. This new characterization has mathematical advantages that will become clear later.

If one identifies a Gibbs state $\langle\cdot\rangle$ with the corresponding measure μ given in Theorem 6.5, then μ should satisfy the infinite-volume version of (6.6): by taking $f = \mathbf{1}_A$, for some local event A, (6.11) becomes

$$\mu(A) = \int \mu^{\omega}_{\Delta;\beta,h}(A)\mu(\mathrm{d}\omega)\,. \tag{6.12}$$

Once again, we can use (6.12) as a set of conditions that *define* those measures that describe the Ising model in infinite volume. We will say that $\mu \in \mathscr{M}_1(\Omega)$ *is a Gibbs measure for the parameters* (β,h) *if* (6.12) *holds for all* $\Delta \Subset \mathbb{Z}^d$ *and all local events* A. An important feature of this point of view is that it characterizes probability measures directly on the infinite lattice $\mathbb{Z}^d$, without assuming them being obtained from a limiting procedure.

This characterization of probability measures for infinite statistical mechanical systems, and the study of their properties, is often called the *DLR formalism*. In Section 6.3, we establish the mathematical framework in which this formalism can be conveniently developed.

6.3 Specifications and measures

We will formulate the DLR approach introduced in the previous section in a more precise and more general way. The theory will apply to a large class of models, containing the Ising model as a particular case. It will also include models with a more complex structure, for example with long-range interactions or interactions between larger collections of spins.

We will proceed in two steps. First, we will generalize the consistency relation (6.7) by introducing the notion of *specification*.

In our discussion of the Ising model, the starting ingredient was the family of finite-volume Gibbs distributions $\{\mu^{\cdot}_{\Lambda;\beta,h}(\cdot)\}_{\Lambda\Subset\mathbb{Z}^d}$, whose main features we gather as follows:

1. For a fixed boundary condition ω, $\mu^{\omega}_{\Lambda;\beta,h}(\cdot)$ is a probability distribution on $(\Omega^{\omega}_{\Lambda}, \mathscr{P}(\Omega^{\omega}_{\Lambda}))$. It can however also be seen as a probability measure on $(\Omega,\mathscr{F})$ by letting, for all $A \in \mathscr{F}$,

$$\mu^{\omega}_{\Lambda;\beta,h}(A) \stackrel{\text{def}}{=} \sum_{\tau_\Lambda\in\Omega_\Lambda} \mu^{\omega}_{\Lambda;\beta,h}(\tau_\Lambda\omega_{\Lambda^c})\mathbf{1}_A(\tau_\Lambda\omega_{\Lambda^c})\,. \tag{6.13}$$

In particular,

$$\forall B \in \mathscr{F}_{\Lambda^c}, \quad \mu^{\omega}_{\Lambda;\beta,h}(B) = \mathbf{1}_B(\omega). \tag{6.14}$$

2. For a fixed $A \in \mathscr{F}$, $\mu^{\omega}_{\Lambda;\beta,h}(A)$ is entirely determined by ω_{Λ^c} (actually, even by $\omega_{\partial^{\text{ex}}\Lambda}$). In particular, $\omega \mapsto \mu^{\omega}_{\Lambda;\beta,h}(A)$ is $\mathscr{F}_{\Lambda^c}$-measurable.

3. When considering regions $\Delta \subset \Lambda \Subset \mathbb{Z}^d$, the consistency condition (6.7) is satisfied.

The maps $\mu^{\cdot}_{\Lambda;\beta,h}(\cdot)$ depend of course on the specific form of the Hamiltonian of the Ising model, but the three properties above can in fact be introduced without reference to any particular Hamiltonian. In a fixed volume, we start by incorporating the first two features in a general definition:

Definition 6.9. *Let $\Lambda \Subset \mathbb{Z}^d$. A* ***probability kernel from $\mathscr{F}_{\Lambda^c}$ to $\mathscr{F}$*** *is a map $\pi_\Lambda : \mathscr{F} \times \Omega \to [0,1]$ with the following properties:*

- *For each $\omega \in \Omega$, $\pi_\Lambda(\cdot\,|\,\omega)$ is a probability measure on $(\Omega, \mathscr{F})$.*
- *For each $A \in \mathscr{F}$, $\pi_\Lambda(A\,|\,\cdot)$ is $\mathscr{F}_{\Lambda^c}$-measurable.*

If, moreover,

$$\pi_\Lambda(B\,|\,\omega) = \mathbf{1}_B(\omega), \qquad \forall B \in \mathscr{F}_{\Lambda^c} \tag{6.15}$$

for all $\omega \in \Omega$, π_Λ is said to be ***proper***.

Note that, if π_Λ is a proper probability kernel from $\mathscr{F}_{\Lambda^c}$ to $\mathscr{F}$, then the probability measure $\pi_\Lambda(\cdot\,|\,\omega)$ is concentrated on the set $\Omega^{\omega}_{\Lambda}$. Indeed, for any $\omega \in \Omega$,

$$\pi_\Lambda(\Omega^{\omega}_{\Lambda}\,|\,\omega) = \mathbf{1}_{\Omega^{\omega}_{\Lambda}}(\omega) = 1, \tag{6.16}$$

since $\Omega^{\omega}_{\Lambda} \in \mathscr{F}_{\Lambda^c}$. For this reason, we will call ω the **boundary condition** of $\pi_\Lambda(\cdot\,|\,\omega)$. Our first example of a proper probability kernel was thus $(A,\omega) \mapsto \mu^{\omega}_{\Lambda;\beta,h}(A)$, defined in (6.13).

For a fixed boundary condition ω, a bounded measurable function $f : \Omega \to \mathbb{R}$ can be integrated with respect to $\pi_\Lambda(\cdot\,|\,\omega)$. We denote by $\pi_\Lambda f$ the $\mathscr{F}_{\Lambda^c}$-measurable function defined by

$$\pi_\Lambda f(\omega) \stackrel{\text{def}}{=} \int f(\eta)\pi_\Lambda(\mathrm{d}\eta\,|\,\omega).$$

Although this integral notation is convenient, our assumption on the finiteness of Ω_0 implies that most of the integrals that will appear in this chapter are actually finite sums. Indeed, we will always work with proper probability kernels and the observation (6.16) implies that the measure $\pi_\Lambda(\cdot\,|\,\omega)$ is entirely characterized by the probability it associates to the configurations in the finite set $\Omega^{\omega}_{\Lambda}$. In particular, we can verify that π_Λ is proper if and only if it is of the form (6.13). Namely, using (6.16), one can compute the probability of any event $A \in \mathscr{F}$ by summing over the configurations in $\Omega^{\omega}_{\Lambda}$:

$$\pi_\Lambda(A\,|\,\omega) = \sum_{\eta \in \Omega^{\omega}_{\Lambda}} \pi_\Lambda(\{\eta\}\,|\,\omega)\mathbf{1}_A(\eta).$$

Since each $\eta \in \Omega^{\omega}_{\Lambda}$ is of the form $\eta = \eta_\Lambda\omega_{\Lambda^c}$, this sum can equivalently be expressed as

$$\pi_\Lambda(A\,|\,\omega) = \sum_{\eta_\Lambda \in \Omega_\Lambda} \pi_\Lambda(\{\eta_\Lambda\omega_{\Lambda^c}\}\,|\,\omega)\mathbf{1}_A(\eta_\Lambda\omega_{\Lambda^c}).$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

In the sequel, *all kernels π_Λ to be considered will be proper*, which, by the above discussion, means that π_Λ is entirely defined by the numbers $\pi_\Lambda(\{\eta_\Lambda\omega_{\Lambda^c}\}\,|\,\omega)$. To lighten the notations, we will abbreviate

$$\pi_\Lambda(\{\eta_\Lambda\omega_{\Lambda^c}\}\,|\,\omega) \equiv \pi_\Lambda(\eta_\Lambda\,|\,\omega)\,.$$

These sums will be used constantly throughout the chapter. We summarize this discussion in the following statement.

Lemma 6.10. *If π_Λ is proper, then, for all $\omega \in \Omega$,*

$$\pi_\Lambda(A\,|\,\omega) = \sum_{\eta_\Lambda\in\Omega_\Lambda} \pi_\Lambda(\eta_\Lambda\,|\,\omega)\mathbf{1}_A(\eta_\Lambda\omega_{\Lambda^c})\,, \qquad \forall A \in \mathscr{F}\,, \tag{6.17}$$

and, for any bounded measurable function $f : \Omega \to \mathbb{R}$,

$$\pi_\Lambda f(\omega) = \sum_{\eta_\Lambda\in\Omega_\Lambda} \pi_\Lambda(\eta_\Lambda\,|\,\omega) f(\eta_\Lambda\omega_{\Lambda^c})\,. \tag{6.18}$$

In order to describe an infinite system on $\mathbb{Z}^d$, we will actually need a *family* of proper probability kernels, $\{\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$, satisfying consistency relations of the type (6.6)–(6.7). These consistency relations are conveniently expressed in terms of the **composition of kernels**: given π_Λ and π_Δ, set

$$\pi_\Lambda\pi_\Delta(A\,|\,\eta) \stackrel{\text{def}}{=} \int \pi_\Delta(A\,|\,\omega)\pi_\Lambda(\mathrm{d}\omega\,|\,\eta)\,.$$

Exercise 6.5. *Let $\Delta \subset \Lambda \Subset \mathbb{Z}^d$. Show that $\pi_\Lambda\pi_\Delta$ is a proper probability kernel from $\mathscr{F}_{\Lambda^c}$ to $\mathscr{F}$.*

In these terms, the generalization of (6.6) can be stated as follows.

Definition 6.11. *A* ***specification*** *is a family $\pi = \{\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ of proper probability kernels that is* ***consistent****, in the sense that*

$$\pi_\Lambda\pi_\Delta = \pi_\Lambda \quad \forall \Delta \subset \Lambda \Subset \mathbb{Z}^d\,.$$

In order to formulate an analogue of (6.12) for probability kernels, it is natural to define, for every kernel π_Λ and every $\mu \in \mathscr{M}_1(\Omega)$, the probability measure $\mu\pi_\Lambda \in \mathscr{M}_1(\Omega)$ via

$$\mu\pi_\Lambda(A) \stackrel{\text{def}}{=} \int \pi_\Lambda(A\,|\,\omega)\,\mu(\mathrm{d}\omega)\,, \quad A \in \mathscr{F}\,. \tag{6.19}$$

Exercise 6.6. *Show that, for every bounded measurable function f, every measure $\mu \in \mathscr{M}_1(\Omega)$ and every kernel π_Λ, $\mu\pi_\Lambda(f) = \mu(\pi_\Lambda f)$.* Hint: *start with $f = \mathbf{1}_A$.*

With a specification at hand, we can now introduce the central definition of this chapter. Expression (6.20) below is the generalization of (6.12).

Definition 6.12. *Let $\pi = \{\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ be a specification. A measure $\mu \in \mathscr{M}_1(\Omega)$ is said to be* ***compatible with*** *(or* ***specified by****) π if*

$$\mu = \mu\pi_\Lambda \quad \forall \Lambda \Subset \mathbb{Z}^d\,. \tag{6.20}$$

The set of measures compatible with π (if any) is denoted by $\mathscr{G}(\pi)$.

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

The above characterization raises several questions, which we shall investigate in quite some generality in the rest of this chapter.

- *Existence.* Is there always at least one measure μ satisfying (6.20)? This problem will be tackled in Section 6.4.

- *Uniqueness.* Can there be several such measures? The uniqueness problem will be considered in Section 6.5, where we will introduce a condition on a specification π which guarantees that $\mathscr{G}(\pi)$ contains exactly one probability measure: $|\mathscr{G}(\pi)| = 1$.

- *Comparison with the former approach.* We will also consider the important question of comparing the approach based on Definition 6.12 with the approach used in Chapter 3, in which infinite-volume states were obtained as the thermodynamic limits of finite-volume ones. We will see that Definition 6.12 yields, in general, a strictly larger set of measures than those produced by the approach via the thermodynamic limit (proof of Theorem 6.26 and Example 6.64). Nevertheless, all the relevant (in a sense to be discussed later) measures in $\mathscr{G}(\pi)$ can in fact be obtained using the latter approach (Section 6.8).

When the specification does not involve interactions between the spins, these questions can be answered easily:

Exercise 6.7. *For each $i \in \mathbb{Z}^d$, let ρ_i be a probability distribution on $\{\pm 1\}$. For each $\Lambda \Subset \mathbb{Z}^d$, define the product distribution ρ^Λ on Ω_Λ by*

$$\rho^\Lambda(\omega_\Lambda) \stackrel{\text{def}}{=} \prod_{i\in\Lambda} \rho_i(\omega_i)\,.$$

For $\tau_\Lambda \in \Omega_\Lambda$ and $\eta \in \Omega$, let

$$\pi_\Lambda(\tau_\Lambda \,|\, \eta) \stackrel{\text{def}}{=} \rho^\Lambda(\tau_\Lambda)\,. \tag{6.21}$$

1. *Show that $\pi = \{\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ is a specification.*

2. *Show that the product measure $\rho^{\mathbb{Z}^d}$ (remember Exercise 6.3) is the unique probability measure specified by π: $\mathscr{G}(\pi) = \{\rho^{\mathbb{Z}^d}\}$.*

In the previous exercise, establishing existence and uniqueness of a probability measure compatible with the specification (6.21) is straightforward, thanks to the independence of the spins. In the next sections, we will introduce a general procedure for constructing specifications corresponding to systems of interacting spins and we will see that existence/uniqueness can be derived for abstract specifications under fairly general assumptions. (Establishing non-uniqueness, on the other hand, usually requires a case-by-case study.)

6.3.1 Kernels vs. conditional probabilities

Before continuing, we emphasize the important relation existing between a specification and the measures it specifies (if any). We first verify the following simple property:

Lemma 6.13. *Assume that π_Λ is proper. Then, for all $A \in \mathscr{F}$ and all $B \in \mathscr{F}_{\Lambda^c}$,*

$$\pi_\Lambda(A \cap B\,|\,\cdot) = \pi_\Lambda(A\,|\,\cdot)\mathbf{1}_B(\cdot)\,. \tag{6.22}$$

Proof. Assume first that $\omega \in B$. Then, since the kernel is proper, B has probability 1 under $\pi_\Lambda(\cdot\,|\,\omega)$: $\pi_\Lambda(B\,|\,\omega) = \mathbf{1}_B(\omega) = 1$. Therefore

$$\pi_\Lambda(A \cap B\,|\,\omega) = \pi_\Lambda(A\,|\,\omega) - \pi_\Lambda(A \cap B^c\,|\,\omega) = \pi_\Lambda(A\,|\,\omega) = \pi_\Lambda(A\,|\,\omega)\mathbf{1}_B(\omega)\,.$$

Similarly, if $\omega \notin B$, $\pi_\Lambda(B\,|\,\omega) = 0$ and thus

$$\pi_\Lambda(A \cap B\,|\,\omega) = 0 = \pi_\Lambda(A\,|\,\omega)\mathbf{1}_B(\omega)\,.$$ □

Now, observe that if $\mu \in \mathscr{G}(\pi)$, then (6.22) implies that, for all $A \in \mathscr{F}_\Lambda$ and $B \in \mathscr{F}_{\Lambda^c}$,

$$\int_B \pi_\Lambda(A\,|\,\omega)\mu(\mathrm{d}\omega) = \int \pi_\Lambda(A \cap B\,|\,\omega)\mu(\mathrm{d}\omega) = \mu\pi_\Lambda(A \cap B) = \mu(A \cap B)\,.$$

But, by definition of the conditional probability,

$$\mu(A \cap B) = \int_B \mu(A\,|\,\mathscr{F}_{\Lambda^c})(\omega)\mu(\mathrm{d}\omega)\,.$$

By the almost sure uniqueness of the conditional expectation (Lemma B.50), we thus see that

$$\mu(A\,|\,\mathscr{F}_{\Lambda^c})(\cdot) = \pi_\Lambda(A\,|\,\cdot)\,, \quad \mu\text{-almost surely.} \tag{6.23}$$

Since $A \mapsto \pi_\Lambda(A\,|\,\omega)$ is a measure for *each* ω, we thus see that π_Λ *provides a regular conditional distribution for* μ*, when conditioned with respect to* $\mathscr{F}_{\Lambda^c}$. On the other hand, if (6.23) holds, then, for all $\Lambda \Subset \mathbb{Z}^d$ and all $A \in \mathscr{F}$,

$$\mu\pi_\Lambda(A) = \int \pi_\Lambda(A\,|\,\omega)\mu(\mathrm{d}\omega) = \int \mu(A\,|\,\mathscr{F}_{\Lambda^c})\mu(\mathrm{d}\omega) = \mu(A)\,,$$

and so $\mu \in \mathscr{G}(\pi)$. We have thus shown that *a measure* μ *is compatible with a specification* $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ *if and only if each kernel* π_Λ *provides a regular version of* $\mu(\cdot\,|\,\mathscr{F}_{\Lambda^c})$.

6.3.2 Gibbsian specifications

Before moving on to the existence problem, we introduce the class of specifications representative of the models studied in this book.

The Ising Hamiltonian $\mathscr{H}_{\Lambda;\beta,h}$ (see (3.1)*) contains two sums: the first one is over pairs of nearest-neighbors $\{i,j\} \in \mathscr{E}^{\mathrm{b}}_\Lambda$, the second one is over single vertices $i \in \Lambda$. It thus contains interactions among pairs, and singletons. This structure can be generalized, including interactions among spins on sets of larger (albeit finite) cardinality.* ⋄

We will define a Hamiltonian by defining the energy of a configuration on each subset $B \Subset \mathbb{Z}^d$, via the notion of *potential*.

Definition 6.14. *If, for each finite* $B \Subset \mathbb{Z}^d$, $\Phi_B : \Omega \to \mathbb{R}$ *is* $\mathscr{F}_B$*-measurable, then the collection* $\Phi = \{\Phi_B\}_{B \Subset \mathbb{Z}^d}$ *is called a* ***potential****. The* ***Hamiltonian*** *in the box* $\Lambda \Subset \mathbb{Z}^d$ *associated to the potential* Φ *is defined by*

$$\mathscr{H}_{\Lambda;\Phi}(\omega) \stackrel{\text{def}}{=} \sum_{\substack{B \Subset \mathbb{Z}^d: \\ B \cap \Lambda \neq \varnothing}} \Phi_B(\omega), \qquad \forall \omega \in \Omega. \tag{6.24}$$

Since the sum (6.24) can a priori contain infinitely many terms, we must guarantee that it converges. Let

$$r(\Phi) \stackrel{\text{def}}{=} \inf\bigl\{R > 0 : \Phi_B \equiv 0 \text{ for all } B \text{ with } \operatorname{diam}(B) > R\bigr\}.$$

If $r(\Phi) < \infty$, Φ has **finite range** and $\mathscr{H}_{\Lambda;\Phi}$ is well defined. If $r(\Phi) = \infty$, Φ has **infinite range** and, for the Hamiltonian to be well defined, we will assume that Φ is **absolutely summable** in the sense that

$$\sum_{\substack{B \Subset \mathbb{Z}^d \\ B \ni i}} \|\Phi_B\|_\infty < \infty, \quad \forall i \in \mathbb{Z}^d, \tag{6.25}$$

(remember that $\|f\|_\infty \stackrel{\text{def}}{=} \sup_\omega |f(\omega)|$) which ensures that the interaction of a spin with the rest of the system is always bounded, and therefore that $\|\mathscr{H}_{\Lambda;\Phi}\|_\infty < \infty$.

We now present a few examples of models discussed in this book with the corresponding potentials.

- The (nearest-neighbor) **Ising model** on $\mathbb{Z}^d$ can be recovered from the potential

$$\Phi_B(\omega) = \begin{cases} -\beta\omega_i\omega_j & \text{if } B = \{i,j\},\, i \sim j, \\ -h\omega_i & \text{if } B = \{i\}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.26}$$

  Observe that the corresponding specification describes a model at specific values of its parameters: in the present case, we get a different specification for each choice of the parameters β and h.

  One can introduce an infinite-range version of the Ising model, by introducing a collection $\{J_{ij}\}_{i,j \in \mathbb{Z}^d}$ of real numbers and setting

$$\Phi_B(\omega) = \begin{cases} -J_{ij}\omega_i\omega_j & \text{if } B = \{i,j\}, \\ -h\omega_i & \text{if } B = \{i\}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.27}$$

- The (nearest-neighbor) ***q*-state Potts model** corresponds to the potential

$$\Phi_B(\omega) = \begin{cases} -\beta\delta_{\omega_i,\omega_j} & \text{if } B = \{i,j\},\, i \sim j, \\ 0 & \text{otherwise.} \end{cases} \tag{6.28}$$

- The (nearest-neighbor) **Blume–Capel model** is characterized by the potential

$$\Phi_B(\omega) = \begin{cases} \beta(\omega_i - \omega_j)^2 & \text{if } B = \{i,j\},\, i \sim j, \\ -h\omega_i - \lambda\omega_i^2 & \text{if } B = \{i\}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.29}$$

  This model will be studied in Chapter 7.

Exercise 6.8. *If $J_{ij} = \|j-i\|_\infty^{-\alpha}$, determine the values of $\alpha > 0$ (depending on the dimension) for which* (6.27) *is absolutely summable.*

In the above examples, the parameters of each model have been introduced according to different sets B. Sometimes, one might want the inverse temperature to be introduced separately, so as to appear as a multiplicative constant in front of the Hamiltonian. This amounts to considering an absolutely summable potential $\Phi = \{\Phi_B\}_{B \Subset \mathbb{Z}^d}$, and to then multiply it by β: $\beta\Phi \equiv \{\beta\Phi_B\}_{B\Subset\mathbb{Z}^d}$.

We now proceed to define a specification $\pi^\Phi = \{\pi^\Phi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ such that $\pi^\Phi_\Lambda(\cdot\,|\,\omega)$ gives to each configuration $\tau_\Lambda\omega_{\Lambda^c}$ a probability proportional to the Boltzmann weight prescribed by equilibrium statistical mechanics:

$$\pi^\Phi_\Lambda(\tau_\Lambda|\omega) \stackrel{\text{def}}{=} \frac{1}{\mathbf{Z}^\omega_{\Lambda;\Phi}} e^{-\mathscr{H}_{\Lambda;\Phi}(\tau_\Lambda\omega_{\Lambda^c})}, \tag{6.30}$$

where we have written explicitly the dependence on ω_{Λ^c}, and where the partition function $\mathbf{Z}^\omega_{\Lambda;\Phi}$ is given by

$$\mathbf{Z}^\omega_{\Lambda;\Phi} \stackrel{\text{def}}{=} \sum_{\tau_\Lambda\in\Omega_\Lambda} \exp(-\mathscr{H}_{\Lambda;\Phi}(\tau_\Lambda\omega_{\Lambda^c})). \tag{6.31}$$

Lemma 6.15. *$\pi^\Phi = \{\pi^\Phi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ is a specification.*

Proof. To lighten the notations, let us omit Φ everywhere from the notations. It will also help to change momentarily the way we denote partition functions, namely, in this proof, we will write

$$\mathbf{Z}_\Lambda(\omega_{\Lambda^c}) \equiv \mathbf{Z}^\omega_{\Lambda;\Phi}.$$

The fact that each π_Λ defines a proper kernel follows by what was said earlier, so it remains to verify consistency. We fix $\Delta \subset \Lambda \Subset \mathbb{Z}^d$, and show that $\pi_\Lambda\pi_\Delta = \pi_\Lambda$. The proof follows the same steps as the one of Lemma 6.7. Using Lemma 6.10,

$$\begin{aligned}\pi_\Lambda\pi_\Delta(A\,|\,\omega) &= \sum_{\tau_\Lambda} \pi_\Lambda(\tau_\Lambda\,|\,\omega)\pi_\Delta(A\,|\,\tau_\Lambda\omega_{\Lambda^c})\\ &= \sum_{\tau_\Lambda}\sum_{\eta_\Delta} \mathbf{1}_A(\eta_\Delta\tau_{\Lambda\setminus\Delta}\omega_{\Lambda^c})\pi_\Lambda(\tau_\Lambda\,|\,\omega)\pi_\Delta(\eta_\Delta\,|\,\tau_{\Lambda\setminus\Delta}\omega_{\Lambda^c}).\end{aligned}$$

We split the first sum in two, writing $\tau_\Lambda = \tau'_\Delta\tau''_{\Lambda\setminus\Delta}$. Using the definition of the kernels π_Λ and π_Δ, the above becomes

$$\sum_{\tau''_{\Lambda\setminus\Delta}}\sum_{\eta_\Delta} \mathbf{1}_A(\eta_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c}) \frac{e^{-\mathscr{H}_\Delta(\eta_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c})}}{\mathbf{Z}_\Lambda(\omega_{\Lambda^c})\mathbf{Z}_\Delta(\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c})} \sum_{\tau'_\Delta} e^{-\mathscr{H}_\Lambda(\tau'_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c})}.$$

But, exactly as in (6.10),

$$\mathscr{H}_\Lambda(\tau'_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c}) - \mathscr{H}_\Delta(\tau'_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c}) = \mathscr{H}_\Lambda(\eta_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c}) - \mathscr{H}_\Delta(\eta_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c}),$$

which gives

$$\sum_{\tau'_\Delta} e^{-\mathscr{H}_\Lambda(\tau'_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c})} = \mathbf{Z}_\Delta(\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c}) e^{-\mathscr{H}_\Lambda(\eta_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c})} e^{\mathscr{H}_\Delta(\eta_\Delta\tau''_{\Lambda\setminus\Delta}\omega_{\Lambda^c})}.$$

Inserting this in the above expression, and renaming $\eta_\Delta\tau''_{\Lambda\setminus\Delta} \equiv \eta'_\Lambda$, we get

$$\pi_\Lambda\pi_\Delta(A\,|\,\omega) = \sum_{\eta'_\Lambda} \mathbf{1}_A(\eta'_\Lambda\omega_{\Lambda^c}) \frac{e^{-\mathscr{H}_\Lambda(\eta'_\Lambda\omega_{\Lambda^c})}}{\mathbf{Z}_\Lambda(\omega_{\Lambda^c})} = \pi_\Lambda(A\,|\,\omega). \qquad \square$$

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

We can now state the general definition of a Gibbs measure.

Definition 6.16. *The specification π^Φ associated to a potential Φ is said to be* ***Gibbsian****. A probability measure μ compatible with the Gibbsian specification π^Φ is said to be an* ***infinite-volume Gibbs measure*** *(or simply a* ***Gibbs measure****) associated to the potential Φ.*

It is customary to use the abbreviation $\mathscr{G}(\Phi) \equiv \mathscr{G}(\pi^\Phi)$. Actually, when the potential is parametrized by a few variables, we will write them rather than Φ. For example, in the case of the (nearest-neighbor) Ising model, whose specification depends on β and h, we will simply write $\mathscr{G}(\beta, h)$.

Remark 6.17. Notice that different potentials can lead to the same specification. For example, in the case of the Ising model, one could as well have considered the potential

$$\tilde{\Phi}_B(\omega) = \begin{cases} -\beta\omega_i\omega_j - \frac{h}{2d}(\omega_i + \omega_j) & \text{if } B = \{i,j\}, i \sim j, \\ 0 & \text{otherwise.} \end{cases}$$

Since they give rise to the same Hamiltonian, up to a term depending only on ω_{Λ^c}, these potentials also give rise to the same specification. They thus describe precisely the same physics. For this reason, they are said to be **physically equivalent**. ⋄

When introducing a model, it is often quite convenient, instead of giving the corresponding potential $\{\Phi_B\}_{B \Subset \mathbb{Z}^d}$, to provide its **formal Hamiltonian**

$$\mathscr{H}(\omega) \stackrel{\text{def}}{=} \sum_{B \Subset \mathbb{Z}^d} \Phi_B(\omega).$$

Of course, this notation is purely formal and does not specify a well-defined function on Ω. It is however possible to read from $\mathscr{H}$ the corresponding potential (up to physical equivalence).

As an example, the effective Hamiltonian of the Ising model on $\mathbb{Z}^d$ may be denoted by

$$-\beta \sum_{\{i,j\} \in \mathscr{E}_{\mathbb{Z}^d}} \sigma_i\sigma_j - h \sum_{i \in \mathbb{Z}^d} \sigma_i.$$

In view of what we saw in Chapter 3, the following is a natural definition of phase transition, in terms of *non-uniqueness* of the Gibbs measure:

Definition 6.18. *If $\mathscr{G}(\Phi)$ contains at least two distinct Gibbs measures, $|\mathscr{G}(\Phi)| > 1$, we say that there is a* ***first-order phase transition*** *for the potential Φ.*

6.4 Existence

Going back to the case of a general specification, we now turn to the problem of determining conditions that ensure the existence of at least one measure compatible with a given specification. As in many existence proofs in analysis and probability theory, this will be based on a *compactness* argument, and thus requires that we introduce a few topological notions. We will take advantage of the fact that the spins take values in a finite set to provide elementary proofs.

The approach is similar to the construction of Gibbs states in Chapter 3. We fix an arbitrary boundary condition $\omega \in \Omega$ and consider the sequence $(\mu_n)_{n\geq 1} \subset \mathscr{M}_1(\Omega)$ defined by

$$\mu_n(\cdot) \stackrel{\text{def}}{=} \pi_{\mathsf{B}(n)}(\cdot\,|\,\omega)\,, \tag{6.32}$$

where, as usual, $\mathsf{B}(n) = \{-n,\dots,n\}^d$. To study this sequence, we will first introduce a suitable notion of convergence for sequences of measures (Definition 6.23). This will make $\mathscr{M}_1(\Omega)$ *sequentially compact*; in particular, there always exists $\mu \in \mathscr{M}_1(\Omega)$ and a subsequence of $(\mu_n)_{n\geq 1}$, say $(\mu_{n_k})_{k\geq 1}$, such that $(\mu_{n_k})_{k\geq 1}$ converges to μ (Theorem 6.24). To guarantee that $\mu \in \mathscr{G}(\pi)$, we will impose a natural condition on π, called *quasilocality*.

6.4.1 Convergence on Ω

We first introduce a **topology** on Ω, that is, a notion of convergence for sequences of configurations.

Definition 6.19. *A sequence $\omega^{(n)} \in \Omega$ **converges** to $\omega \in \Omega$ if*

$$\lim_{n\to\infty} \omega_j^{(n)} = \omega_j\,, \qquad \forall j \in \mathbb{Z}^d.$$

We then write $\omega^{(n)} \to \omega$.

Since $\{\pm 1\}$ is a finite set, this convergence can be reformulated as follows: $\omega^{(n)} \to \omega$ if and only if, for all N, there exists n_0 such that

$$\omega^{(n)}_{\mathsf{B}(N)} = \omega_{\mathsf{B}(N)} \qquad \text{for all } n \geq n_0\,.$$

The notion of neighborhood in this topology should thus be understood as follows: two configurations are close to each other if they coincide on a large region containing the origin. The following exercise shows that this topology is *metrizable*.

Exercise 6.9. *For $\omega,\eta \in \Omega$, let*

$$d(\omega,\eta) \stackrel{\text{def}}{=} \sum_{i\in\mathbb{Z}^d} 2^{-\|i\|_\infty} \mathbf{1}_{\{\omega_i \neq \eta_i\}}\,. \tag{6.33}$$

Show that $d(\cdot,\cdot)$ is a distance on Ω, and that $\omega^{(n)} \to \omega^*$ if and only if $d(\omega^{(n)},\omega^*) \to 0$.

Another consequence of the finiteness of the spin space is that Ω is compact in the topology just introduced:

Proposition 6.20 (Compactness of Ω)**.** *With the above notion of convergence, Ω is **sequentially compact**: for every sequence $(\omega^{(n)})_{n\geq 1} \subset \Omega$, there exists $\omega^* \in \Omega$ and a subsequence $(\omega^{(n_k)})_{k\geq 1}$ such that $\omega^{(n_k)} \to \omega^*$ when $k \to \infty$.*

Proof. We use a standard diagonalization argument. Consider $(\omega^{(n)})_{n\geq 1} \subset \Omega$ and let $i_1, i_2, \dots$ be an arbitrary enumeration of $\mathbb{Z}^d$. Then $(\omega^{(n)}_{i_1})_{n\geq 1}$ is a sequence in $\{\pm 1\}$, from which we can extract a subsequence $(\omega^{(n_{1,j})}_{i_1})_{j\geq 1}$ which converges (in fact, it can be taken constant). We then consider $(\omega_{i_2}^{(n_{1,j})})_{j\ge1}$, from which we extract a converging subsequence $(\omega_{i_2}^{(n_{2,j})})_{j\ge1}$, etc., until we have, for each k, a converging subsequence $(\omega_{i_k}^{(n_{k,j})})_{j\ge1}$. Let $\omega^* \in \Omega$ be defined by

$$\omega^*_{i_k} \stackrel{\text{def}}{=} \lim_{j\to\infty} \omega_{i_k}^{(n_{k,j})}, \qquad \forall k \ge 1\,.$$

Now, the diagonal subsequence $(\omega^{(n_{j,j})})_{j\ge1}$ is a subsequence of $(\omega^{(n)})_{n\ge1}$ and satisfies $\omega^{(n_{j,j})} \to \omega^*$ as $j\to\infty$. □

We can now define a function $f : \Omega \to \mathbb{R}$ to be **continuous** if $\omega^{(n)} \to \omega$ implies $f(\omega^{(n)}) \to f(\omega)$. The set of continuous functions on Ω is denoted by $C(\Omega)$.

Exercise 6.10. *Show that each $f \in C(\Omega)$ is measurable.* Hint: *first show that $\mathscr{C} \subset \{\text{open sets}\} \subset \mathscr{F}$, where the open sets are those associated to the topology defined above.*

We say that f is **uniformly continuous** (see Appendix B.4) if

$$\forall\epsilon > 0, \text{ there exists } \delta > 0 \text{ such that } d(\omega,\eta) \le \delta \text{ implies } |f(\omega) - f(\eta)| \le \epsilon.$$

Exercise 6.11. *Using Proposition 6.20, give a direct proof of the following facts: if f is continuous, it is also uniformly continuous, bounded, and it attains its supremum and its infimum.*

Local functions are clearly continuous (since they do not depend on remote spins); they are in fact *dense* in $C(\Omega)$ [2]:

Lemma 6.21. *$f \in C(\Omega)$ if and only if it is **quasilocal**, that is, if and only if there exists a sequence of local functions $(g_n)_{n\ge1}$ such that $\|g_n - f\|_\infty \to 0$.*

Proof. Let $f : \Omega \to \mathbb{R}$ be continuous. Fix some $\epsilon > 0$. Since f is also uniformly continuous, there exists some $\Lambda \Subset \mathbb{Z}^d$ such that $|f(\omega) - f(\eta)| \le \epsilon$ for any pair η and ω coinciding on Λ. Therefore, if one chooses some arbitrary $\tilde\omega \in \Omega$ and introduces the local function $g(\omega) \stackrel{\text{def}}{=} f(\omega_\Lambda \tilde\omega_{\Lambda^c})$, we have that $|f(\omega) - g(\omega)| \le \epsilon\ \forall\omega \in \Omega$. Conversely, let $(g_n)_{n\ge1}$ be a sequence of local functions such that $\|g_n - f\|_\infty \to 0$. Fix $\epsilon > 0$ and let n be such that $\|g_n - f\|_\infty \le \epsilon$. Since g_n is uniformly continuous, let $\delta > 0$ be such that $d(\omega,\eta) \le \delta$ implies $|g_n(\omega) - g_n(\eta)| \le \epsilon$. For each such pair ω,η we also have

$$|f(\omega) - f(\eta)| \le |f(\omega) - g_n(\omega)| + |g_n(\omega) - g_n(\eta)| + |g_n(\eta) - f(\eta)| \le 3\epsilon\,.$$

Since this can be done for all $\epsilon > 0$, we have shown that $f \in C(\Omega)$. □

We will often use the fact that probability measures on $(\Omega,\mathscr{F})$ are uniquely determined by their action on cylinders, or by the value they associate to the expectation of local or continuous functions.

Lemma 6.22. *If $\mu,\nu \in \mathscr{M}_1(\Omega)$, then the following are equivalent:*

1. $\mu = \nu$
2. $\mu(C) = \nu(C)$ *for all cylinders* $C \in \mathscr{C}$.
3. $\mu(g) = \nu(g)$ *for all local functions* g.
4. $\mu(f) = \nu(f)$ *for all* $f \in C(\Omega)$.

Proof. 1⇒2 is trivial, and 2⇒1 is a consequence of the Uniqueness Theorem for measures (Corollary B.37). 2⇔3 is immediate, since the indicator of a cylinder is a local function. 3⇒4: Let $f \in C(\Omega)$ and let $(g_n)_{n\geq 1}$ be a sequence of local functions such that $\|g_n - f\|_\infty \to 0$ (Lemma 6.21). This implies $|\mu(g_n) - \mu(f)| \leq \|g_n - f\|_\infty \to 0$. Similarly, $|\nu(g_n) - \nu(f)| \to 0$. Therefore,

$$\mu(f) = \lim_{n\to\infty} \mu(g_n) = \lim_{n\to\infty} \nu(g_n) = \nu(f)\,.$$

Finally, 4⇒3 holds because local functions are continuous. □

6.4.2 Convergence on $\mathscr{M}_1(\Omega)$

The topology on $\mathscr{M}_1(\Omega)$ will be the following:

Definition 6.23. *A sequence $(\mu_n)_{n\geq 1} \subset \mathscr{M}_1(\Omega)$ **converges to** $\mu \in \mathscr{M}_1(\Omega)$ if*

$$\lim_{n\to\infty} \mu_n(C) = \mu(C)\,, \quad \text{for all cylinders } C \in \mathscr{C}\,.$$

We then write $\mu_n \Rightarrow \mu$.

The fact that the convergence of a sequence of measures is tested on local events (the cylinders) should remind the reader of the convergence encountered in Chapter 3 (Definition 3.14), where a similar notion of convergence was introduced to define Gibbs states.

Before pursuing, we let the reader check the following equivalent characterizations of convergence on $\mathscr{M}_1(\Omega)$.

Exercise 6.12. *Show the equivalence between:*

1. $\mu_n \Rightarrow \mu$
2. $\mu_n(f) \to \mu(f)$ *for all local functions f.*
3. $\mu_n(f) \to \mu(f)$ *for all $f \in C(\Omega)$.*
4. $\rho(\mu_n, \mu) \to 0$, *where we defined, for all $\mu, \nu \in \mathscr{M}_1(\Omega)$, the distance*

$$\rho(\mu,\nu) \stackrel{\text{def}}{=} \sup_{k\geq 1} \frac{1}{k} \max_{C\in\mathscr{C}(\mathrm{B}(k))} |\mu(C) - \nu(C)|\,.$$

Theorem 6.24 (Compactness of $\mathscr{M}_1(\Omega)$)**.** *With the above notion of convergence, $\mathscr{M}_1(\Omega)$ is sequentially compact: for every sequence $(\mu_n)_{n\geq 1} \subset \mathscr{M}_1(\Omega)$, there exist $\mu \in \mathscr{M}_1(\Omega)$ and a subsequence $(\mu_{n_k})_{k\geq 1}$ such that $\mu_{n_k} \Rightarrow \mu$ when $k \to \infty$.*

Since the proof of this result is similar, in spirit, to the one used in the proof of the compactness of Ω, we postpone it to Section 6.12.

6.4.3 Existence and quasilocality

We will see below that the following condition on a specification π guarantees that $\mathscr{G}(\pi) \neq \varnothing$.

Definition 6.25. *A specification $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is* ***quasilocal*** *if each kernel π_Λ is continuous with respect to its boundary condition. That is, if for all $C \in \mathscr{C}$, $\omega \mapsto \pi_\Lambda(C\,|\,\omega)$ is continuous.*

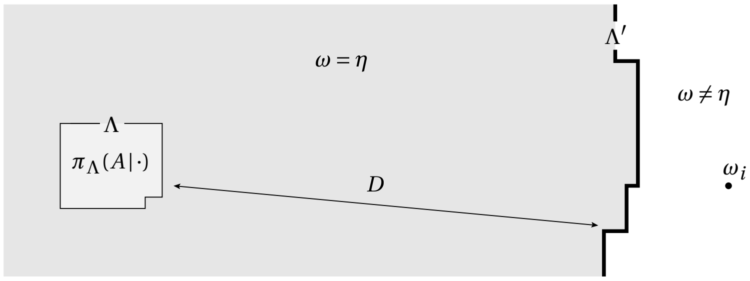


Figure 6.1: Understanding quasilocality: when D is large, $\pi_\Lambda(A\,|\,\omega)$ depends weakly on the values of ω_i for all i at distance larger than D from Λ (assuming all closer spins are fixed). In other words, for all $\epsilon > 0$, if ω and η coincide on a sufficiently large region $\Lambda' \supset \Lambda$, then $|\pi_\Lambda(A\,|\,\omega) - \pi_\Lambda(A\,|\,\eta)| \le \epsilon$.

The next exercise shows that quasilocal specifications map continuous (and, in particular, local) functions to continuous functions.

Exercise 6.13. *Let $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ be quasilocal and fix some $\Lambda \Subset \mathbb{Z}^d$. Show that $f \in C(\Omega)$ implies $\pi_\Lambda f \in C(\Omega)$. (This property is sometimes referred to as the* ***Feller property****.)*

We can now state the main existence theorem.

Theorem 6.26. *If $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is quasilocal, then $\mathscr{G}(\pi) \neq \varnothing$.*

Proof. Fix an arbitrary $\omega \in \Omega$ and let $\mu_n(\cdot) \stackrel{\text{def}}{=} \pi_{\mathsf{B}(n)}(\cdot\,|\,\omega)$. (One could also choose a different ω for each n.) Observe that, by the consistency assumption of the kernels forming π, we have that, once n is so large that $\mathsf{B}(n) \supset \Lambda$,

$$\mu_n \pi_\Lambda = \pi_{\mathsf{B}(n)} \pi_\Lambda(\cdot\,|\,\omega) = \pi_{\mathsf{B}(n)}(\cdot\,|\,\omega) = \mu_n \,. \tag{6.34}$$

By Theorem 6.24, there exist $\mu \in \mathscr{M}_1(\Omega)$ and a subsequence $(\mu_{n_k})_{k\ge 1}$ such that $\mu_{n_k} \Rightarrow \mu$ as $k \to \infty$. We prove that $\mu \in \mathscr{G}(\pi)$. Fix $f \in C(\Omega)$, $\Lambda \Subset \mathbb{Z}^d$. Since π is quasilocal, Exercise 6.13 shows that $\pi_\Lambda f \in C(\Omega)$. Therefore,

$$\mu\pi_\Lambda(f) = \mu(\pi_\Lambda f) = \lim_{k\to\infty} \mu_{n_k}(\pi_\Lambda f) = \lim_{k\to\infty} \mu_{n_k}\pi_\Lambda(f) = \lim_{k\to\infty} \mu_{n_k}(f) = \mu(f) \,.$$

We used Exercise 6.6 for the first and third identities. The fourth identity follows from (6.34). By Lemma 6.22, we conclude that $\mu\pi_\Lambda = \mu$. Since this holds for all $\Lambda \Subset \mathbb{Z}^d$, this shows that $\mu \in \mathscr{G}(\pi)$. □

Since ω should be interpreted as a boundary condition, a Gibbs measure μ constructed as in the above proof,

$$\pi_{\mathsf{B}(n_k)}(\cdot\,|\,\omega) \Rightarrow \mu \,,$$

is said to be **prepared with the boundary condition ω**. A priori, μ can depend on the chosen boundary condition, and should therefore be denoted by μ^ω. A fundamental question, of course, is to determine whether uniqueness holds, or whether under certain conditions there exist distinct boundary conditions ω, ω' for which $\mu^\omega \neq \mu^{\omega'}$.

Adapting the proof of the above theorem allows to obtain the following topological property of $\mathscr{G}(\pi)$ (completing the proof is left as an exercise):

Lemma 6.27. *Let π be a quasilocal specification. Then, $\mathscr{G}(\pi)$ is a closed subset of $\mathscr{M}_1(\Omega)$.*

A class of quasilocal specifications of central importance is provided by the Gibbsian specifications:

Lemma 6.28. *If Φ is absolutely summable, then π^Φ is quasilocal.*

Proof. Fix $\Lambda \Subset \mathbb{Z}^d$. Let ω be fixed, and ω' another configuration which coincides with ω on a region $\Delta \supset \Lambda$. Let $\tau_\Lambda \in \Omega_\Lambda$. We can write

$$\left|\pi_\Lambda^\Phi(\tau_\Lambda \,|\, \omega) - \pi_\Lambda^\Phi(\tau_\Lambda \,|\, \omega')\right| = \left|\int_0^1 \Big\{\frac{\mathrm{d}}{\mathrm{d}t}\frac{e^{-h_t(\tau_\Lambda)}}{z_t}\Big\}\mathrm{d}t\right|, \tag{6.35}$$

where we have set, for $0 \le t \le 1$, $h_t(\tau_\Lambda) \stackrel{\text{def}}{=} t\mathscr{H}_{\Lambda;\Phi}(\tau_\Lambda\omega_{\Lambda^c}) + (1-t)\mathscr{H}_{\Lambda;\Phi}(\tau_\Lambda\omega'_{\Lambda^c})$, and $z_t \stackrel{\text{def}}{=} \sum_{\tau_\Lambda} e^{-h_t(\tau_\Lambda)}$. As can be easily verified,

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}t}\frac{e^{-h_t(\tau_\Lambda)}}{z_t}\right| &\le 2\max_{\eta_\Lambda\in\Omega_\Lambda}\left|\mathscr{H}_{\Lambda;\Phi}(\eta_\Lambda\omega_{\Lambda^c}) - \mathscr{H}_{\Lambda;\Phi}(\eta_\Lambda\omega'_{\Lambda^c})\right| \\ &\le 4|\Lambda|\max_{i\in\Lambda}\sum_{\substack{B\Subset\mathbb{Z}^d, B\ni i\\ \mathrm{diam}(B)\ge D}}\|\Phi_B\|_\infty,\end{aligned}$$

where D is the distance between Λ and Δ^c. Due to the absolute summability of Φ, this last series goes to 0 when $D \to \infty$. As a consequence, $\pi_\Lambda^\Phi(\tau_\Lambda \,|\, \cdot)$ is continuous at ω. This implies that $\pi_\Lambda^\Phi(C \,|\, \cdot)$ is continuous for all $C \in \mathscr{C}$. □

Lemma 6.28 provides an efficient solution to the problem of constructing quasilocal specifications. Coupled with Theorem 6.26, it provides a general approach to the construction of Gibbs measures. [3]

In Chapter 3, we considered also other types of boundary conditions, namely free and periodic. It is not difficult to show, arguing similarly as in the proof of Theorem 6.26, that these also lead to Gibbs measures:

Exercise 6.14. *Use the finite-volume Gibbs distributions of the Ising model with free boundary condition, $\mu^\varnothing_{\Lambda;\beta,h}$, and the thermodynamic limit, to construct a measure $\mu^\varnothing_{\beta,h}$. Show that $\mu^\varnothing_{\beta,h} \in \mathscr{G}(\beta,h)$.*

The following exercise [4] shows that existence is not guaranteed in the absence of quasilocality.

Exercise 6.15. *Let η^- denote the configuration in which all spins are -1, and let $\eta^{-,i}$ denote the configuration in which all spins are -1, except at the vertex i, at which it is $+1$. For $\Lambda \Subset \mathbb{Z}^d$, let*

$$\pi_\Lambda(A\,|\,\omega) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{|\Lambda|}\sum_{i\in\Lambda} \mathbf{1}_A(\eta^{-,i}) & \textit{if } \omega_{\Lambda^c} = \eta^-_{\Lambda^c} \\ \mathbf{1}_A(\eta^-_\Lambda \omega_{\Lambda^c}) & \textit{otherwise.} \end{cases}$$

Show that $\pi = \{\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ is a specification and explain why it describes a system consisting of a single + spin, located anywhere on $\mathbb{Z}^d$, in a sea of − spins. Show that π is not quasilocal and that $\mathscr{G}(\pi) = \varnothing$. Hint: *Let $N^+(\omega)$ denote the number of vertices $i \in \mathbb{Z}^d$ at which $\omega_i = +1$. Assume $\mu \in \mathscr{G}(\pi)$, and show that $\mu(\{N^+ = 0\} \cup \{N^+ = 1\} \cup \{N^+ \geq 2\}) = 0$, which gives $\mu(\Omega) = 0$.*

6.5 Uniqueness

Now that we have a way of ensuring that $\mathscr{G}(\pi)$ contains at least one measure, we describe further conditions on π which ensure that this measure is actually unique. As will be seen later, the measure, when it is unique, inherits several useful properties.

Remark 6.29. We continue using Ising spins, but emphasize, however, that all statements and proofs in this section remain valid for any finite single-spin space. This matters, since, in contrast to most results in this chapter, some of the statements below are not of a qualitative nature, but involve quantitative criteria. The point is, then, that these criteria still apply verbatim to this more general setting. ⋄

6.5.1 Uniqueness vs. sensitivity to boundary conditions

The following result shows that when (and only when) there is a unique Gibbs measure, the system enjoys a very strong form of lack of sensitivity to boundary condition: any sequence of finite-volume Gibbs distributions converges.

Lemma 6.30. *The following are equivalent.*

1. *Uniqueness holds: $\mathscr{G}(\pi) = \{\mu\}$.*
2. *For all ω, all $\Lambda_n \uparrow \mathbb{Z}^d$ and all local functions f,*

$$\pi_{\Lambda_n} f(\omega) \to \mu(f)\,. \tag{6.36}$$

The convergence *for all ω* is essential here. We will see later, in Section 6.8.2, that convergence can also be guaranteed to occur in other important situations, but only for suitable sets of boundary conditions.

Proof. Fix some boundary condition ω. Remember from the proof of Theorem 6.26 that, from any sequence $(\pi_{\Lambda_n}(\cdot\,|\,\omega))_{n\geq1}$, one can extract a subsequence converging to some element of $\mathscr{G}(\pi)$. If $\mathscr{G}(\pi) = \{\mu\}$, all these subsequences must have the same limit μ. Therefore, the sequence itself converges to μ.

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

On the other hand, if $\mu, \nu \in \mathscr{G}(\pi)$, then one can write, for all local functions f,

$$\begin{aligned}|\mu(f) - \nu(f)| &= \big|\mu\pi_\Lambda(f) - \nu\pi_\Lambda(f)\big| \\ &= \Big|\int \big\{\pi_\Lambda f(\omega) - \pi_\Lambda f(\eta)\big\}\mu(\mathrm{d}\omega)\nu(\mathrm{d}\eta)\Big| \\ &\le \int \big|\pi_\Lambda f(\omega) - \pi_\Lambda f(\eta)\big|\mu(\mathrm{d}\omega)\nu(\mathrm{d}\eta)\,.\end{aligned}$$

Since $|\pi_\Lambda f(\cdot)| \le \|f\|_\infty$, we can use dominated convergence and (6.36) to conclude that $\mu(f) = \nu(f)$. Since this holds for all local functions, it follows that $\mu = \nu$. □

6.5.2 Dobrushin's Uniqueness Theorem

Our first uniqueness criterion will be formulated in terms of the one-vertex kernels $\pi_{\{i\}}(\cdot\,|\,\omega)$, which for simplicity will be denoted by π_i. Each $\pi_i(\cdot\,|\,\omega)$ should be considered as a distribution for the spin at vertex i, with boundary condition ω. We will measure the dependence of $\pi_i(\cdot\,|\,\omega)$ on the value of the boundary condition ω at other vertices. We will measure the proximity between two such distributions using the **total variation distance** (see Section B.10)

$$\|\pi_i(\cdot\,|\,\omega) - \pi_i(\cdot\,|\,\omega')\|_{TV} \stackrel{\text{def}}{=} \sum_{\eta_i=\pm1} \big|\pi_i(\eta_i\,|\,\omega) - \pi_i(\eta_i\,|\,\omega')\big|\,.$$

We can then introduce

$$c_{ij}(\pi) \stackrel{\text{def}}{=} \sup_{\substack{\omega,\omega'\in\Omega:\\ \omega_k=\omega'_k\,\forall k\neq j}} \|\pi_i(\cdot\,|\,\omega) - \pi_i(\cdot\,|\,\omega')\|_{TV}\,,$$

and

$$c(\pi) \stackrel{\text{def}}{=} \sup_{i\in\mathbb{Z}^d} \sum_{j\in\mathbb{Z}^d} c_{ij}(\pi)\,.$$

Theorem 6.31. *Let π be a quasilocal specification satisfying* ***Dobrushin's Condition of Weak Dependence:***

$$c(\pi) < 1\,. \tag{6.37}$$

Then the probability measure specified by π is unique: $|\mathscr{G}(\pi)| = 1$.

Before starting the proof, we need to introduce a few notions. Define the **oscillation** of $f : \Omega \to \mathbb{R}$ at $i \in \mathbb{Z}^d$ by

$$\delta_i(f) \stackrel{\text{def}}{=} \sup_{\substack{\omega,\eta\in\Omega\\ \omega_k=\eta_k\,\forall k\neq i}} |f(\omega) - f(\eta)|\,. \tag{6.38}$$

The oscillation enables us to quantify the variation of $f(\omega)$ when one changes ω into another configuration by successive spin flips. Namely, if $\omega_{\Lambda^c} = \eta_{\Lambda^c}$, then

$$|f(\omega) - f(\eta)| \le \sum_{i\in\Lambda} \delta_i(f)\,. \tag{6.39}$$

It is thus natural to define the **total oscillation** of f by

$$\Delta(f) \stackrel{\text{def}}{=} \sum_{i\in\mathbb{Z}^d} \delta_i(f)\,. \tag{6.40}$$

We denote the space of functions with finite total oscillation by $\mathscr{O}(\Omega)$. All local functions have finite total oscillation; by Lemma 6.21, this implies that $\mathscr{O}(\Omega)$ is dense in $C(\Omega)$. Nevertheless,

Exercise 6.16. *Show that* $\mathscr{O}(\Omega) \not\subset C(\Omega)$ *and* $\mathscr{O}(\Omega) \not\supset C(\Omega)$.

Intuitively, $\Delta(f)$ measures how far f is from being a constant. This is made clear in the following lemma. Letting $C_{\mathscr{O}}(\Omega) \stackrel{\text{def}}{=} C(\Omega) \cap \mathscr{O}(\Omega)$, we have:

Lemma 6.32. *Let* $f \in C_{\mathscr{O}}(\Omega)$. *Then* $\Delta(f) \geq \sup f - \inf f$.

Proof. Let $f \in C_{\mathscr{O}}(\Omega)$. By Exercise 6.11, f attains its supremum and its infimum. In particular there exist, for all $\epsilon > 0$, two configurations ω^1, ω^2 such that $\omega^1_{\Lambda^c} = \omega^2_{\Lambda^c}$ for some sufficiently large box Λ, and such that $\sup f \leq f(\omega^1) + \epsilon$, $\inf f \geq f(\omega^2) - \epsilon$. Then, using (6.39),

$$\sup f - \inf f \leq f(\omega^1) - f(\omega^2) + 2\epsilon \leq \sum_{i \in \Lambda} \delta_i(f) + 2\epsilon \leq \Delta(f) + 2\epsilon . \qquad \square$$

Using Lemma 6.32, we can always write

$$|\mu(f) - \nu(f)| \leq \Delta(f), \qquad \forall f \in C_{\mathscr{O}}(\Omega). \tag{6.41}$$

Proposition 6.33. *Assume* (6.37). *Let* $\mu, \nu \in \mathscr{G}(\pi)$ *be such that*

$$|\mu(f) - \nu(f)| \leq \alpha \Delta(f), \qquad \forall f \in C_{\mathscr{O}}(\Omega), \tag{6.42}$$

for some constant $\alpha \leq 1$. *Then,*

$$|\mu(f) - \nu(f)| \leq c(\pi) \alpha \Delta(f), \qquad \forall f \in C_{\mathscr{O}}(\Omega). \tag{6.43}$$

Assuming, for the moment, the validity of this proposition, we can easily conclude the proof of Theorem 6.31.

Proof of Theorem 6.31: Let $\mu, \nu \in \mathscr{G}(\pi)$ and let f be a local function. (6.41) shows that (6.42) holds with $\alpha = 1$. Since $c(\pi) < 1$, we can apply repeatedly Proposition 6.33:

$$\begin{aligned} |\mu(f) - \nu(f)| \leq \Delta(f) &\implies |\mu(f) - \nu(f)| \leq c(\pi)\Delta(f) \\ &\implies |\mu(f) - \nu(f)| \leq c(\pi)^2 \Delta(f) \\ &\implies |\mu(f) - \nu(f)| \leq c(\pi)^n \Delta(f), \quad \forall n \geq 0. \end{aligned}$$

Since $\Delta(f) < \infty$ and $c(\pi) < 1$, taking $n \to \infty$ leads to $\mu(f) = \nu(f)$. By Lemma 6.22, $\mu = \nu$. $\square$

The proof of Proposition 6.33 relies on a technical estimate:

Lemma 6.34. *Let* $f \in C_{\mathscr{O}}(\Omega)$. *Then,* $\delta_j(\pi_j f) = 0$ *for all* j *and, for any* $i \neq j$,

$$\delta_i(\pi_j f) \leq \delta_i(f) + c_{ji}(\pi) \delta_j(f). \tag{6.44}$$

The content of this lemma can be given an intuitive meaning, as follows. If f is constant, then $\delta_i(f) = 0$ for all $i \in \mathbb{Z}^d$. If f is non-constant, each oscillation $\delta_i(f)$ can be seen as a quantity of dust present at i, measuring how far f is from being constant: the less dust, the closer f is to a constant function. With this interpretation, the map $f \mapsto \pi_j f$ can be interpreted as a dusting of f at vertex j. Namely, before the dusting at j, the oscillation at any given point i is $\delta_i(f)$. After the dusting at j, Lemma 6.34 says that the amount of dust at j becomes zero ($\delta_j(\pi_j f) = 0$) and that the total amount of dust every other point $i \neq j$ is incremented, at most, by a fraction $c_{ji}(\pi)$ of the dust present at j before the dusting. For this reason, Lemma 6.34 is often called the ***dusting lemma***. ⋄

Proof of Lemma 6.34: If $i = j$, then $\delta_j(\pi_j f) = 0$ (remember that the function $\pi_j f$ is $\mathscr{F}_{\{j\}^c}$-measurable). Let us thus assume that $i \neq j$. Let ω, ω' be two configurations which agree everywhere outside i. We write

$$\begin{aligned}\pi_j f(\omega) - \pi_j f(\omega') &= \sum_{\eta_j = \pm 1} \big\{\pi_j(\eta_j \,|\, \omega) f(\eta_j \omega_{\{j\}^c}) - \pi_j(\eta_j \,|\, \omega') f(\eta_j \omega'_{\{j\}^c})\big\} \\ &= \sum_{\eta_j = \pm 1} \big\{\pi_j(\eta_j \,|\, \omega) \tilde{f}(\eta_j \omega_{\{j\}^c}) - \pi_j(\eta_j \,|\, \omega') \tilde{f}(\eta \omega'_{\{j\}^c})\big\},\end{aligned}$$

where $\tilde{f}(\cdot) \stackrel{\text{def}}{=} f(\cdot) - m$, for some constant m to be chosen later. We add and subtract $\pi_j(\eta_j \,|\, \omega) \tilde{f}(\eta_j \omega'_{\{j\}^c})$ from each term of the last sum and use

$$\begin{aligned}|\tilde{f}(\eta_j \omega_{\{j\}^c}) - \tilde{f}(\eta_j \omega'_{\{j\}^c})| &= |f(\eta_j \omega_{\{j\}^c}) - f(\eta_j \omega'_{\{j\}^c})| \leq \delta_i(f), \\ \sum_{\eta_j = \pm 1} |\pi_j(\eta_j \,|\, \omega) - \pi_j(\eta_j \,|\, \omega')| &= \|\pi_j(\cdot \,|\, \omega) - \pi_j(\cdot \,|\, \omega')\|_{TV} \leq c_{ji}(\pi).\end{aligned}$$

Since $\sum_{\eta_j} \pi_j(\eta_j \,|\, \omega) = 1$,

$$\delta_i(\pi_j f) \leq \delta_i(f) + c_{ji}(\pi) \max_{\eta_j} |\tilde{f}(\eta_j \omega'_{\{j\}^c})|.$$

Choosing $m = f((+1)_j \omega'_{\{j\}^c})$, we have $\max_{\eta_j} |\tilde{f}(\eta_j \omega'_{\{j\}^c})| \leq \delta_j(f)$ and (6.44) follows. □

Proof of Proposition 6.33: Fix an arbitrary total order on $\mathbb{Z}^d$, denoted $\succ$, in which the smallest element is the origin. We first prove that, when (6.42) holds, one has, for all $i \in \mathbb{Z}^d$,

$$|\mu(f) - \nu(f)| \leq c(\pi)\alpha \sum_{k \prec i} \delta_k(f) + \alpha \sum_{k \succeq i} \delta_k(f), \qquad \forall f \in C_{\mathscr{O}}(\Omega). \tag{6.45}$$

When $i = 0$, the first sum is empty and the claim reduces to our assumption (6.42). Let us thus assume that (6.45) has been proved for i.

Observe that, for all k, $\pi_k f \in C_{\mathscr{O}}(\Omega)$. Indeed, on the one hand, $\pi_k f$ is continuous since π is quasilocal. On the other hand, by (6.44),

$$\Delta(\pi_k f) = \sum_j \delta_j(\pi_k f) \leq \sum_j \delta_j(f) + c(\pi)\delta_k(f) < \infty.$$

Using (6.45) with f replaced by $\pi_i f$, and since $\delta_i(\pi_i f) = 0$,

$$|\mu(f) - \nu(f)| = |\mu(\pi_i f) - \nu(\pi_i f)| \leq c(\pi)\alpha \sum_{k \prec i} \delta_k(\pi_i f) + \alpha \sum_{k \succ i} \delta_k(\pi_i f).$$

Using again (6.44),

$$|\mu(f)-\nu(f)| \le c(\pi)\alpha\sum_{k<i}\delta_k(f)+\alpha\sum_{k\ge i}\delta_k(f) + \alpha\delta_i(f)c(\pi)\sum_{k<i}c_{ik}(\pi)+\alpha\delta_i(f)\sum_{k>i}c_{ik}(\pi)\,.$$

Now, observe that, since $c(\pi)<1$,

$$c(\pi)\sum_{k<i}c_{ik}(\pi)+\sum_{k>i}c_{ik}(\pi)\le\sum_{k\in\mathbb{Z}^d}c_{ik}(\pi)=c(\pi)\,,$$

which yields

$$\begin{aligned}|\mu(f)-\nu(f)| &\le c(\pi)\alpha\sum_{k<i}\delta_k(f)+\alpha\sum_{k>i}a_k\delta_k(f)+c(\pi)\alpha\delta_i(f)\\ &= c(\pi)\alpha\sum_{k\le i}\delta_k(f)+\alpha\sum_{k>i}\delta_k(f)\,.\end{aligned}$$

This shows that (6.45) holds for all $i\in\mathbb{Z}^d$. Since $\sum_k\delta_k(f)=\Delta(f)<\infty$, (6.43) now follows by letting i increase to infinity (with respect to $\succ$) in (6.42). □

6.5.3 Application to Gibbsian specifications

Theorem 6.31 is very general. We will now apply it to several Gibbsian specifications. We will start with regimes in which the Gibbs measure is unique despite possibly strong interactions between the spins.

Exercise 6.17. *Consider the Ising model ($d\ge 1$) with a magnetic field $h>0$ and arbitrary inverse temperature β. Use Theorem 6.35 to show that $|\mathscr{G}(\beta,h)|=1$ for all large enough h. (Contrast this result with the corresponding one obtained in Theorem 3.25, where it was shown that uniqueness holds for all $h\neq 0$.)*

Exercise 6.18. *Consider the Blume–Capel model in $d\ge 1$ (see (6.29)).*

1. *Consider first $(\lambda,h)=(0,0)$, and give a range of values of β for which uniqueness holds.*
2. *Then, fix $(\lambda,h)=t\mathbf{e}$, with $t>0$. Show that if $\mathbf{e}\in\mathbb{S}^1$ points in any direction different from $(1,0)$, $(-1,1)$ or $(-1,-1)$, then for all $\beta>0$, the Gibbs measure is unique as soon as t is sufficiently large. (See Figure 6.2.)*

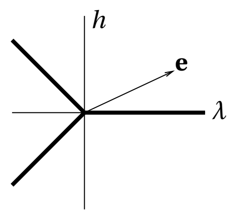


Figure 6.2: The Blume–Capel with parameters $(\lambda,h)=t\mathbf{e}$ has a unique Gibbs measure when $t>0$ is large enough, and when $\mathbf{e}$ points to any direction distinct from those indicated by the bold line.

Exercise 6.19. *The (nearest-neighbor)* ***Potts antiferromagnet*** *on* $\mathbb{Z}^d$ *at inverse temperature* $\beta \geq 0$ *has single spin space* $\Omega_0 \stackrel{\text{def}}{=} \{0, \dots, q-1\}$ *and is associated to the potential*

$$\Phi_B(\omega) = \begin{cases} +\beta\delta_{\omega_i,\omega_j} & \text{if } B = \{i,j\}, i \sim j, \\ 0 & \text{otherwise.} \end{cases}$$

Show that this model has a unique Gibbs measure for all $\beta \in \mathbb{R}_{\geq 0}$*, provided that* $q > 6d$.

Let us now formulate the criterion of Theorem 6.31 in a form better suited to the treatment of weak interactions. Let $\delta(f) \stackrel{\text{def}}{=} \sup_{\eta',\eta''} |f(\eta') - f(\eta'')|$.

Theorem 6.35. *Assume that* $\Phi = \{\Phi_B\}_{B \Subset \mathbb{Z}^d}$ *is absolutely summable and satisfies*

$$\sup_{i \in \mathbb{Z}^d} \sum_{j \neq i} \sum_{B \supset \{i,j\}} \delta(\Phi_B) < 1. \tag{6.46}$$

Then Dobrushin's condition of weak dependence is satisfied, and therefore there is a unique Gibbs measure specified by π^Φ.

Proof. Fix some $i \in \mathbb{Z}^d$, and let ω and ω' coincide everywhere except at some vertex $j \neq i$. Starting as in the proof of Lemma 6.28,

$$\|\pi_i^\Phi(\cdot\,|\,\omega) - \pi_i^\Phi(\cdot\,|\,\omega')\|_{TV} \leq \int_0^1 \Big\{ \sum_{\eta_i} \Big| \frac{\mathrm{d}\nu_t(\eta_i)}{\mathrm{d}t} \Big| \Big\} \mathrm{d}t$$

where, for $0 \leq t \leq 1$, $\nu_t(\eta_i) \stackrel{\text{def}}{=} \frac{e^{-h_t(\eta_i)}}{z_t}$, with

$$h_t(\eta_i) \stackrel{\text{def}}{=} t\mathscr{H}_{\{i\};\Phi}(\eta_i\omega_{\{i\}^c}) + (1-t)\mathscr{H}_{\{i\};\Phi}(\eta_i\omega'_{\{i\}^c}),$$

and $z_t \stackrel{\text{def}}{=} \sum_{\eta_i} e^{-h_t(\eta_i)}$. A straightforward computation shows that

$$\frac{\mathrm{d}\nu_t(\eta_i)}{\mathrm{d}t} = \big\{\Delta\mathscr{H}_i - E_{\nu_t}[\Delta\mathscr{H}_i]\big\}\nu_t(\eta_i),$$

where $\Delta\mathscr{H}_i(\eta_i) \stackrel{\text{def}}{=} \mathscr{H}_{\{i\};\Phi}(\eta_i\omega'_{\{i\}^c}) - \mathscr{H}_{\{i\};\Phi}(\eta_i\omega_{\{i\}^c})$. We therefore have

$$\begin{aligned} \sum_{\eta_i} \Big| \frac{\mathrm{d}\nu_t(\eta_i)}{\mathrm{d}t} \Big| &= E_{\nu_t}\Big[\big|\Delta\mathscr{H}_i - E_{\nu_t}[\Delta\mathscr{H}_i]\big|\Big] \\ &\leq E_{\nu_t}\Big[\big(\Delta\mathscr{H}_i - E_{\nu_t}[\Delta\mathscr{H}_i]\big)^2\Big]^{1/2} \\ &\leq E_{\nu_t}\big[\big(\Delta\mathscr{H}_i - m\big)^2\big]^{1/2}, \end{aligned}$$

where the first inequality follows from the Cauchy-Schwartz Inequality, and we introduced an arbitrary number $m \in \mathbb{R}$ (remember that $m \mapsto E[(X-m)^2]$ is minimal when $m = E[X]$). Choosing $m = (\max \Delta\mathscr{H}_i + \min \Delta\mathscr{H}_i)/2$, we have

$$|\Delta\mathscr{H}_i - m| \leq \tfrac{1}{2} \max_{\eta_i,\eta'_i} \big|\Delta\mathscr{H}_i(\eta_i) - \Delta\mathscr{H}_i(\eta'_i)\big|.$$

Each $\Delta\mathscr{H}_i(\cdot)$ contains a sum over sets $B \ni i$, and notice that those sets B which do not contain j do not contribute. We can thus restrict to those sets $B \supset \{i, j\}$ and get

$$\begin{aligned}
&|\Delta\mathscr{H}_i(\eta_i) - \Delta\mathscr{H}_i(\eta'_i)| \\
&\qquad \le \sum_{B\supset\{i,j\}} \left\{|\Phi_B(\eta_i\omega'_{\{i\}^c}) - \Phi_B(\eta'_i\omega'_{\{i\}^c})| + |\Phi_B(\eta_i\omega_{\{i\}^c}) - \Phi_B(\eta'_i\omega_{\{i\}^c})|\right\} \\
&\qquad \le 2\sum_{B\supset\{i,j\}} \delta(\Phi_B)\,.
\end{aligned}$$

This proves (6.46). □

Let us give a simple example of application of the above criterion.

Example 6.36. Consider first the nearest-neighbor Ising model with $h = 0$ on $\mathbb{Z}^d$, whose potential was given in (6.26). The only sets B that contribute to the sum (6.46) are the nearest neighbors $B = \{i, j\}$, $i \sim j$, for which $\delta(\Phi_B) = 2\beta$. (6.46) therefore reads, since each i has $2d$ neighbors,

$$2\beta \cdot 2d < 1\,.$$

In $d = 1$, this means that uniqueness holds when $\beta < \frac{1}{4}$, although we know from the results of Chapter 3 that uniqueness holds at *all* temperatures. In $d = 2$, the above guarantees uniqueness when $\beta < \frac{1}{8} = 0.125$, which should be compared with the exact range, known to be $\beta \le \beta_c(2) = 0.4406\ldots$. ⋄

We will actually see in Corollary 6.41 that, for finite-range models, uniqueness holds at all temperatures when $d = 1$.

More generally, the above criterion allows one to prove uniqueness at sufficiently high temperature for a wide class of models. A slight rewriting of the condition makes the application more immediate. If one changes the order of summation in the double sum in (6.46), the latter becomes

$$\sum_{j\neq i}\sum_{B\supset\{i,j\}} \delta(\Phi_B) = \sum_{B\ni i}(|B|-1)\delta(\Phi_B)\,. \tag{6.47}$$

We can thus state a general, easily applicable high-temperature uniqueness result. Remember that the inverse temperature β can always be associated to a potential $\Phi \stackrel{\text{def}}{=} \{\Phi_B\}_{B\Subset\mathbb{Z}^d}$, by multiplication: $\beta\Phi \stackrel{\text{def}}{=} \{\beta\Phi_B\}_{B\Subset\mathbb{Z}^d}$.

Corollary 6.37. *Let $\Phi = \{\Phi_B\}_{B\Subset\mathbb{Z}^d}$ be an absolutely summable potential satisfying*

$$b \stackrel{\text{def}}{=} \sup_{i\in\mathbb{Z}^d}\sum_{B\ni i}(|B|-1)\|\Phi_B\|_\infty < \infty\,, \tag{6.48}$$

and let $\beta_0 \stackrel{\text{def}}{=} \frac{1}{2b}$. Then, for all $\beta < \beta_0$, there is a unique measure compatible with $\pi^{\beta\Phi}$.

Proof. It suffices to use (6.47) in Theorem 6.35, with $\delta(\Phi_B) \le 2\|\Phi_B\|_\infty$. □

Exercise 6.20. *Consider the long-range Ising model introduced in* (6.27)*, with $h = 0$ and $J_{ij} = \|j-i\|_\infty^{-\alpha}$. Find a range of values of $\alpha > 0$ (depending on the dimension) for which* (6.48) *holds, and deduce a range of values of* $0 < \beta < \infty$ *for which uniqueness holds.*

In dimension 1, the previous exercise guarantees that uniqueness holds at sufficiently high temperature whenever $\alpha > 1$. We will prove later that, when $\alpha > 2$, uniqueness actually holds for all positive temperatures; see Example 6.42.

6.5.4 Uniqueness at high temperature via cluster expansion

In this section, we consider an alternative approach, relying on the cluster expansion, to establish uniqueness of the Gibbs measure at sufficiently high temperature. We have seen in Lemma 6.30 that $\mathscr{G}(\beta\Phi) = \{\mu\}$ if and only if

$$\pi_{\Lambda_n}^{\beta\Phi} f(\omega) \to \mu(f), \qquad \forall \omega \in \Omega, \tag{6.49}$$

for every local function f. Here we provide a direct way of proving such a convergence.

Theorem 6.38. *Assume that $\Phi = \{\Phi_B\}_{B \Subset \mathbb{Z}^d}$ satisfies*

$$\sup_{i \in \mathbb{Z}^d} \sum_{B \ni i} \|\Phi_B\|_\infty e^{4|B|} < \infty. \tag{6.50}$$

Then, there exists $0 < \beta_1 < \infty$ such that, for all $\beta \le \beta_1$, (6.49) holds. As a consequence: $\mathscr{G}(\beta\Phi) = \{\mu\}$. Moreover, when Φ has finite range, the convergence in (6.49) is exponential: for all sufficiently large Λ,

$$\left|\pi_{\Lambda}^{\beta\Phi} f(\omega) - \mu(f)\right| \le D\|f\|_\infty e^{-Cd(\mathrm{supp}(f), \Lambda^c)}, \qquad \forall \omega \in \Omega, \tag{6.51}$$

where $C > 0$ and D depend on Φ.

Since $\mu(f) = \int \pi_{\Lambda}^{\beta\Phi}(f\,|\,\omega)\mu(\mathrm{d}\omega)$ for any $\mu \in \mathscr{G}(\beta\Phi)$, Theorem 6.38 is a consequence of the following proposition, whose proof relies on the cluster expansion and provides an explicit expression for $\mu(f)$. In order not to delve here into the technicalities of the cluster expansion, we postpone this proof to the end of Section 6.12.

Proposition 6.39. *If (6.50) holds, then there exists $0 < \beta_1 < \infty$ such that, for all $\beta \le \beta_1$, the following holds. Fix some ω. For every local function f, there exists $c(f)$ (independent of ω) such that*

$$\lim_{\Lambda \uparrow \mathbb{Z}^d} \pi_{\Lambda}^{\beta\Phi} f(\omega) = c(f). \tag{6.52}$$

Moreover, if Φ has finite range,

$$\left|\pi_{\Lambda}^{\beta\Phi} f(\omega) - c(f)\right| \le D\|f\|_\infty e^{-Cd(\mathrm{supp}(f), \Lambda^c)}, \tag{6.53}$$

where C and D depend on Φ.

6.5.5 Uniqueness in one dimension

In one dimension, the criterion (6.46) implies uniqueness for any model with absolutely convergent potential, but only at sufficiently high temperatures (small values of β).

We know that the nearest-neighbor Ising model on $\mathbb{Z}$ has a unique Gibbs measure at *all* temperatures, so the proof given above, relying on Dobrushin's Condition of Weak Dependence, ignores some important features of one-dimensional systems. We now establish another criterion, of less general applicability, but providing considerably stronger results when $d = 1$.

Theorem 6.40. *Let Φ be an absolutely summable potential such that*

$$D \stackrel{\text{def}}{=} \sup_n \sup_{\substack{\omega_{\mathsf{B}(n)} \\ \eta_{\mathsf{B}(n)^c}, \eta'_{\mathsf{B}(n)^c}}} \left| \mathscr{H}_{\mathsf{B}(n);\Phi}(\omega_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c}) - \mathscr{H}_{\mathsf{B}(n);\Phi}(\omega_{\mathsf{B}(n)}\eta'_{\mathsf{B}(n)^c}) \right| < \infty . \tag{6.54}$$

Then there is a unique Gibbs measure compatible with π^Φ.

Since

$$\left| \mathscr{H}_{\mathsf{B}(n);\Phi}(\omega_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c}) - \mathscr{H}_{\mathsf{B}(n);\Phi}(\omega_{\mathsf{B}(n)}\eta'_{\mathsf{B}(n)^c}) \right| \le 2 \sum_{\substack{A \cap \mathsf{B}(n) \neq \varnothing \\ A \cap \mathsf{B}(n)^c \neq \varnothing}} \|\Phi_A\|_\infty , \tag{6.55}$$

condition (6.54) will be satisfied for one-dimensional systems in which the interaction between the inside and the outside of any interval $\mathsf{B}(n) = \{-n, \dots, n\}$ is uniformly bounded in n, meaning that *boundary effects are negligible*.

The sum in the right-hand side of (6.55) is of course finite when Φ has finite range, which allows to state a general uniqueness result for one-dimensional systems:

Corollary 6.41. *($d = 1$) If Φ is any finite-range potential, then $|\mathscr{G}(\Phi)| = 1$.*

However, the sum in the right-hand side of (6.55) can contain infinitely many terms, as long as these decay sufficiently fast, as the next example shows.

Example 6.42. Consider the one-dimensional long-range Ising model (6.27), with

$$J_{ij} = |j - i|^{-(2+\epsilon)} ,$$

with $\epsilon > 0$. Using (6.55),

$$\begin{aligned} \left| \mathscr{H}_{\mathsf{B}(n);\beta\Phi}(\omega_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c}) - \mathscr{H}_{\mathsf{B}(n);\beta\Phi}(\omega_{\mathsf{B}(n)}\eta'_{\mathsf{B}(n)^c}) \right| &\le 2 \sum_{i \in \mathsf{B}(n)} \sum_{j \in \mathsf{B}(n)^c} \frac{\beta}{|j-i|^{2+\epsilon}} \\ &\le 2\beta \sum_{k \ge 1} \sum_{\substack{i \in \mathsf{B}(n): \\ d(i, \mathsf{B}(n)^c) = k}} \sum_{r \ge k} \frac{1}{r^{2+\epsilon}} \\ &\le 2\beta c_\epsilon \sum_{k \ge 1} \frac{1}{k^{1+\epsilon}} < \infty \end{aligned}$$

Theorem 6.40 implies uniqueness for all finite values of $\beta \ge 0$ whenever $\epsilon > 0$. Remarkably, this is a sharp result, as it can be shown that uniqueness fails at large values of β whenever $\epsilon \le 0$ [5]. ⋄

Since we do not yet have all the necessary tools, we postpone the proof of Theorem 6.40 to the end of Section 6.8.4 (p. 296). It will rely on the following ingredient:

Lemma 6.43. *Let D be defined as in* (6.54)*. Then, for all $\omega, \eta \in \Omega$ and all cylinders $C \in \mathscr{C}$, for all large enough n,*

$$e^{-2D} \pi^\Phi_{\mathsf{B}(n)}(C \,|\, \eta) \le \pi^\Phi_{\mathsf{B}(n)}(C \,|\, \omega) \le e^{2D} \pi^\Phi_{\mathsf{B}(n)}(C \,|\, \eta) . \tag{6.56}$$

Proof. Using (6.54) in the Boltzmann weight, we obtain

$$e^{-D} e^{-\mathscr{H}_{\mathsf{B}(n);\Phi}(\tau_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c})} \le e^{-\mathscr{H}_{\mathsf{B}(n);\Phi}(\tau_{\mathsf{B}(n)}\omega_{\mathsf{B}(n)^c})} \le e^{D} e^{-\mathscr{H}_{\mathsf{B}(n);\Phi}(\tau_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c})} .$$

This yields $\mathbf{Z}^{\omega}_{\mathsf{B}(n);\Phi} \le \mathbf{Z}^{\eta}_{\mathsf{B}(n);\Phi} e^{D}$. Thus, $\pi^{\Phi}_{\mathsf{B}(n)}(\tau_{\mathsf{B}(n)}\,|\,\omega) \ge e^{-2D}\pi^{\Phi}_{\mathsf{B}(n)}(\tau_{\mathsf{B}(n)}\,|\,\eta)$. Let $C \in \mathscr{C}$. If n is large enough for $\mathsf{B}(n)$ to contain the base of C, then $\mathbf{1}_C(\tau_{\mathsf{B}(n)}\omega_{\mathsf{B}(n)^c}) = \mathbf{1}_C(\tau_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c})$. Therefore,

$$\begin{aligned}\pi^{\Phi}_{\mathsf{B}(n)}(C\,|\,\omega) &= \sum_{\tau_{\mathsf{B}(n)}} \pi^{\Phi}_{\mathsf{B}(n)}(\tau_{\mathsf{B}(n)}\,|\,\omega)\mathbf{1}_C(\tau_{\mathsf{B}(n)}\omega_{\mathsf{B}(n)^c})\\ &\ge e^{-2D}\sum_{\tau_{\mathsf{B}(n)}} \pi^{\Phi}_{\mathsf{B}(n)}(\tau_{\mathsf{B}(n)}\,|\,\eta)\mathbf{1}_C(\tau_{\mathsf{B}(n)}\eta_{\mathsf{B}(n)^c}) = e^{-2D}\pi^{\Phi}_{\mathsf{B}(n)}(C\,|\,\eta). \qquad \square\end{aligned}$$

6.6 Symmetries

In this section, we study how the presence of symmetries in a specification π can extend to the measures in $\mathscr{G}(\pi)$. We do not assume that the single-spin space Ω_0 is necessarily $\{\pm 1\}$.

We will be interested in the action of a group $(\mathsf{G},\cdot)$ on the set of configurations Ω. That is, we consider a family $(\tau_{\mathsf{g}})_{\mathsf{g}\in\mathsf{G}}$ of maps $\tau_{\mathsf{g}}:\Omega\to\Omega$ such that

1. $(\tau_{\mathsf{g}_1}\circ\tau_{\mathsf{g}_2})\omega = \tau_{\mathsf{g}_1\cdot\mathsf{g}_2}\omega$ for all $\mathsf{g}_1,\mathsf{g}_2\in\mathsf{G}$, and
2. $\tau_{\mathsf{e}}\omega = \omega$ for all $\omega\in\Omega$, where e is the neutral element of G.

Note that $\tau_{\mathsf{g}}^{-1} = \tau_{\mathsf{g}^{-1}}$ for all $\mathsf{g}\in\mathsf{G}$. The action of the group can be extended to functions and measures. For all $\mathsf{g}\in\mathsf{G}$, all functions $f:\Omega\to\mathbb{R}$ and all $\mu\in\mathscr{M}_1(\Omega)$, we define

$$\tau_{\mathsf{g}}f(\omega) \stackrel{\text{def}}{=} f(\tau_{\mathsf{g}}^{-1}\omega), \qquad \tau_{\mathsf{g}}\mu(A) \stackrel{\text{def}}{=} \mu(\tau_{\mathsf{g}}^{-1}A),$$

for all $\omega\in\Omega$ and all $A\in\mathscr{F}$. Of course, we then have $\tau_{\mathsf{g}}\mu(f) = \mu(\tau_{\mathsf{g}}^{-1}f)$, for all integrable functions f.

We will use τ_{g} to act on a specification $\pi = \{\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$, and turn it into a new specification $\tau_{\mathsf{g}}\pi = \{\tau_{\mathsf{g}}\pi_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$. We will mainly consider two types of transformations, *internal* and *spatial*.

1. An **internal transformation** starts with a group G acting on the single-spin space Ω_0. The action of G is then extended to Ω by setting, for all $\mathsf{g}\in\mathsf{G}$ and all $\omega\in\Omega$,
$$(\tau_{\mathsf{g}}\omega)_i \stackrel{\text{def}}{=} \tau_{\mathsf{g}}\omega_i \quad \forall i\in\mathbb{Z}^d.$$
(We use the same notation for the action on both Ω_0 and Ω as this will never lead to ambiguity.) The action of τ_{g} on a kernel π_Λ is defined by
$$(\tau_{\mathsf{g}}\pi)_\Lambda(A\,|\,\omega) \stackrel{\text{def}}{=} \pi_\Lambda(\tau_{\mathsf{g}}^{-1}A\,|\,\tau_{\mathsf{g}}^{-1}\omega). \tag{6.57}$$
2. In the case of a **spatial transformation**, we start with a group G acting on $\mathbb{Z}^d$ and we extend its action to Ω by setting, for all $\mathsf{g}\in\mathsf{G}$ and all $\omega\in\Omega$,
$$(\tau_{\mathsf{g}}\omega)_i \stackrel{\text{def}}{=} \omega_{\tau_{\mathsf{g}}^{-1}i} \quad \forall i\in\mathbb{Z}^d.$$
Basic examples of spatial transformations are: translations, rotations and reflections. The action of τ_{g} on π_Λ is defined by
$$(\tau_{\mathsf{g}}\pi)_\Lambda(A\,|\,\omega) \stackrel{\text{def}}{=} \pi_{\tau_{\mathsf{g}}^{-1}\Lambda}(\tau_{\mathsf{g}}^{-1}A\,|\,\tau_{\mathsf{g}}^{-1}\omega). \tag{6.58}$$

A **general transformation** is then a composition of these two types of transformations. To simplify the exposition, for the rest of this section, we focus on internal transformations. However, everything can be extended in a straightforward way to the other cases. Invariance under translations will play an important role in the rest of the book. For that reason, we will describe this type of spatial transformation in more detail in Section 6.7, together with translation-invariant specifications and Gibbs measures.

So, until the end of this section, we assume that the actions τ_{g} are associated to an internal transformation group.

Definition 6.44. *π is* **G-*invariant*** *if $(\tau_{\mathsf{g}}\pi)_\Lambda = \pi_\Lambda$ for all $\Lambda \Subset \mathbb{Z}^d$ and all $\mathsf{g} \in \mathsf{G}$.*

The most important example is that of a Gibbsian specification associated to a potential that is invariant under the action of G. Namely, consider an absolutely summable potential $\Phi = \{\Phi_A\}_{A\Subset\mathbb{Z}^d}$, and let us assume that $\tau_{\mathsf{g}}\Phi_A = \Phi_A$ for all $A \Subset \mathbb{Z}^d$ and all $\mathsf{g} \in \mathsf{G}$. It then follows that, for all $\Lambda \Subset \mathbb{Z}^d$ and all $\mathsf{g} \in \mathsf{G}$,

$$\mathscr{H}_\Lambda(\omega) = \mathscr{H}_\Lambda(\tau_{\mathsf{g}}\omega) \quad \forall \omega \in \Omega\,.$$

As a consequence, the associated specification is G-invariant: $\tau_{\mathsf{g}}\pi^\Phi = \pi^\Phi$ for all $\mathsf{g} \in \mathsf{G}$. Let us mention a few specific examples.

- **The Ising model with $h = 0$.** In this case, the internal symmetry group is given by the cyclic group Z_2, that is, the group with two elements: the neutral element e and the **spin flip** f which acts on Ω_0 via $\tau_{\mathsf{f}}\omega_0 = -\omega_0$. As already discussed in Chapter 3, the Hamiltonian is invariant under the global spin flip,
$$\mathscr{H}_{\Lambda;\beta,0}(\omega) = \mathscr{H}_{\Lambda;\beta,0}(\tau_{\mathsf{f}}\omega)\,.$$
and the specification of the Ising model with $h = 0$ is therefore invariant under the action of Z_2. When $h \neq 0$, this is of course no longer true.

- **The Potts model.** In this case, the internal symmetry group is S_q, the group of all permutations on the set $\Omega_0 = \{0,\dots,q-1\}$. It is immediate to check that the potential defining the Potts model (see (6.28)) is S_q-invariant.

- **The Blume–Capel model with $h = 0$.** As in the Ising model, the potential is invariant under the action of Z_2, the spin flip acting again on $\Omega_0 = \{-1,0,1\}$ via $\tau_{\mathsf{f}}\omega_0 = -\omega_0$. That is, the model is invariant under the interchange of $+$ and $-$ spins (leaving the 0 spins unchanged).

At the end of the chapter, we will also consider models in which the spin-space is not a finite set.

6.6.1 Measures compatible with a G-invariant specification

Theorem 6.45. *Let G be an internal transformation group and π be a G-invariant specification. Then, $\mathscr{G}(\pi)$ is preserved by G:*

$$\mu \in \mathscr{G}(\pi) \Rightarrow \tau_{\mathsf{g}}\mu \in \mathscr{G}(\pi) \quad \forall \mathsf{g} \in \mathsf{G}\,.$$

Proof. Let $\mathrm{g} \in \mathsf{G}$, $\Lambda \Subset \mathbb{Z}^d$, $\omega \in \Omega$ and $A \in \mathscr{F}$. Since $(\tau_{\mathrm{g}}\mu)(f) = \mu(f \circ \tau_{\mathrm{g}})$,

$$(\tau_{\mathrm{g}}\mu)\pi_\Lambda(A) = \int \pi_\Lambda(A \,|\, \tau_{\mathrm{g}}\omega)\,\mu(\mathrm{d}\omega)$$
$$= \int \pi_\Lambda(\tau_{\mathrm{g}}^{-1}A \,|\, \omega)\,\mu(\mathrm{d}\omega) = \mu\pi_\Lambda(\tau_{\mathrm{g}}^{-1}A) = \mu(\tau_{\mathrm{g}}^{-1}A) = \tau_{\mathrm{g}}\mu(A)\,.$$

It follows that $\tau_{\mathrm{g}}\mu \in \mathscr{G}(\pi)$. □

The above result does not necessarily mean that $\tau_{\mathrm{g}}\mu = \mu$ for all $\mathrm{g} \in \mathsf{G}$, but this property is of course verified when uniqueness holds: in this case, the unique Gibbs measure inherits all the symmetries of the Hamiltonian.

Corollary 6.46. *Assume that $\mathscr{G}(\pi) = \{\mu\}$. If π is G-invariant, then μ is G-invariant: $\tau_{\mathrm{g}}\mu = \mu$ for all $\mathrm{g} \in \mathsf{G}$.*

However, when there are multiple measures compatible with a given specification, it can happen that some of these measures are not G-invariant.

Definition 6.47. *Let π be G-invariant. If there exists $\mu \in \mathscr{G}(\pi)$ for which $\tau_{\mathrm{g}}\mu \neq \mu$, the associated symmetry is said to be **spontaneously broken** under μ.*

*"Spontaneous" is used here to distinguish this phenomenon from an **explicit symmetry breaking**. The latter occurs, for example, when one introduces a nonzero magnetic field h in the Ising model, thereby deliberately destroying the symmetry present when $h = 0$.* ⋄

Example 6.48. We have seen that, when $h = 0$, the interactions of the Ising model treat + and − spins in a completely symmetric way: $\mathscr{H}_{\Lambda;\beta,0}(\tau_{\mathrm{f}}\omega) = \mathscr{H}_{\Lambda;\beta,0}(\omega)$, where f denotes the global spin flip. Nevertheless, when $d \geq 2$ and $\beta > \beta_{\mathrm{c}}(d)$, we know that the associated Gibbs measures $\mu^+_{\beta,0} \neq \mu^-_{\beta,0}$ are not invariant under a global spin flip, since $\langle\sigma_0\rangle^+_{\beta,0} > 0 > \langle\sigma_0\rangle^-_{\beta,0}$: the symmetry is spontaneously broken. We nevertheless have that $\tau_{\mathrm{f}}\mu^+_{\beta;0} = \mu^-_{\beta;0}$, in complete accordance with the claim of Theorem 6.45. ⋄

6.7 Translation invariant Gibbs measures

The theory of Gibbs measures often becomes simpler once restricted to translation-invariant measures. We will see for instance in Section 6.9 that, in this framework, Gibbs measures can be characterized in an alternative way, allowing us to establish a close relation between the DLR formalism and thermostatics.

Translations on $\mathbb{Z}^d$ are a particular type of spatial transformation group, as described in the previous section. Remember from Chapter 3 (see (3.15)) that the **translation by $j \in \mathbb{Z}^d$**, denoted $\theta_j : \mathbb{Z}^d \to \mathbb{Z}^d$, is defined by

$$\theta_j i \stackrel{\text{def}}{=} i + j\,,$$

and can be seen as an action of $\mathbb{Z}^d$ on itself. Notice that $\theta_j^{-1} = \theta_{-j}$.

Definition 6.49. *$\mu \in \mathscr{M}_1(\Omega)$ is **translation invariant** if $\theta_j\mu = \mu$ for all $j \in \mathbb{Z}^d$.*

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

Example 6.50. The product measure $\rho^{\mathbb{Z}^d}$ obtained with $\rho_i \equiv \rho_0$ (some fixed distribution on Ω_0) is translation invariant. ⋄

Example 6.51. The Gibbs measures of the Ising model, $\mu^+_{\beta,h}$ and $\mu^-_{\beta,h}$, are translation invariant. Namely, we saw in Theorem 3.17 that $\langle\cdot\rangle^+_{\beta,h}$ is invariant under any translation θ_j. Therefore, for each cylinder $C \in \mathscr{C}$, since $\mathbf{1}_C$ is local,

$$\theta_j \mu^+_{\beta,h}(C) = \mu^+_{\beta,h}(\theta_j^{-1}C) = \langle \mathbf{1}_C \circ \theta_j\rangle^+_{\beta,h} = \langle \mathbf{1}_C\rangle^+_{\beta,h} = \mu^+_{\beta,h}(C)\,.$$

This implies that $\theta_j\mu^+_{\beta,h}$ and $\mu^+_{\beta,h}$ coincide on cylinders. Since the cylinders generate $\mathscr{F}$, Corollary B.37 implies $\theta_j\mu^+_{\beta,h} = \mu^+_{\beta,h}$. The same can be done with $\mu^-_{\beta,h}$. ⋄

We will sometimes use the following notation:

$$\mathscr{M}_{1,\theta}(\Omega) \stackrel{\text{def}}{=} \big\{\mu \in \mathscr{M}_1(\Omega) : \mu \text{ is translation invariant}\big\}\,.$$

The study of translation-invariant measures is simplified thanks to the fact that *spatial averages of local observables*,

$$\lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|} \sum_{j\in\mathsf{B}(n)} \theta_j f\,,$$

exist almost surely and can be related to their expectation. To formulate this precisely, let $\mathscr{I}$ denote the **$\boldsymbol{\sigma}$-algebra of translation-invariant events**:

$$\mathscr{I} \stackrel{\text{def}}{=} \big\{A \in \mathscr{F} : \theta_j A = A,\ \forall j \in \mathbb{Z}\big\}\,.$$

The following result is called the *multidimensional ergodic theorem*. We state it without proof. [6]

Theorem 6.52. *Let $\mu \in \mathscr{M}_{1,\theta}(\Omega)$. Then, for any $f \in L^1(\mu)$,*

$$\frac{1}{|\mathsf{B}(n)|} \sum_{j\in\mathsf{B}(n)} \theta_j f \to \mu(f\,|\,\mathscr{I}) \quad \mu\text{-a.s. and in } L^1(\mu)\,. \tag{6.59}$$

Note that the limit (6.59) remains random in general. However, it becomes deterministic if one assumes that μ satisfies one further property.

Definition 6.53. *$\mu \in \mathscr{M}_{1,\theta}(\Omega)$ is **ergodic** if each translation-invariant event A has probability $\mu(A) = 0$ or 1.*

Theorem 6.54. *If $\mu \in \mathscr{M}_{1,\theta}(\Omega)$ is ergodic, then, for any $f \in L^1(\mu)$,*

$$\frac{1}{|\mathsf{B}(n)|} \sum_{j\in\mathsf{B}(n)} \theta_j f \to \mu(f) \quad \mu\text{-a.s. and in } L^1(\mu)\,.$$

Proof. By Theorem 6.52, we only need to show that $\mu(f\,|\,\mathscr{I}) = \mu(f)$ almost surely. Notice that $g \stackrel{\text{def}}{=} \mu(f\,|\,\mathscr{I})$ is $\mathscr{I}$-measurable, and therefore $\{g \le \alpha\} \in \mathscr{I}$ for all $\alpha \in \mathbb{R}$, giving $\mu(g \le \alpha) \in \{0,1\}$. Since $\alpha \mapsto \mu(g \le \alpha)$ is non-decreasing, there exists some $\alpha_* \in \mathbb{R}$ for which $\mu(g = \alpha_*) = 1$. But since $\mu(g) = \mu(f)$, we have $\alpha_* = \mu(f)$. □

6.7.1 Translation invariant specifications

The action of a translation θ_j on a kernel π_Λ takes the form (see (6.58))

$$(\theta_j\pi)_\Lambda(A\,|\,\omega) \stackrel{\text{def}}{=} \pi_{\theta_j^{-1}\Lambda}(\theta_j^{-1}A\,|\,\theta_j^{-1}\omega)\,. \tag{6.60}$$

Say that $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ is **translation invariant** if $\theta_j\pi_\Lambda = \pi_\Lambda$ for all Λ and all $j \in \mathbb{Z}^d$.

Theorem 6.45 and its corollary also hold in this situation:

Exercise 6.21. *Show that if π is translation invariant and $\mu \in \mathscr{G}(\pi)$, then $\theta_j\mu \in \mathscr{G}(\pi)$ for all $j \in \mathbb{Z}^d$. In particular, if $\mathscr{G}(\pi) = \{\mu\}$, then μ is translation invariant.*

For example, if $\Phi = \{\Phi_B\}_{B\Subset\mathbb{Z}^d}$ is a translation-invariant (for all $B \Subset \mathbb{Z}^d$, $\Phi_{\theta_j B}(\omega) = \Phi_B(\theta_{-j}\omega)$) absolutely summable potential, then π^Φ is translation invariant, and $\theta_j\mu \in \mathscr{G}(\Phi)$ for each $\mu \in \mathscr{G}(\Phi)$.

We will sometimes use the following notation:

$$\mathscr{G}_\theta(\pi) \stackrel{\text{def}}{=} \big\{\mu \in \mathscr{G}(\pi) : \mu \text{ translation invariant}\big\}\,.$$

We leave it as an exercise to check that translation-invariant measures compatible with a translation-invariant quasilocal specification always exist:

Exercise 6.22. *Show that if π is translation invariant and quasilocal, then $\mathscr{G}_\theta(\pi) \neq \varnothing$.* Hint: *Take $\mu \in \mathscr{G}(\pi)$, and use $\mu_n \stackrel{\text{def}}{=} \frac{1}{|\mathrm{B}(n)|}\sum_{j\in\mathrm{B}(n)}\theta_j\mu$.*

Let us stress, as we did in the previous section in the case of internal transformations, that translation-invariant specifications do not necessarily yield translation-invariant measures:

Example 6.55. The specification associated to the Ising antiferromagnet defined in (3.76) is clearly translation invariant. Nevertheless, neither of the Gibbs measures μ_β^{even} and μ_β^{odd} constructed in Exercise 3.33 is translation invariant. ⋄

6.8 Convexity and Extremal Gibbs measures

We now investigate general properties of $\mathscr{G}(\pi)$, without assuming either symmetry or uniqueness, and derive fundamental properties of the measures $\mu \in \mathscr{G}(\pi)$.

Let $\nu_1, \nu_2 \in \mathscr{M}_1(\Omega)$, and $\lambda \in [0,1]$. Then the **convex combination** $\lambda\nu_1 + (1-\lambda)\nu_2$ is defined as follows: for $A \in \mathscr{F}$,

$$\big(\lambda\nu_1 + (1-\lambda)\nu_2\big)(A) \stackrel{\text{def}}{=} \lambda\nu_1(A) + (1-\lambda)\nu_2(A).$$

A set $\mathscr{M}' \subset \mathscr{M}_1(\Omega)$ is **convex** if it is stable under convex combination of its elements, that is, if $\nu_1, \nu_2 \in \mathscr{M}'$ and $\lambda \in (0,1)$ imply $\lambda\nu_1 + (1-\lambda)\nu_2 \in \mathscr{M}'$.

The following is a nice feature of the DLR approach, which the definition of Gibbs states in Chapter 3 does not enjoy in general. Let π be any specification.

Theorem 6.56. *$\mathscr{G}(\pi)$ is convex.*

Proof. Let $\mu = \lambda\nu_1 + (1-\lambda)\nu_2$, with $\nu_1, \nu_2 \in \mathscr{G}(\pi)$. For all $\Lambda \Subset \mathbb{Z}^d$,

$$\mu\pi_\Lambda = \lambda\nu_1\pi_\Lambda + (1-\lambda)\nu_2\pi_\Lambda = \lambda\nu_1 + (1-\lambda)\nu_2 = \mu\,,$$

and so $\mu \in \mathscr{G}(\pi)$. □

Since $\mathscr{G}(\pi)$ is convex, it is natural to distinguish the measures that cannot be expressed as a non-trivial convex combination of other measures of $\mathscr{G}(\pi)$.

Definition 6.57. *$\mu \in \mathscr{G}(\pi)$ is **extremal** if any decomposition of the form $\mu = \lambda\nu_1 + (1-\lambda)\nu_2$ (with $\lambda \in (0,1)$ and $\nu_1, \nu_2 \in \mathscr{G}(\pi)$) implies that $\mu = \nu_1 = \nu_2$. The set of extremal elements of $\mathscr{G}(\pi)$ is denoted by* $\mathrm{ex}\mathscr{G}(\pi)$.

This in turn raises the following questions:

1. Is $\mathrm{ex}\mathscr{G}(\pi)$ non-empty?
2. Are there properties that distinguish the elements of $\mathrm{ex}\mathscr{G}(\pi)$ from the non-extremal ones?
3. Do extremal measures have special physical significance?

We will first answer the last two questions.

6.8.1 Properties of extremal Gibbs measures

We will see that extremal Gibbs measures are characterized by the fact that they possess *deterministic* macroscopic properties. The latter properties correspond to the following family of events, called the **tail-$\boldsymbol{\sigma}$-algebra**:

$$\mathscr{T}_\infty \stackrel{\text{def}}{=} \bigcap_{\Lambda \Subset \mathbb{Z}^d} \mathscr{F}_{\Lambda^c}\,; \tag{6.61}$$

its elements are called **tail** (or **macroscopic**) **events**. Remembering that $\mathscr{F}_{\Lambda^c}$ is the σ-algebra of events that only depend on spins located outside Λ, we see that tail events are those *whose occurrence is not altered by local changes*: if $A \in \mathscr{T}_\infty$ and if ω and ω' coincide everywhere but on a finite set of vertices, then

$$\mathbf{1}_A(\omega) = \mathbf{1}_A(\omega')\,.$$

The σ-algebra $\mathscr{T}_\infty$ contains many important events. For example, particularly relevant in view of what we saw in Chapter 3, the event "the infinite-volume magnetization exists and is positive",

$$\Big\{\omega \in \Omega : \lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|} \sum_{j \in \mathsf{B}(n)} \omega_j \text{ exists and is positive}\Big\}$$

belongs to $\mathscr{T}_\infty$. Indeed, neither the existence nor the sign of the limit are altered if any finite number of spins are changed. The $\mathscr{T}_\infty$-measurable functions $f : \Omega \to \mathbb{R}$ are also called **macroscopic observables**, since they are not altered by local changes in a configuration. As the following exercise shows, their behavior contrasts sharply with that of local functions.

Exercise 6.23. *Show that non-constant $\mathscr{T}_\infty$-measurable functions are everywhere discontinuous.*

We now present the main features that characterize the elements of $\mathrm{ex}\mathscr{G}(\pi)$.

Theorem 6.58. *Let π be a specification. Let $\mu \in \mathscr{G}(\pi)$. The following conditions are equivalent characterizations of extremality.*

1. *μ is extremal.*
2. *μ is **trivial on** $\mathscr{T}_\infty$: if $A \in \mathscr{T}_\infty$, then $\mu(A)$ is either 1 or 0.*
3. *All $\mathscr{T}_\infty$-measurable functions are μ-almost surely constant.*
4. *μ has **short-range correlations**: for all $A \in \mathscr{F}$ (or, equivalently, for all $A \in \mathscr{C}$),*

$$\lim_{\Lambda \uparrow \mathbb{Z}^d} \sup_{B \in \mathscr{F}_{\Lambda^c}} \left| \mu(A \cap B) - \mu(A)\mu(B) \right| = 0 . \tag{6.62}$$

A few remarks need to be made:

- These characterizations all express the fact that, whenever a system is described by an extremal Gibbs measure, its macroscopic properties are deterministic: every macroscopic event occurs with probability 0 or 1. This is clearly a very desirable feature: as discussed at the beginning of Chapter 1, all observables associated to a given phase of a macroscopic system in thermodynamic equilibrium are determined once the thermodynamic parameters characterizing the macrostate (for example, (β, h) for an Ising ferromagnet) are fixed. Note however that, as mentioned there, the macrostate does not fully characterize the macroscopic state of the system when there is a first-order phase transition. The reason for that is made clear in the present context: all macroscopic observables are deterministic under each extremal measure, but the macrostate does not specify which of these measures is realized.
- The statement (6.62) implies that local events become asymptotically independent as the distance separating their support diverges. In fact, it even applies to non-local events, although the interpretation of the statement becomes more difficult.
- Notice also that condition 2 above provides a remarkable and far-reaching generalization of a famous result in probability theory: *Kolmogorov's* 0-1 *law*. Indeed, combined with Exercise 6.7, Theorem 6.58 implies triviality of the tail-σ-algebra associated to a collection of independent random variables indexed by $\mathbb{Z}^d$.

To prove Theorem 6.58, we first need two preliminary propositions. Since it will be convenient to specify the σ-algebra on which measures are defined, we temporarily write $\mathscr{M}_1(\Omega, \mathscr{F})$ instead of $\mathscr{M}_1(\Omega)$.

Let $\Lambda \Subset \mathbb{Z}^d$. We define the **restriction** $\mathsf{r}_\Lambda : \mathscr{M}_1(\Omega, \mathscr{F}) \to \mathscr{M}_1(\Omega, \mathscr{F}_{\Lambda^c})$ by

$$\mathsf{r}_\Lambda \mu(B) \stackrel{\text{def}}{=} \mu(B) , \quad \forall B \in \mathscr{F}_{\Lambda^c} .$$

Observe that if $g : \Omega \to \mathbb{R}$ is $\mathscr{F}_{\Lambda^c}$-measurable, then $\mathsf{r}_\Lambda \mu(g) = \mu(g)$. Using a specification π, one can define for each $\Lambda \Subset \mathbb{Z}^d$ the **extension** $\mathsf{t}^\pi_\Lambda : \mathscr{M}_1(\Omega, \mathscr{F}_{\Lambda^c}) \to \mathscr{M}_1(\Omega, \mathscr{F})$ by

$$\mathsf{t}^\pi_\Lambda \nu(A) \stackrel{\text{def}}{=} \nu\pi_\Lambda(A) , \quad \forall A \in \mathscr{F} .$$

Note that the composition of t^π_Λ with r_Λ is such that $\mathsf{t}^\pi_\Lambda \mathsf{r}_\Lambda : \mathscr{M}_1(\Omega, \mathscr{F}) \to \mathscr{M}_1(\Omega, \mathscr{F})$. We will prove the following new characterization of $\mathscr{G}(\pi)$:

Proposition 6.59. *$\mu \in \mathscr{G}(\pi)$ if and only if $\mu = \mathrm{t}^{\pi}_{\Lambda}\mathrm{r}_{\Lambda}\mu$ for all $\Lambda \Subset \mathbb{Z}^d$.*

This characterization [7] of $\mathscr{G}(\pi)$ can be interpreted as follows. Given a measure μ on $(\Omega, \mathscr{F})$, the restriction r_Λ results in a loss of information: from the measure $r_\Lambda\mu$, nothing can be said about what happens inside Λ. However, when $\mu \in \mathscr{G}(\pi)$, that lost information can be recovered using t^{π}_{Λ}: $t^{\pi}_{\Lambda} r_\Lambda \mu = \mu$. ◇

Proof of Proposition 6.59: Composing $\mathrm{t}^{\pi}_{\Lambda}$ with r_Λ gives, for all $A \in \mathscr{F}$,

$$\mathrm{t}^{\pi}_{\Lambda}\mathrm{r}_\Lambda\mu(A) = (\mathrm{r}_\Lambda\mu)\pi_\Lambda(A) = \int \pi_\Lambda(A\,|\,\omega)\mathrm{r}_\Lambda\mu(\mathrm{d}\omega) = \int \pi_\Lambda(A\,|\,\omega)\mu(\mathrm{d}\omega) = \mu\pi_\Lambda(A)\,.$$

In the third identity, we used the $\mathscr{F}_{\Lambda^{\mathrm{c}}}$-measurability of $\pi_\Lambda(A\,|\,\cdot)$. □

Let $\hat{\mathscr{F}}$ be a sub-σ-algebra of $\mathscr{F}$ and let $\mu \in \mathscr{M}_1(\Omega, \hat{\mathscr{F}})$. For a nonnegative $\hat{\mathscr{F}}$-measurable function $f : \Omega \to \mathbb{R}$ that satisfies $\mu(f) = 1$, let $f\mu \in \mathscr{M}_1(\Omega, \hat{\mathscr{F}})$ denote the probability measure whose density with respect to μ is f:

$$f\mu(A) \stackrel{\text{def}}{=} \int_A f(\omega)\mu(\mathrm{d}\omega)\,, \qquad \forall A \in \hat{\mathscr{F}}\,.$$

Observe that $f_1\mu = f_2\mu$ if and only if $f_1 = f_2$ μ-almost surely (Lemma B.42).

Lemma 6.60. *Let $\Lambda \Subset \mathbb{Z}^d$.*

1. *Let $\mu \in \mathscr{M}_1(\Omega, \mathscr{F})$ and let $f : \Omega \to \mathbb{R}_{\geq 0}$ be an $\mathscr{F}$-measurable function such that $\mu(f) = 1$. Then*
$$\mathrm{r}_\Lambda(f\mu) = \mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})\mathrm{r}_\Lambda\mu\,.$$
2. *Let $\nu \in \mathscr{M}_1(\Omega, \mathscr{F}_{\Lambda^{\mathrm{c}}})$ and let $g : \Omega \to \mathbb{R}_{\geq 0}$, be an $\mathscr{F}_{\Lambda^{\mathrm{c}}}$-measurable function such that $\nu(g) = 1$. Then*
$$\mathrm{t}^{\pi}_{\Lambda}(g\nu) = g \cdot \mathrm{t}^{\pi}_{\Lambda}\nu\,.$$

Proof. For the first item, take $B \in \mathscr{F}_{\Lambda^{\mathrm{c}}}$ and use the definition of conditional expectation:

$$\mathrm{r}_\Lambda(f\mu)(B) = \int_B f(\omega)\mu(\mathrm{d}\omega) = \int_B \mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})(\omega)\mu(\mathrm{d}\omega) = \int_B \mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})(\omega)\mathrm{r}_\Lambda\mu(\mathrm{d}\omega)\,.$$

For the second item, take $A \in \mathscr{F}$ and compute:

$$\mathrm{t}^{\pi}_{\Lambda}(g\nu)(A) = (g\nu)\pi_\Lambda(A) = g(\nu\pi_\Lambda)(A) = g \cdot \mathrm{t}^{\pi}_{\Lambda}\nu(A)\,. \tag{6.63}$$

□

Exercise 6.24. *Justify the second identity in* (6.63).

Proposition 6.61. *Let π be a specification.*

1. *Let $\mu \in \mathscr{G}(\pi)$. Let $f : \Omega \to \mathbb{R}_{\geq 0}$, $\mathscr{F}$-measurable, such that $\mu(f) = 1$. Then $f\mu \in \mathscr{G}(\pi)$ if and only if f is equal μ-almost everywhere to a $\mathscr{T}_\infty$-measurable function.*
2. *Let $\mu, \nu \in \mathscr{G}(\pi)$ be two probability measures that coincide on $\mathscr{T}_\infty$: $\mu(A) = \nu(A)$ for all $A \in \mathscr{T}_\infty$. Then $\mu = \nu$.*

We will need the following classical result, called the **backward martingale convergence theorem**: for all measurable $f:\Omega\to\mathbb{R}$, integrable with respect to μ,

$$\mu(f\,|\,\mathscr{F}_{\mathsf{B}(n)^{\mathrm{c}}}) \xrightarrow{n\to\infty} \mu(f\,|\,\mathscr{T}_\infty)\,,\ \mu\text{-a.s. and in } L^1(\mu)\,. \tag{6.64}$$

See also Theorem B.52 in Appendix B.5.

Proof of Proposition 6.61: 1. If $f\mu\in\mathscr{G}(\pi)$, we use Lemma 6.60 and Proposition 6.59 to get, for all $\Lambda\Subset\mathbb{Z}^d$,

$$f\mu = \mathsf{t}^\pi_\Lambda \mathsf{r}_\Lambda(f\mu) = \mathsf{t}^\pi_\Lambda\big\{\mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})\mathsf{r}_\Lambda\mu\big\} = \mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})\cdot \mathsf{t}^\pi_\Lambda\{\mathsf{r}_\Lambda\mu\} = \mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})\cdot\mu\,.$$

Therefore, again by Lemma B.42, this implies $f=\mu(f\,|\,\mathscr{F}_{\Lambda^{\mathrm{c}}})$ μ-almost surely. Since this holds in particular when $\Lambda=\mathsf{B}(n)$, and since $\mu(f\,|\,\mathscr{F}_{\mathsf{B}(n)^{\mathrm{c}}})\to\mu(f\,|\,\mathscr{T}_\infty)$ almost surely as $n\to\infty$ (see (6.64)), we have shown that $f=\mu(f\,|\,\mathscr{T}_\infty)$ μ-almost surely. The latter is $\mathscr{T}_\infty$-measurable, which proves the claim. Inversely, if f coincides μ-almost surely with a $\mathscr{T}_\infty$-measurable function $\tilde f$, then $f\mu=\tilde f\mu$, and, since $\tilde f$ is $\mathscr{F}_{\Lambda^{\mathrm{c}}}$-measurable for all $\Lambda\Subset\mathbb{Z}^d$,

$$(f\mu)\pi_\Lambda(A) = (\tilde f\mu)\pi_\Lambda(A) = \tilde f(\mu\pi_\Lambda)(A) = \tilde f\mu(A) = f\mu(A)\,,$$

for all $A\in\mathscr{F}$ (we used Exercise 6.24 for the second identity) and so $f\mu\in\mathscr{G}(\pi)$.

2. Define $\lambda\stackrel{\text{def}}{=}\frac12(\mu+\nu)$. Then, $\lambda\in\mathscr{G}(\pi)$ and both μ and ν are absolutely continuous with respect to λ. By the Radon–Nikodým Theorem (Theorem B.41), there exist $f,g\ge 0$, $\lambda(f)=\lambda(g)=1$, such that $\mu=f\lambda$, $\nu=g\lambda$. For all $A\in\mathscr{T}_\infty$,

$$\int_A (f-g)\,\mathrm{d}\lambda = \mu(A)-\nu(A) = 0\,.$$

But, by item 1, there exist two $\mathscr{T}_\infty$-measurable functions $\tilde f$ and $\tilde g$, λ-almost surely equal to f, respectively g. Since $A=\{\tilde f>\tilde g\}\in\mathscr{T}_\infty$, we conclude that $\lambda(f>g)=\lambda(\tilde f>\tilde g)=0$. In the same way, $\lambda(f<g)=0$ and therefore $f=g$ λ-almost surely, which implies that $\mu=\nu$. □

Proof of Theorem 6.58: 1 ⇒ 2: Assume there exists $A\in\mathscr{T}_\infty$ such that $\alpha=\mu(A)\in(0,1)$. By item 1 of Proposition 6.61, $\mu_1\stackrel{\text{def}}{=}\frac{1}{\alpha}\mathbf{1}_A\mu$ and $\mu_2\stackrel{\text{def}}{=}\frac{1}{1-\alpha}\mathbf{1}_{A^{\mathrm{c}}}\mu$ are both in $\mathscr{G}(\pi)$. But since $\mu=\alpha\mu_1+(1-\alpha)\mu_2$, μ cannot be extremal.

2 ⇒ 1: Let μ be trivial on $\mathscr{T}_\infty$, and assume that $\mu=\alpha\mu_1+(1-\alpha)\mu_2$, with $\alpha\in(0,1)$ and $\mu_1,\mu_2\in\mathscr{G}(\pi)$. Then μ_1 and μ_2 are absolutely continuous with respect to μ. Let now $A\in\mathscr{T}_\infty$. Then, since $\mu(A)$ can be either 0 or 1, $\mu_1(A)$ and $\mu_2(A)$ are either both 0, or both 1. By item 2 of Proposition 6.61, $\mu=\mu_1=\mu_2$.

2 ⇒ 3: If f is $\mathscr{T}_\infty$-measurable, each $\{f\le c\}\in\mathscr{T}_\infty$ and thus $\mu(f\le c)\in\{0,1\}$ for all c. Setting $c_*=\inf\{c:\mu(f\le c)=1\}$, we get $\mu(f=c_*)=1$.

3 ⇒ 2: If $A\in\mathscr{T}_\infty$, then $\mathbf{1}_A$ is $\mathscr{T}_\infty$-measurable. Since it must be μ-almost surely constant, we necessarily have that $\mu(A)\in\{0,1\}$.

2 ⇒ 4: Let $A\in\mathscr{F}$, $\epsilon>0$. Using (6.64) with $f=\mathbf{1}_A$, one can take n large enough so that

$$\|\mu(A\,|\,\mathscr{F}_{\mathsf{B}(n)^{\mathrm{c}}})-\mu(A\,|\,\mathscr{T}_\infty)\|_1\le\epsilon\,. \tag{6.65}$$

Since $\mu(A\,|\,\mathscr{T}_\infty)$ is $\mathscr{T}_\infty$-measurable, item 3 implies that it is μ-almost surely constant. This constant can only be $\mu(A)$, since $\mu(\mu(A\,|\,\mathscr{T}_\infty))=\mu(A)$. Then, for all

$B \in \mathscr{F}_{\mathsf{B}(n)^c}$,

$$\begin{aligned}\bigl|\mu(A\cap B)-\mu(A)\mu(B)\bigr| &= \Bigl|\int_B \{\mathbf{1}_A-\mu(A)\}\,\mathrm{d}\mu\Bigr| \\ &= \Bigl|\int_B \{\mu(A\,|\,\mathscr{F}_{\mathsf{B}(n)^c})-\mu(A\,|\,\mathscr{T}_\infty)\}\,\mathrm{d}\mu\Bigr| \le \epsilon\,.\end{aligned}$$

4 ⇒ 2: Suppose that (6.62) holds for all $A \in \mathscr{C}$. Then, $\mu(A\cap B)=\mu(A)\mu(B)$ for all $A\in\mathscr{C}$ and all $B\in\mathscr{T}_\infty$ (since $B\in\mathscr{F}_{\Lambda^c}$ for all $\Lambda \Subset \mathbb{Z}^d$). If this can be extended to

$$\mu(A\cap B)=\mu(A)\mu(B)\,,\qquad \forall A\in\mathscr{F}, B\in\mathscr{T}_\infty\,, \tag{6.66}$$

then, taking $A=B$ implies $\mu(B)=\mu(B\cap B)=\mu(B)^2$, which is only possible if $\mu(B)\in\{0,1\}$ for all $B\in\mathscr{T}_\infty$, that is, if μ is trivial on $\mathscr{T}_\infty$.

To prove (6.66), fix $B\in\mathscr{T}_\infty$ and define

$$\mathscr{D} \stackrel{\text{def}}{=} \bigl\{A\in\mathscr{F} : \mu(A\cap B)=\mu(A)\mu(B)\bigr\}\,.$$

If $A, A' \in \mathscr{D}$, with $A\subset A'$, then $\mu\bigl((A'\setminus A)\cap B\bigr)=\mu(A'\cap B)-\mu(A\cap B)=\mu(A'\setminus A)\mu(B)$, showing that $A'\setminus A\in\mathscr{D}$. Moreover, for any sequence $(A_n)_{n\ge1}\subset\mathscr{D}$ such that $A_n\uparrow A$, we have that $\mu(A\cap B)=\lim_n\mu(A_n\cap B)=\lim_n\mu(A_n)\mu(B)=\mu(A)\mu(B)$, and so $A\in\mathscr{D}$. This implies that $\mathscr{D}$ is a **Dynkin system** (see Appendix B.5). Since $\mathscr{C}\subset\mathscr{D}$ by assumption, and since $\mathscr{C}$ is an algebra, we conclude that, $\mathscr{D}=\sigma(\mathscr{C})=\mathscr{F}$ (Theorem B.36), so (6.66) holds. □

In the following exercise, we consider a non-extremal measure for the Ising model, and we provide an example of events for which the property of short-range correlations does not hold.

Exercise 6.25. *Consider the two-dimensional Ising model with* $h=0$ *and* $\beta>\beta_c(2)$. *Take any* $\lambda\in(0,1)$ *and consider the (non-extremal) Gibbs measure*

$$\mu=\lambda\mu^+_{\beta,0}+(1-\lambda)\mu^-_{\beta,0}\,.$$

Show that μ *does not satisfy* (6.62), *by taking* $A=\{\sigma_0=1\}$, $B_i=\{\sigma_i=1\}$ *and verifying that*

$$\liminf_{\|i\|_1\to\infty}\bigl|\mu(A\cap B_i)-\mu(A)\mu(B_i)\bigr|>0\,.$$

Hint: *Use the symmetry between* $\mu^+_{\beta,0}$ *and* $\mu^-_{\beta,0}$ *and the FKG inequality.*

To end this section, we mention that extremal measures of $\mathscr{G}(\pi)$ can be distinguished from each other by only considering tail events:

Lemma 6.62. *Distinct extremal measures* $\mu,\nu\in\mathrm{ex}\mathscr{G}(\pi)$ *are* ***singular****: there exists a tail event* $A\in\mathscr{T}_\infty$ *such that* $\mu(A)=0$ *and* $\nu(A)=1$.

Proof. If $\mu,\nu\in\mathscr{G}(\pi)$ are distinct, then item 2 of Proposition 6.61 shows that there must exist $A\in\mathscr{T}_\infty$ such that $\mu(A)\ne\nu(A)$. But if μ and ν are extremal, they are trivial on $\mathscr{T}_\infty$ (Theorem 6.58, item 2), so either $\mu(A)=0$ and $\nu(A)=1$, or $\mu(A^c)=0$ and $\nu(A^c)=1$. □

6.8.2 Extremal Gibbs measures and the thermodynamic limit

Since real macroscopic systems are always finite (albeit very large), the most physically relevant Gibbs measures are those that can be approximated by finite-volume Gibbs distributions, that is, those that can be obtained by a thermodynamic limit with some fixed boundary condition. It turns out that all *extremal* Gibbs measures enjoy this property:

Theorem 6.63. *Let $\mu \in \mathrm{ex}\mathscr{G}(\pi)$. Then, for μ-almost all ω,*

$$\pi_{\mathsf{B}(n)}(\cdot\,|\,\omega) \Rightarrow \mu .$$

Proof. We need to prove that, for μ-almost all ω,

$$\pi_{\mathsf{B}(n)}(C\,|\,\omega) \overset{n\to\infty}{\longrightarrow} \mu(C) \quad \forall C \in \mathscr{C} . \tag{6.67}$$

Let $C \in \mathscr{C}$. On the one hand (see (6.23)), there exists $\Omega_{n,C}$, $\mu(\Omega_{n,C}) = 1$, such that $\pi_{\mathsf{B}(n)}(C\,|\,\omega) = \mu(C\,|\,\mathscr{F}_{\mathsf{B}(n)^c})(\omega)$ for all $\omega \in \Omega_{n,C}$. On the other hand, the extremality of μ (see item 3 of Theorem 6.58) guarantees that there exists Ω_C, $\mu(\Omega_C) = 1$, such that $\mu(C) = \mu(C\,|\,\mathscr{T}_\infty)(\omega)$ for all $\omega \in \Omega_C$. Using (6.64) with $f = \mathbf{1}_C$, there also exists $\tilde\Omega_C$, $\mu(\tilde\Omega_C) = 1$, such that

$$\mu(C\,|\,\mathscr{F}_{\mathsf{B}(n)^c})(\omega) \to \mu(C\,|\,\mathscr{T}_\infty)(\omega) \quad \forall \omega \in \tilde\Omega_C .$$

Therefore, for all ω that belong to the countable intersection of all the sets Ω_C, $\tilde\Omega_C$ and $\Omega_{n,C}$, which has μ-measure 1, (6.67) holds. □

The above theorem shows yet another reason that extremal Gibbs measures are natural to consider: they can be prepared by taking limits of finite-volume systems. However, we will see in Example 6.68 that the converse statement is not true: not all limits of finite-volume systems lead to extremal states.

A more basic question at this stage is whether all Gibbs measures can be obtained with the thermodynamic limit. The following example shows that this is not the case: $\mathscr{G}(\pi)$ can contain measures that do not appear in the approach of Chapter 3 relying on the thermodynamic limit.

Example 6.64. [8] Let us consider the 3-dimensional Ising model, with $\beta > \beta_c(3)$ and $h = 0$, in the box $\mathsf{B}(n)$. We have seen in Section 3.10.7 that the sequence of finite-volume Gibbs distributions with Dobrushin boundary condition admits a converging subsequence, defining a Gibbs measure $\mu^{\mathrm{Dob}}_{\beta,0}$ satisfying, for any $\epsilon > 0$,

$$\langle \sigma_{(0,0,0)}\sigma_{(0,0,-1)} \rangle^{\mathrm{Dob}}_{\beta,0} \le -1 + \epsilon , \tag{6.68}$$

once β is large enough (see Theorem 3.60). Let us denote by $\mu^{\mathrm{Dob}}_{\beta,0}$ the corresponding Gibbs measure. Applying a global spin flip, we obtain another Gibbs measure, $\mu^{-\mathrm{Dob}}_{\beta,0} \overset{\text{def}}{=} \tau_{\mathsf{f}}\mu^{\mathrm{Dob}}_{\beta,0}$, also satisfying (6.68). Since $\mathscr{G}(\beta,0)$ is convex, $\mu \overset{\text{def}}{=} \frac12\mu^{\mathrm{Dob}}_{\beta,0} + \frac12\mu^{-\mathrm{Dob}}_{\beta,0} \in \mathscr{G}(\beta,0)$. We show that it cannot be obtained as a thermodynamic limit. Notice that $\langle\sigma_i\rangle_\mu = 0$ for all $i \in \mathbb{Z}^d$ and that one has, for any $\epsilon > 0$,

$$\langle \sigma_{(0,0,0)}\sigma_{(0,0,-1)} \rangle_\mu \le -1 + \epsilon , \tag{6.69}$$

once β is large enough. Suppose there exists a sequence $(\mu^{\eta_k}_{\mathsf{B}(n_k);\beta,0})_{k\ge1}$ converging to μ. By the FKG inequality, for all $k \ge 1$,

$$\langle \sigma_{(0,0,0)}\sigma_{(0,0,-1)} \rangle^{\eta_k}_{\mathsf{B}(n_k);\beta,0} \ge \langle \sigma_{(0,0,0)} \rangle^{\eta_k}_{\mathsf{B}(n_k);\beta,0} \langle \sigma_{(0,0,-1)} \rangle^{\eta_k}_{\mathsf{B}(n_k);\beta,0} ,$$

and thus

$$\langle \sigma_{(0,0,0)}\sigma_{(0,0,-1)}\rangle_\mu \geq \langle \sigma_{(0,0,0)}\rangle_\mu \langle \sigma_{(0,0,-1)}\rangle_\mu = 0\,.$$

This contradicts (6.69). ◇

So far, we have described general properties of extremal Gibbs measures. We still need to determine whether such measures exist in general, and what role they play in the description of $\mathscr{G}(\pi)$. Before pursuing with the general description of the theory, we illustrate some of the ideas presented so far on our favorite example.

6.8.3 More on $\mu^+_{\beta,h}$, $\mu^-_{\beta,h}$ and $\mathscr{G}(\beta,h)$

In this section, using tools specific to the Ising model, we provide more informations about $\mu^+_{\beta,h}$ and $\mu^-_{\beta,h}$.

Lemma 6.65. *$\mu^+_{\beta,h}, \mu^-_{\beta,h}$ are extremal.*

Proof. We consider $\mu^+_{\beta,h}$. We start by showing that, for any $\nu \in \mathscr{G}(\beta,h)$,

$$\nu(f) \leq \mu^+_{\beta,h}(f) \quad \text{for every nondecreasing local function } f\,. \tag{6.70}$$

Remember that, for all $\Lambda \Subset \mathbb{Z}^d$ and all boundary condition η, the FKG inequality implies that $\mu^\eta_{\Lambda;\beta,h}(f) \leq \mu^+_{\Lambda;\beta,h}(f)$ for every nondecreasing local function f (see Lemma 3.23). Therefore,

$$\nu(f) = \int \mu^\eta_{\Lambda;\beta,h}(f)\,\nu(\mathrm{d}\eta) \leq \mu^+_{\Lambda;\beta,h}(f)\,.$$

Since $\lim_{\Lambda\uparrow\mathbb{Z}^d} \mu^+_{\Lambda;\beta,h}(f) = \mu^+_{\beta,h}(f)$, this establishes (6.70). Now assume that $\mu^+_{\beta,h}$ is not extremal:

$$\mu^+_{\beta,h} = \lambda\nu_1 + (1-\lambda)\nu_2\,,$$

where $\lambda \in (0,1)$ and $\nu_1, \nu_2 \in \mathscr{G}(\beta,h)$ are both distinct from $\mu^+_{\beta,h}$. We use (6.70) as follows. First, since $\nu_1 \neq \mu^+_{\beta,h}$, there must exist a local function f_* such that $\nu_1(f_*) \neq \mu^+_{\beta,h}(f_*)$. From Lemma 3.19, we can assume that f_* is nondecreasing. Therefore, (6.70) implies that $\nu_1(f_*) < \mu^+_{\beta,h}(f_*)$ and $\nu_2(f_*) \leq \mu^+_{\beta,h}(f_*)$. Consequently,

$$\mu^+_{\beta,h}(f_*) = \lambda\nu_1(f_*) + (1-\lambda)\nu_2(f_*) < \mu^+_{\beta,h}(f_*)\,,$$

a contradiction. We conclude that $\mu^+_{\beta,h}$ is extremal. □

Since $\mu^+_{\beta,h}$ is extremal, it inherits all the properties described in Theorem 6.58. For example, property (4) of that theorem implies that the truncated 2-point function,

$$\langle \sigma_i;\sigma_j\rangle^+_{\beta,h} \stackrel{\text{def}}{=} \langle \sigma_i\sigma_j\rangle^+_{\beta,h} - \langle \sigma_i\rangle^+_{\beta,h}\langle \sigma_j\rangle^+_{\beta,h}\,,$$

tends to zero when $\|j-i\|_\infty \to \infty$. (Note that this claim was already established, by other means, in Exercise 3.15.) This can be used to obtain a Weak Law of Large Numbers:

Exercise 6.26. *Consider*

$$m_{\mathsf{B}(n)} \stackrel{\text{def}}{=} \frac{1}{|\mathsf{B}(n)|} \sum_{j \in \mathsf{B}(n)} \sigma_j \,.$$

Show that $m_{\mathsf{B}(n)} \to \mu^+_{\beta,h}(\sigma_0)$ *in* $\mu^+_{\beta,h}$*-probability. That is, for all* $\epsilon > 0$*,*

$$\mu^+_{\beta,h}\big(|m_{\mathsf{B}(n)} - \mu^+_{\beta,h}(\sigma_0)| \geq \epsilon\big) \to 0 \quad \textit{when } n \to \infty \,.$$

Hint: *Show that the variance of* $m_{\mathsf{B}(n)}$ *vanishes as* $n \to \infty$*.*

One may wonder whether the convergence of $m_{\mathsf{B}(n)}$ to $\mu^+_{\beta,h}(\sigma_0)$ proved in the previous exercise also holds almost surely. Actually, we know that

$$m \stackrel{\text{def}}{=} \limsup_{n\to\infty} m_{\mathsf{B}(n)} \tag{6.71}$$

is almost surely constant, since it is a macroscopic observable. To show that the lim sup in (6.71) is a true limit, we will use a further property of $\mu^+_{\beta,h}$.

Lemma 6.66. $\mu^+_{\beta,h}$ *and* $\mu^-_{\beta,h}$ *are ergodic.*

We start by proving the following general fact:

Lemma 6.67. *Let* $\mu \in \mathscr{M}_1(\Omega, \mathscr{F})$ *be invariant under translations. Then, for all* $A \in \mathscr{I}$*, there exists* $B \in \mathscr{T}_\infty$ *such that* $\mu(A \triangle B) = 0$*; in particular,* $\mu(A) = \mu(B)$*.*

Proof. By Lemma B.34, there exists a sequence $(C_n)_{n\geq 1} \subset \mathscr{C}$ such that $\mu(A \triangle C_n) \leq 2^{-n}$. For each n, let $\Lambda(n) \Subset \mathbb{Z}^d$ be such that $C_n \in \mathscr{C}(\Lambda(n))$. By a property already used in Lemma 6.2, we can assume that $\Lambda(n) \uparrow \mathbb{Z}^d$. For each n, let $i_n \in \mathbb{Z}^d$ be such that $\Lambda(n) \cap \theta_{i_n}\Lambda(n) = \varnothing$. Let $C'_n = \theta_{i_n} C_n$. Since A and μ are invariant,

$$\mu(A \triangle C'_n) = \mu(\theta_{i_n}(\theta_{-i_n} A \triangle C_n)) = \mu(\theta_{i_n}(A \triangle C_n)) = \mu(A \triangle C_n) \leq 2^{-n} \,.$$

Since $C'_n \in \mathscr{F}_{\Lambda(n)^c}$, we have $B \stackrel{\text{def}}{=} \bigcap_n \bigcup_{m\geq n} C'_m \in \mathscr{T}_\infty$. Moreover,

$$\mu(A \triangle B) \leq \lim_{n\to\infty} \sum_{m \geq n} \mu(A \triangle C'_m) = 0 \,. \qquad \square$$

Proof of Lemma 6.66: Since the measure $\mu^+_{\beta,h}$ is invariant under translations, it follows from Lemma 6.67 that, for all $A \in \mathscr{I}$, there exists $B \in \mathscr{T}_\infty$ such that $\mu^+_{\beta,h}(A) = \mu^+_{\beta,h}(B)$. But $\mu^+_{\beta,h}$ is extremal, therefore $\mu^+_{\beta,h}(B) \in \{0,1\}$. $\square$

Since $\mu^+_{\beta,h}$ is ergodic, and since one can always write $\sigma_j = \sigma_0 \circ \theta_{-j}$, we deduce from Theorem 6.54 that the **infinite-volume magnetization**

$$m = \lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|} \sum_{j \in \mathsf{B}(n)} \sigma_j$$

exists $\mu^+_{\beta,h}$-almost surely, and equals $\mu^+_{\beta,h}(\sigma_0)$. A similar statement holds for $\mu^-_{\beta,h}$. Since $\mu^+_{\beta,0} \neq \mu^-_{\beta,0}$ when $\beta > \beta_c(d)$, we know from Lemma 6.62 that they are also singular. The events $\{m > 0\}$ and $\{m < 0\}$ thus provide examples of two tail events on which these measures differ.

Digression: on the significance of non-extremal Gibbs measures

With the properties of extremal measures described in detail above, we can now understand better the significance of *non-extremal* Gibbs measures. We continue illustrating things on the Ising model, but this discussion applies to more general situations.

Let $\lambda \in (0,1)$ and consider the convex combination

$$\mu \stackrel{\text{def}}{=} \lambda\mu^+_{\beta,0} + (1-\lambda)\mu^-_{\beta,0}\,.$$

Assume that $d \geq 2$ and $\beta > \beta_c(d)$, so that $\mu^+_{\beta,0} \neq \mu^-_{\beta,0}$. As explained below, a natural interpretation of the coefficient λ (respectively $1-\lambda$) is as the probability that a configuration sampled from μ is "typical" of $\mu^+_{\beta,0}$ (respectively $\mu^-_{\beta,0}$). The only minor difficulty is to give a reasonable meaning to the word "typical". One possible way to do that is to consider two tail-measurable events T^+ and T^- such that

$$\mu^+_{\beta,0}(T^+) = \mu^-_{\beta,0}(T^-) = 1, \quad \mu^+_{\beta,0}(T^-) = \mu^-_{\beta,0}(T^+) = 0\,.$$

In other words the event T^+ encodes macroscopic properties that are typically (that is, almost surely) verified under $\mu^+_{\beta,0}$, and similarly for T^-; moreover T^+ and T^- allow us to distinguish between these two measures. A configuration $\omega \in \Omega$ will then be said to be typical for $\mu^+_{\beta,0}$ (resp. $\mu^-_{\beta,0}$) if $\omega \in T^+$ (resp. T^-).

Since $\mu^+_{\beta,h}$ and $\mu^-_{\beta,h}$ are extremal and distinct, we know by Lemma 6.62 that events like T^+ and T^- always exist. In the case of the Ising model, we can be more explicit. For example, since $\mu^+_{\beta,0}$ and $\mu^-_{\beta,0}$ are characterized by the probability they associate to cylinders, one can take

$$T^\pm = \bigcap_{C\in\mathscr{C}} \Big\{ \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \sum_{i\in\Lambda} \mathbf{1}_C \circ \theta_i \text{ exists and equals } \mu^\pm_{\beta,0}(C) \Big\}\,.$$

It is easy to verify that T^+ and T^- enjoy all the desired properties. First, Theorem 6.54 guarantees that $\mu^\pm_{\beta,0}(T^\pm) = 1$. Moreover, $T^+ \cap T^- = \varnothing$, since for example $\mu^+_{\beta,0}(\sigma_0) > 0 > \mu^-_{\beta,0}(\sigma_0)$.

Let us then check that if we sample a configuration according to μ, then it will be almost surely typical for either $\mu^+_{\beta,0}$ or $\mu^-_{\beta,0}$:

$$\mu(T^+ \cup T^-) \geq \lambda\,\mu^+_{\beta,0}(T^+) + (1-\lambda)\,\mu^-_{\beta,0}(T^-) = 1\,.$$

Moreover, λ is the probability that the sampled configuration is typical for $\mu^+_{\beta,0}$:

$$\mu(T^+) = \lambda\,\mu^+_{\beta,0}(T^+) = \lambda \in (0,1)\,.$$

In the same way, $1-\lambda$ is the probability that the sampled configuration is typical of $\mu^-_{\beta,0}$. Let us then ask the following question: *If the configuration sampled (under μ) was in T^+, what else can be said about its properties?* Since $\mu^+_{\beta,0}(T^+) = 1$ and $\mu^-_{\beta,0}(T^+) = 0$, we have, for all $B \in \mathscr{F}$,

$$\mu(B \cap T^+) = \lambda\,\mu^+_{\beta,0}(B \cap T^+) = \lambda\,\mu^+_{\beta,0}(B) = \mu(T^+)\,\mu^+_{\beta,0}(B)\,.$$

Therefore,

$$\mu(B \,|\, T^+) = \mu^+_{\beta,0}(B)\,.$$

In other words, *conditionally on the fact that one observes a configuration typical for $\mu^+_{\beta,0}$, the distribution is precisely given by $\mu^+_{\beta,0}$.*

For example, taking $T^+ = \{m > 0\}$, $T^- = \{m < 0\}$,

$$\mu(\cdot \mid m > 0) = \mu^+_{\beta,0}\,, \qquad \mu(\cdot \mid m < 0) = \mu^-_{\beta,0}\,.$$

This discussion shows that non-extremal Gibbs measures do not bring any new physics: everything that can be observed under such a measure is typical for one of the extremal Gibbs measures that appears in its decomposition. In this sense, the physically relevant elements of $\mathscr{G}(\pi)$ are the extremal ones. [9]

Digression: on the simplex structure for the Ising Model in $d = 2$

The nearest-neighbor Ising model on $\mathbb{Z}^2$ happens to be one of the very few models of equilibrium statistical mechanics for which the exact structure $\mathscr{G}(\pi)$ is known. We make a few comments on this fact, whose full description is beyond the scope of this book, as it might help the reader to understand the following section on the extreme decomposition. This is also very closely related to the discussion in Section 3.10.8.

We continue with $h = 0$. The following can be proved for any $\beta > \beta_c(2)$:

1. *$\mu^+_{\beta,0}$ and $\mu^-_{\beta,0}$ are the* only *extremal Gibbs states*:

   $$\mathrm{ex}\mathscr{G}(\beta,0) = \{\mu^-_{\beta,0}, \mu^+_{\beta,0}\}\,.$$

   This follows from the discussion in Section 3.10.8.

2. *Any non-extremal Gibbs measure can be expressed in a unique manner as a convex combination of those two extremal elements*: if $\mu \in \mathscr{G}(\beta,0)$, then there exists $\lambda \in [0,1]$ such that

   $$\forall B \in \mathscr{F}\,, \quad \mu(B) = \lambda \mu^+_{\beta,0}(B) + (1-\lambda)\mu^-_{\beta,0}(B)\,. \tag{6.72}$$

   This representation induces in fact a one-to-one correspondence between measures in $\mathscr{G}(\beta)$ and the corresponding coefficient $\lambda \in [0,1]$. Indeed, taking $B = \{\sigma_0 = 1\}$ in (6.72) shows that the coefficient λ associated to a measure $\mu \in \mathscr{G}(\beta)$ can be expressed as

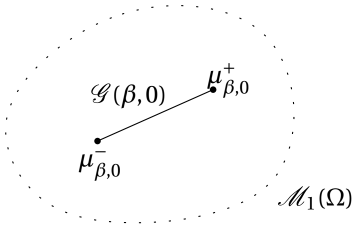
   $$\lambda = \frac{\mu(\sigma_0 = +1) - \mu^-_{\beta,0}(\sigma_0 = +1)}{\mu^+_{\beta,0}(\sigma_0 = +1) - \mu^-_{\beta,0}(\sigma_0 = +1)}\,.$$

   In view of the discussion of the previous subsection, it is natural to interpret the pair $(\lambda, 1-\lambda)$ as a probability distribution on $\mathrm{ex}\mathscr{G}(\beta,0)$.

All this can be compactly summarized by writing

$$\mathscr{G}(\beta,0) = \big\{\lambda\mu^+_{\beta,0} + (1-\lambda)\mu^-_{\beta,0},\, \lambda \in [0,1]\big\}\,. \tag{6.73}$$

This means that $\mathscr{G}(\beta,0)$ is a **simplex**: it is a closed (Lemma 6.27), convex subset of $\mathscr{M}_1(\Omega)$, which is the convex hull of its extremal elements (that is, each of its elements can be written, in a unique way, as a convex combination of the extremal elements). Schematically,



Accepting (6.73), we can clarify a point raised after Theorem 6.63: the thermodynamic limit does not always lead to an extremal state.

Example 6.68. Let $\mu^{\varnothing}_{\beta,0}$ be the Gibbs measure of the nearest-neighbor Ising model on $\mathbb{Z}^2$ prepared with free boundary condition, which is constructed in Exercise 6.14. By (6.73), $\mu^{\varnothing}_{\beta,0}$ must be a convex combination of $\mu^+_{\beta,0}$ and $\mu^-_{\beta,0}$. But by symmetry, the only possibility is that $\mu^{\varnothing}_{\beta,0} = \frac{1}{2}\mu^+_{\beta,0} + \frac{1}{2}\mu^-_{\beta,0}$. Therefore, $\mu^{\varnothing}_{\beta,0}$ is not extremal as soon as $\mu^+_{\beta,0} \neq \mu^-_{\beta,0}$ (that is, when $\beta > \beta_c(2)$), although it was constructed using the thermodynamic limit. ⋄

6.8.4 Extremal decomposition

There are unfortunately very few non-trivial specifications π for which $\mathscr{G}(\pi)$ can be determined explicitly. However, as will explained now, one can show in great generality that something similar to what we just saw in the case of the two-dimensional Ising model occurs: the set $\mathrm{ex}\mathscr{G}(\pi) \neq \varnothing$ and $\mathscr{G}(\pi)$ is always a simplex (although often an infinite-dimensional one).

Heuristics

Throughout the section, we assume that π is a specification for which $\mathscr{G}(\pi) \neq \varnothing$. (One can assume, for example, that π is quasilocal, but quasilocality itself is not necessary for the forthcoming results.) Our aim is to show that $\mathrm{ex}\mathscr{G}(\pi) \neq \varnothing$, and that any $\mu \in \mathscr{G}(\pi)$ can be expressed in a unique way as a convex combination of elements of $\mathrm{ex}\mathscr{G}(\pi)$. A priori, there can be uncountably many extremal Gibbs measures, so one can expect the combination to take the form of an integral:

$$\forall B \in \mathscr{F}, \quad \mu(B) = \int_{\mathrm{ex}\mathscr{G}(\pi)} \nu(B)\lambda_\mu(\mathrm{d}\nu). \tag{6.74}$$

Here, $\lambda_\mu(\cdot)$ is a probability distribution on $\mathrm{ex}\mathscr{G}(\pi)$ (the measurable structure on sets of probability measures will be introduced later) that plays the role of the coefficients $(\lambda, 1-\lambda)$ in (6.72); in particular,

$$\lambda_\mu(\mathrm{ex}\mathscr{G}(\pi)) = 1. \tag{6.75}$$

The main steps leading to (6.74) will be as follows. To start, for each $B \in \mathscr{F}$, the definition of the conditional expectation allows us to write

$$\mu(B) = \int \mu(B\,|\,\mathscr{T}_\infty)(\omega)\mu(\mathrm{d}\omega).$$

The central ingredient will be to show that there exists a *regular version* of $\mu(\cdot\,|\,\mathscr{T}_\infty)$. This means that one can associate to each ω a probability measure $Q^\omega \in \mathscr{M}_1(\Omega)$ in such a way that

$$\mu(\cdot\,|\,\mathscr{T}_\infty)(\omega') = Q^{\omega'}(\cdot)\,, \quad \text{for } \mu\text{-almost every } \omega'.$$

When such a family of measures Q^ω exists,

$$\mu(B) = \int Q^\omega(B)\mu(\mathrm{d}\omega)\,, \tag{6.76}$$

which is a first step towards the decomposition of $\mu(B)$ we are after.

The idea behind the construction of $Q^\cdot$ given below can be illustrated as follows (although the true construction will be more involved). Consider the basic local property characterizing the measures of $\mathscr{G}(\pi)$, written in its integral form: for all $\Lambda \Subset \mathbb{Z}^d$,

$$\mu(B) = \int \pi_\Lambda(B\,|\,\omega)\mu(\mathrm{d}\omega)\,.$$

Then, taking $\Lambda \uparrow \mathbb{Z}^d$ formally in the previous display yields

$$\mu(B) = \int \underbrace{\lim_{\Lambda\uparrow\mathbb{Z}^d} \pi_\Lambda(B\,|\,\omega)}_{Q^\omega(B)}\,\mu(\mathrm{d}\omega)\,.$$

◇

Let us give two arguments in favor of the fact that, under the map $\omega \mapsto Q^\omega$, most of the configurations ω are mapped to a $Q^\omega \in \mathrm{ex}\mathscr{G}(\pi)$, in the sense that

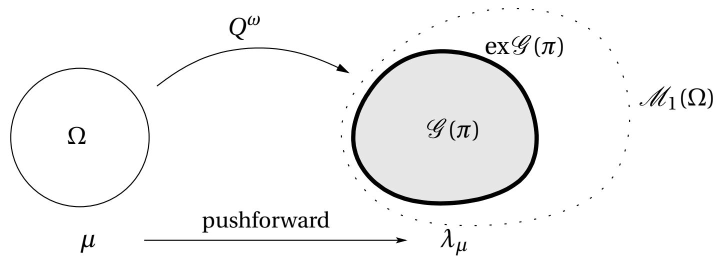
$$\mu(Q^\cdot \in \mathrm{ex}\mathscr{G}(\pi)) = 1\,.$$

1. We have already seen (remember (6.64)) that $\mu(\cdot\,|\,\mathscr{T}_\infty)$ can be expressed as a limit:
$$\mu(\cdot\,|\,\mathscr{F}_{\mathrm{B}(n)^\mathrm{c}}) \to \mu(\cdot\,|\,\mathscr{T}_\infty)\,.$$
But since $\mu \in \mathscr{G}(\pi)$, we have $\mu(\cdot\,|\,\mathscr{F}_{\mathrm{B}(n)^\mathrm{c}})(\omega) = \pi_{\mathrm{B}(n)}(\cdot\,|\,\omega)$ for μ-almost all ω and for all n. We have also seen in Theorem 6.26 that the limits of sequences $\pi_{\mathrm{B}(n)}(\cdot\,|\,\omega)$, when they exist, belong to $\mathscr{G}(\pi)$. We therefore expect that
$$Q^\omega(\cdot) \in \mathscr{G}(\pi)\,, \qquad \mu\text{-a.a.}\,\omega\,. \tag{6.77}$$

2. Moreover, if $A \in \mathscr{T}_\infty$, then $\mathbf{1}_A$ is $\mathscr{T}_\infty$-measurable and so $\mu(A\,|\,\mathscr{T}_\infty) = \mathbf{1}_A$ almost surely, which suggests that $Q^\omega(A) = \mathbf{1}_A(\omega)$; in other words, $Q^\omega(\cdot)$ should be trivial on $\mathscr{T}_\infty$, which by Theorem 6.58 means that
$$Q^\omega \in \mathrm{ex}\mathscr{G}(\pi)\,, \qquad \mu\text{-a.a.}\,\omega\,.$$

The implementation of the above argument leads to a natural way of obtaining extremal elements: first take any $\mu \in \mathscr{G}(\pi)$, then condition it with respect to $\mathscr{T}_\infty$ and get (almost surely): $\mu(\cdot\,|\,\mathscr{T}_\infty) \in \mathrm{ex}\mathscr{G}(\pi)$. ◇

One should thus consider $\omega \mapsto Q^\omega$, roughly, as a mapping from Ω to $\mathrm{ex}\mathscr{G}(\pi)$:



We would like to push μ forward onto $\mathscr{G}(\pi)$. Leaving aside the measurability issues, we proceed by letting, for $M \subset \mathscr{G}(\pi)$,

$$\lambda_\mu(M) \stackrel{\text{def}}{=} \mu(Q^\cdot \in M). \tag{6.78}$$

We then proceed as in elementary probability, and push the integration of μ over Ω onto an integration of λ_μ over $\mathrm{ex}\mathscr{G}(\pi)$. Namely [1], for a function $\varphi : \mathscr{M}_1(\Omega) \to \mathbb{R}$,

$$\int_\Omega \varphi(Q^\omega)\mu(\mathrm{d}\omega) = \int_{\mathrm{ex}\mathscr{G}(\pi)} \varphi(\nu)\lambda_\mu(\mathrm{d}\nu). \tag{6.79}$$

If one defines, for all $B \in \mathscr{F}$, the **evaluation map** $e_B : \mathscr{M}_1(\Omega) \to [0,1]$ by

$$e_B(\nu) \stackrel{\text{def}}{=} \nu(B), \tag{6.80}$$

then (6.79) with $\varphi = e_B$ and (6.76) give (6.74).

Implementing the idea exposed in the two arguments given above is not trivial (albeit mostly technical); it will be rigorously established in Propositions 6.69 and 6.70 below.

Construction and properties of the kernel $Q^\cdot$

The family $\{Q^\omega\}_{\omega\in\Omega}$ is nothing but a regular conditional distribution for $\mu(\cdot \mid \mathscr{T}_\infty)$; it will be constructed using only the kernels of π. $Q^\cdot$ will be defined by a **probability kernel from $\mathscr{T}_\infty$ to $\mathscr{F}$**, which, similarly to the kernels introduced in Definition 6.9, is a mapping $\mathscr{F} \times \Omega \to [0,1]$, $(B,\omega) \mapsto Q^\omega(B)$ with the following properties:

- For each $\omega \in \Omega$, $B \mapsto Q^\omega(B)$ is a probability measure on $(\Omega, \mathscr{F})$.
- For each $B \in \mathscr{F}$, $\omega \mapsto Q^\omega(B)$ is $\mathscr{T}_\infty$-measurable.

Proposition 6.69. *There exists, for each $\omega \in \Omega$, a probability kernel Q^ω from $\mathscr{T}_\infty$ to $\mathscr{F}$ such that, for each $\mu \in \mathscr{G}(\pi)$,*

1. *For every bounded measurable $f : \Omega \to \mathbb{R}$,*

$$\mu(f \mid \mathscr{T}_\infty)(\cdot) = Q^\cdot(f), \quad \mu\text{-almost surely}, \tag{6.81}$$

2. *$\{Q^\cdot \in \mathscr{G}(\pi)\} \in \mathscr{T}_\infty$, and $\mu(Q^\cdot \in \mathscr{G}(\pi)) = 1$.*

[1] What we are doing here is the exact analogue of the standard operation in probability theory. There, one defines the distribution of a random variable X, $\lambda_X(\cdot) \stackrel{\text{def}}{=} P(X \in \cdot)$, and uses it to express the expectation of functions of X as integrations over $\mathbb{R}$:

$$\int_\Omega g(X(\omega))P(\mathrm{d}\omega) = \int_{\mathbb{R}} g(x)\lambda_X(\mathrm{d}x).$$

Proof. ► *Construction of* Q^ω. Let $\pi = \{\pi_\Lambda\}_{\Lambda \Subset \mathbb{Z}^d}$ and let

$$\Omega_\pi \stackrel{\text{def}}{=} \bigcap_{C \in \mathscr{C}} \bigl\{\omega \in \Omega : \lim_{n\to\infty} \pi_{\mathsf{B}(n)}(C\,|\,\omega) \text{ exists}\bigr\}.$$

Clearly, $\Omega_\pi \in \mathscr{T}_\infty$. When $\omega \in \Omega_\pi^{\mathrm{c}}$, we define $Q^\omega \stackrel{\text{def}}{=} \mu_0$, where μ_0 is any fixed probability measure on Ω. When $\omega \in \Omega_\pi$, we define

$$Q^\omega(C) \stackrel{\text{def}}{=} \lim_{n\to\infty} \pi_{\mathsf{B}(n)}(C\,|\,\omega),$$

for each $C \in \mathscr{C}$. By construction, Q^ω is a probability measure on $\mathscr{C}$. By Theorem 6.96, it extends uniquely to $\mathscr{F}$. To prove $\mathscr{T}_\infty$-measurability, let

$$\mathscr{D} \stackrel{\text{def}}{=} \bigl\{B \in \mathscr{F} : \omega \mapsto Q^\omega(B) \text{ is } \mathscr{T}_\infty\text{-measurable}\bigr\}.$$

When $C \in \mathscr{C}$, we have, for all α, $\{Q^\cdot(C) \le \alpha\} = \bigl(\{Q^\cdot(C) \le \alpha\} \cap \Omega_\pi\bigr) \cup \bigl(\{Q^\cdot(C) \le \alpha\} \cap \Omega_\pi^{\mathrm{c}}\bigr) \in \mathscr{T}_\infty$. Therefore, $\mathscr{C} \subset \mathscr{D}$. We verify that $\mathscr{D}$ is a Dynkin class (see Appendix B.5): if $B, B' \in \mathscr{D}$, with $B \subset B'$, then $Q^\cdot(B' \setminus B) = Q^\cdot(B') - Q^\cdot(B)$ is $\mathscr{T}_\infty$-measurable, giving $B' \setminus B \in \mathscr{D}$. Then, if $(B_n)_{n\ge1} \subset \mathscr{D}$, $B_n \subset B_{n+1}$, $B_n \uparrow B$, then $Q^\cdot(B) = \lim_n Q^\cdot(B_n)$, and so $B \in \mathscr{D}$. Since $\mathscr{C}$ is stable under intersections, Theorem B.36 implies that $\mathscr{D} = \mathscr{F}$.

► *Relating* Q^ω *to* $\mathscr{G}(\pi)$. Let now $\mu \in \mathscr{G}(\pi)$. We have already seen in the proof of Theorem 6.63 that, on a set of μ-measure 1,

$$\pi_{\mathsf{B}(n)}(C\,|\,\cdot) = \mu(C\,|\,\mathscr{F}_{\mathsf{B}(n)^{\mathrm{c}}})(\cdot) \xrightarrow{n\to\infty} \mu(C\,|\,\mathscr{T}_\infty)(\cdot) \quad \text{for all } C \in \mathscr{C}.$$

In particular, $\mu(\Omega_\pi) = 1$ and $\mu(C\,|\,\mathscr{T}_\infty) = Q^\cdot(C)$ μ-almost surely, for all $C \in \mathscr{C}$. Again, we can show that

$$\mathscr{D}' \stackrel{\text{def}}{=} \bigl\{B \in \mathscr{F} : \mu(B\,|\,\mathscr{T}_\infty) = Q^\cdot(B)\,\mu\text{-a.s.}\bigr\}$$

is a Dynkin class containing $\mathscr{C}$, giving $\mathscr{D}' = \mathscr{F}$. To show (6.81), one can assume that f is non-negative, and take any sequence of simple functions $f_n \uparrow f$. Since each f_n is a finite sum of indicators and since $\mu(B\,|\,\mathscr{T}_\infty) = Q^\cdot(B)$ μ-a.s. for all $B \in \mathscr{F}$, it follows that $\mu(f_n\,|\,\mathscr{T}_\infty) = Q^\cdot(f_n)$ almost surely. The result follows by the monotone convergence theorem.

To show that $\{Q^\cdot \in \mathscr{G}(\pi)\} \in \mathscr{T}_\infty$, we observe that

$$\begin{aligned}\{Q^\cdot \in \mathscr{G}(\pi)\} &= \bigcap_{\Lambda \Subset \mathbb{Z}^d} \bigcap_{A \in \mathscr{F}} \{Q^\cdot \pi_\Lambda(A) = Q^\cdot(A)\}\\ &= \bigcap_{\Lambda \Subset \mathbb{Z}^d} \bigcap_{C \in \mathscr{C}} \{Q^\cdot \pi_\Lambda(C) = Q^\cdot(C)\}. \end{aligned} \tag{6.82}$$

We used Lemma 6.22 in the second equality to obtain a countable intersection (over $C \in \mathscr{C}$). Since each $Q^\cdot(C)$ is $\mathscr{T}_\infty$-measurable, $Q^\cdot\pi_\Lambda(C)$ also is. Indeed, one can consider a sequence of simple functions $f_n \uparrow \pi_\Lambda(C\,|\,\cdot)$, giving $Q^\cdot\pi_\Lambda(C) = \lim_n Q^\cdot(f_n)$. Since each $Q^\cdot(f_n)$ is $\mathscr{T}_\infty$-measurable, its limit also is. This implies that each set $\{Q^\cdot\pi_\Lambda(C) = Q^\cdot(C)\} \in \mathscr{T}_\infty$ and, therefore, $\{Q^\cdot \in \mathscr{G}(\pi)\} \in \mathscr{T}_\infty$.

Now, if $\mu \in \mathscr{G}(\pi)$, we will show that $\mu(Q^\cdot\pi_\Lambda(C) = Q^\cdot(C)) = 1$ for all $C \in \mathscr{C}$, which with (6.82) implies $\mu(Q^\cdot \in \mathscr{G}(\pi)) = 1$, thus completing the proof of the proposition. Using (6.81), $\mathscr{F}_{\Lambda^{\mathrm{c}}} \supset \mathscr{T}_\infty$ and the tower property of conditional expectation (the third to fifth inequalities below hold for μ-almost all ω),

$$\begin{aligned} Q^\omega \pi_\Lambda(C) &= Q^\omega\big(\pi_\Lambda(C\,|\,\cdot)\big) \\ &= Q^\omega\big(\mu(C\,|\,\mathscr{F}_{\Lambda^c})\big) \\ &= \mu\big(\mu(C\,|\,\mathscr{F}_{\Lambda^c})\,|\,\mathscr{T}_\infty\big)(\omega) \\ &= \mu(C\,|\,\mathscr{T}_\infty)(\omega) \\ &= Q^\omega(C)\,. \end{aligned}$$

□

Proposition 6.70. *If* $\mu \in \mathscr{G}(\pi)$*, then* $\mu(Q^\cdot \in \mathrm{ex}\mathscr{G}(\pi)) = 1$.

Note that this result has the following immediate, but crucial, consequence:

Corollary 6.71. *If* $\mathscr{G}(\pi) \neq \varnothing$*, then* $\mathrm{ex}\mathscr{G}(\pi) \neq \varnothing$.

The proof of Proposition 6.70 will rely partly on the characterization of $\mathscr{T}_\infty$ given in Theorem 6.58: extremal measures of $\mathscr{G}(\pi)$ are those that are trivial on $\mathscr{T}_\infty$. Furthermore, we have:

Exercise 6.27. *Show that* $\nu \in \mathscr{M}_1(\Omega)$ *is trivial on* $\mathscr{T}_\infty$ *if and only if, for all* $B \in \mathscr{F}$, $\nu(B\,|\,\mathscr{T}_\infty) = \nu(B)$ *ν-almost surely.* Hint: *half of the claim was already given in the proof of Theorem 6.58.*

Proof of Proposition 6.70: Using Exercise 6.27,

$$\begin{aligned} \mathrm{ex}\mathscr{G}(\pi) &= \big\{\nu \in \mathscr{G}(\pi) : \nu \text{ is trivial on } \mathscr{T}_\infty\big\} \\ &= \big\{\nu \in \mathscr{G}(\pi) : \forall A \in \mathscr{F}, \nu(A\,|\,\mathscr{T}_\infty) = \nu(A)\,, \nu\text{-a.s.}\big\} \\ &= \big\{\nu \in \mathscr{G}(\pi) : \forall C \in \mathscr{C}, \nu(C\,|\,\mathscr{T}_\infty) = \nu(C)\,, \nu\text{-a.s.}\big\} \\ &= \big\{\nu \in \mathscr{G}(\pi) : \forall C \in \mathscr{C}, Q^\cdot(C) = \nu(C)\,, \nu\text{-a.s.}\big\}\,. \end{aligned} \tag{6.83}$$

To prove the third identity, define $\mathscr{D}'' \stackrel{\text{def}}{=} \big\{A \in \mathscr{F} : \nu(A\,|\,\mathscr{T}_\infty) = \nu(A)\,, \nu\text{-a.s.}\big\}$. Since $\mathscr{D}'' \supset \mathscr{C}$ and since $\mathscr{D}''$ is a Dynkin class (as can be verified easily), we have $\mathscr{D}'' = \mathscr{F}$. For all $C \in \mathscr{C}$, $\nu \in \mathscr{M}_1(\Omega)$, let $V_C(\nu)$ denote the variance of $Q^\cdot(C)$ under ν:

$$V_C(\nu) \stackrel{\text{def}}{=} E_\nu\big[(Q^\cdot(C) - E_\nu[Q^\cdot(C)])^2\big]\,.$$

If $\nu \in \mathscr{G}(\pi)$, then $E_\nu[Q^\cdot(C)] = \nu(C)$ (because of (6.81)), and so

$$\mathrm{ex}\mathscr{G}(\pi) = \mathscr{G}(\pi) \cap \bigcap_{C \in \mathscr{C}} \big\{\nu \in \mathscr{M}_1(\Omega) : V_C(\nu) = 0\big\}\,.$$

Let $\mu \in \mathscr{G}(\pi)$. Since $\mu(Q^\cdot \in \mathscr{G}(\pi)) = 1$ (Proposition 6.69), we need to show that $\mu(V_C(Q^\cdot) = 0) = 1$ for each $C \in \mathscr{C}$. Since $V_C \geq 0$, it suffices to show that $\mu(V_C(Q^\cdot)) = 0$:

$$\begin{aligned} \mu(V_C(Q^\cdot)) &= \int \big\{E_{Q^\omega}[Q^\cdot(C)^2] - Q^\omega(C)^2\big\}\,\mu(\mathrm{d}\omega) \\ &= \int \big\{E_\mu[Q^\cdot(C)^2\,|\,\mathscr{T}_\infty](\omega) - Q^\omega(C)^2\big\}\,\mu(\mathrm{d}\omega) = 0\,, \end{aligned}$$

where we used (6.81) for the second identity. □

Construction and uniqueness of the decomposition

Let $\mu \in \mathscr{G}(\pi)$. Since we are interested in having (6.79) valid for the evaluation maps e_B, we consider the smallest σ-algebra on $\mathscr{M}_1(\Omega)$ for which all the maps $\{e_B, B \in \mathscr{F}\}$ are measurable. λ_μ in (6.78) then defines a probability measure on this σ-algebra, and satisfies (6.75).

Theorem 6.72. *For all $\mu \in \mathscr{G}(\pi)$,*

$$\forall B \in \mathscr{F}, \quad \mu(B) = \int_{\mathrm{ex}\mathscr{G}(\pi)} \nu(B)\, \lambda_\mu(\mathrm{d}\nu)\,. \tag{6.84}$$

Moreover, λ_μ is the unique measure on $\mathscr{M}_1(\Omega)$ for which such a representation holds.

Proof. The construction of (6.79) is standard. The definition $\lambda_\mu(M) \stackrel{\text{def}}{=} \mu(Q^\cdot \in M)$ can be expressed in terms of indicators:

$$\int_\Omega \mathbf{1}_M(Q^\omega)\, \mu(\mathrm{d}\omega) = \int_{\mathrm{ex}\mathscr{G}(\pi)} \mathbf{1}_M(\nu)\, \lambda_\mu(\mathrm{d}\nu).$$

An arbitrary bounded measurable function $\varphi : \mathscr{M}_1(\Omega) \to \mathbb{R}$ can be approximated by a sequence of finite linear combinations of indicator functions $\mathbf{1}_M$. Applying this with $\varphi = e_B$ yields (6.84).

Assume now that there exists another measure λ'_μ such that (6.84) holds with λ'_μ in place of λ_μ. Observe that any $\nu \in \mathrm{ex}\mathscr{G}(\pi)$ satisfies $\nu(Q^\cdot = \nu) = 1$ (see (6.83)). This implies that, for a measurable $M \subset \mathscr{M}_1(\Omega)$, $\nu(Q^\cdot \in M) = \mathbf{1}_M(\nu)$. Therefore,

$$\begin{aligned}\lambda'_\mu(M) &= \int_{\mathrm{ex}\mathscr{G}(\pi)} \mathbf{1}_M(\nu)\, \lambda'_\mu(\mathrm{d}\nu)\\ &= \int_{\mathrm{ex}\mathscr{G}(\pi)} \nu(Q^\cdot \in M)\, \lambda'_\mu(\mathrm{d}\nu) = \mu(Q^\cdot \in M) = \lambda_\mu(M)\,,\end{aligned}$$

where we used (6.84) for the third identity. This shows that $\lambda_\mu = \lambda'_\mu$. □

The fact that any $\mu \in \mathscr{G}(\pi)$ can be decomposed over the extremal elements of $\mathscr{G}(\pi)$ is convenient when trying to establish uniqueness. Indeed, to show that $\mathscr{G}(\pi)$ is a singleton, by Theorem 6.72, it suffices to show that it contains a unique extremal element. Since the latter have distinguishing properties, proving that there is only one is often simpler. This is seen in the following proof of our result on uniqueness for one-dimensional systems, stated in Section 6.5.5.

Proof of Theorem 6.40: The proof consists in showing that $\mathscr{G}(\Phi)$ has a unique *extremal* measure. By Theorem 6.72, this implies that $\mathscr{G}(\Phi)$ is a singleton.

Let therefore $\mu, \nu \in \mathrm{ex}\mathscr{G}(\Phi)$. By Theorem 6.63, μ and ν can be constructed as thermodynamic limits. Let ω (respectively η) be such that $\pi^\Phi_{\mathsf{B}(N)}(\cdot\,|\,\omega) \Rightarrow \mu$ (respectively $\pi^\Phi_{\mathsf{B}(N)}(\cdot\,|\,\eta) \Rightarrow \nu$) as $N \to \infty$. By Lemma 6.43, we thus have, for all cylinders $C \in \mathscr{C}$,

$$\mu(C) = \lim_{N\to\infty} \pi^\Phi_{\mathsf{B}(N)}(C\,|\,\omega) \geq e^{-2D} \lim_{N\to\infty} \pi^\Phi_{\mathsf{B}(N)}(C\,|\,\eta) = e^{-2D}\nu(C)\,.$$

It is easy to verify that $\mathscr{D} = \{A \in \mathscr{F} : \mu(A) \geq e^{-2D}\nu(A)\}$ is a monotone class. Since it contains the algebra $\mathscr{C}$, it also coincides with $\mathscr{F}$, and so $\mu \geq e^{-2D}\nu$. In particular, ν is absolutely continuous with respect to μ. Since two distinct extremal measures are mutually singular (see Lemma 6.62), we conclude that $\mu = \nu$. □

6.9 The variational principle

The DLR formalism studied in the present chapter characterizes the Gibbs measures describing infinite systems through a collection of local conditions: $\mu\pi_\Lambda^\Phi = \mu$ for all $\Lambda \Subset \mathbb{Z}^d$. In this section, we present an alternative, variational characterization of translation-invariant Gibbs measures, that allows to establish a relationship between the DLR formalism and the way equilibrium is described in thermostatics, as had been presented in the introduction.

The idea behind the variational principle is of a different nature, and has a more thermodynamical flavor. It will only apply to *translation-invariant Gibbs measures*, and consists in defining an appropriate functional on the set of all translation-invariant probability measures, $\mathscr{W} : \mathscr{M}_{1,\theta}(\Omega) \to \mathbb{R}$, of the form

$$\mu \mapsto \mathscr{W}(\mu) = \text{Entropy}(\mu) - \beta \times \text{Energy}(\mu)\,. \tag{6.85}$$

By analogy with (1.17) of Chapter 1, $-\frac{1}{\beta}\mathscr{W}(\mu)$ can be interpreted as the *free energy*. As seen at various places in Chapter 1, in particular in Section 1.3, equilibrium states are characterized as those that minimize the free energy; here, we will see that translation-invariant Gibbs measures are the minimizers of $-\frac{1}{\beta}\mathscr{W}(\cdot)$.

Remark 6.73. Since the relation between the material presented in this chapter and thermostatics is important from the physical point of view, we will again resort to the physicists' conventions regarding the inverse temperature and write, in particular, the potential as $\beta\Phi$. To ease notations, the temperature will usually not be explicitly indicated. ◇

6.9.1 Formulation in the finite case

To illustrate the content of the variational principle, let us consider the simplest case of a system living in a finite set $\Lambda \Subset \mathbb{Z}^d$. Let $\mathscr{M}_1(\Omega_\Lambda)$ denote the set of all probability distributions on Ω_Λ and let $\mathscr{H}_\Lambda : \Omega_\Lambda \to \mathbb{R}$ be a Hamiltonian. Define, for each $\mu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$,

$$\mathscr{W}_\Lambda(\mu_\Lambda) \stackrel{\text{def}}{=} S_\Lambda(\mu_\Lambda) - \beta\langle\mathscr{H}_\Lambda\rangle_{\mu_\Lambda}\,, \tag{6.86}$$

where $S_\Lambda(\mu_\Lambda)$ is the Shannon entropy of μ_Λ, which was already considered in Chapter 1. Here, we denote it by

$$S_\Lambda(\mu_\Lambda) \stackrel{\text{def}}{=} -\sum_{\omega_\Lambda\in\Omega_\Lambda} \mu_\Lambda(\omega_\Lambda)\log\mu_\Lambda(\omega_\Lambda)\,, \tag{6.87}$$

and $\langle\mathscr{H}_\Lambda\rangle_{\mu_\Lambda}$ represents the **average energy** under μ_Λ. Our goal is to *maximize* $\mathscr{W}_\Lambda(\cdot)$ over all probability distributions $\mu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$.

Notice that, at high temperature (small β), the dominant term is the entropy and $\mathscr{W}_\Lambda$ is maximal for the uniform distribution (remember Lemma 1.9). On the other hand, at low temperature (β large) the dominant term is the energy and $\mathscr{W}_\Lambda$ is maximal for distributions with a minimal energy. ◇

As we have already seen in Chapter 1, when μ_Λ is the Gibbs distribution associated to the Hamiltonian $\mathscr{H}_\Lambda$, $\mu_\Lambda^{\text{Gibbs}}(\omega_\Lambda) \stackrel{\text{def}}{=} \frac{e^{-\beta\mathscr{H}_\Lambda(\omega_\Lambda)}}{\mathbf{Z}_\Lambda}$, a simple computation shows that $\mathscr{W}(\mu_\Lambda^{\text{Gibbs}})$ coincides with the pressure of the system:

$$\mathscr{W}_\Lambda(\mu_\Lambda^{\text{Gibbs}}) = \log \mathbf{Z}_\Lambda\,.$$

Lemma 6.74 (Variational principle, finite version)**.** *For all* $\mu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$*,*

$$\mathscr{W}_\Lambda(\mu_\Lambda) \le \mathscr{W}_\Lambda(\mu_\Lambda^{\mathrm{Gibbs}}).$$

Moreover, $\mu_\Lambda^{\mathrm{Gibbs}}$ *is the unique maximizer of* $\mathscr{W}_\Lambda(\cdot)$*.*

Proof. Since $\log(\cdot)$ is concave, Jensen's inequality gives, for all $\mu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$,

$$\mathscr{W}_\Lambda(\mu_\Lambda) = \sum_{\omega_\Lambda\in\Omega_\Lambda} \mu_\Lambda(\omega_\Lambda) \log \frac{e^{-\beta\mathscr{H}_\Lambda(\omega_\Lambda)}}{\mu_\Lambda(\omega_\Lambda)} \le \log \sum_{\omega_\Lambda\in\Omega_\Lambda} e^{-\beta\mathscr{H}_\Lambda(\omega_\Lambda)} = \log \mathbf{Z}_\Lambda\,,$$

and equality holds if and only if $\frac{e^{-\beta\mathscr{H}_\Lambda(\omega_\Lambda)}}{\mu_\Lambda(\omega_\Lambda)}$ is constant, that is, if $\mu_\Lambda = \mu_\Lambda^{\mathrm{Gibbs}}$. □

The variational principle can be expressed in a slightly different form, useful to understand what will be done later. Let us define the **relative entropy** of two distributions $\mu_\Lambda, \nu_\Lambda \in \mathscr{M}_1(\Omega_\Lambda)$ by

$$\mathsf{H}_\Lambda(\mu_\Lambda\,|\,\nu_\Lambda) \stackrel{\mathrm{def}}{=} \begin{cases} \sum_{\omega_\Lambda\in\Omega_\Lambda} \mu_\Lambda(\omega_\Lambda) \log \frac{\mu_\Lambda(\omega_\Lambda)}{\nu_\Lambda(\omega_\Lambda)} & \text{if } \mu_\Lambda \ll \nu_\Lambda\,,\\ +\infty & \text{otherwise.}\end{cases} \tag{6.88}$$

(The interested reader can find a discussion of relative entropy and its basic properties in Appendix B.12.) First, H_Λ can be related to S_Λ by noting that, if λ_Λ denotes the uniform measure on Ω_Λ,

$$\mathsf{H}_\Lambda(\mu_\Lambda\,|\,\lambda_\Lambda) = \log|\Omega_\Lambda| - S_\Lambda(\mu_\Lambda)\,. \tag{6.89}$$

Observe also that

$$\mathsf{H}_\Lambda(\mu_\Lambda\,|\,\mu_\Lambda^{\mathrm{Gibbs}}) = \mathscr{W}_\Lambda(\mu_\Lambda^{\mathrm{Gibbs}}) - \mathscr{W}_\Lambda(\mu_\Lambda)\,, \tag{6.90}$$

so that the variational principle above can be reformulated as follows:

$$\mathsf{H}_\Lambda(\mu_\Lambda\,|\,\mu_\Lambda^{\mathrm{Gibbs}}) \ge 0\,, \quad \textit{with equality if and only if} \quad \mu_\Lambda = \mu_\Lambda^{\mathrm{Gibbs}}\,. \tag{6.91}$$

Exercise 6.28. *Let* $\mathscr{H}^{\mathrm{per}}_{V_n;\beta,h}$ *be the Hamiltonian of the Ising model in* $V_n \stackrel{\mathrm{def}}{=} \{0,\dots,n-1\}^d$ *with periodic boundary condition. Show that, among all product probability measures* $\mu_{V_n} = \bigotimes_{i\in V_n}\rho_i$ *on* $\{\pm1\}^{V_n}$ *(where all* ρ_i *are equal), the unique measure maximizing*

$$\mathscr{W}_{V_n}(\mu_{V_n}) \stackrel{\mathrm{def}}{=} S_{V_n}(\mu_{V_n}) - \beta\big\langle \mathscr{H}^{\mathrm{per}}_{V_n;\beta,h}\big\rangle_{\mu_{V_n}}$$

is the measure such that $\rho_i = \nu$ *for all* $i \in V_n$*, where* ν *is the probability measure on* $\{\pm1\}$ *with mean* m *satisfying* $m = \tanh(2d\beta m + h)$*. In other words, the maximum is achieved by the product measure obtained through the "naive mean-field approach" of Section 2.5.1. In this sense, the latter is the best approximation of the original model among all product measures.*

The rest of this section consists in extending this point of view to infinite systems. To formulate the variational principle for infinite-volume Gibbs measures, we will need to introduce notions playing the role of the entropy and average energy for infinite systems. This will be done by considering the corresponding *densities*:

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} S_\Lambda(\mu_\Lambda)\,, \qquad \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \langle \mathscr{H}_\Lambda\rangle_{\mu_\Lambda}\,.$$

The existence of these two limits will be established when μ_Λ is the marginal of a translation-invariant measure, in Propositions 6.75 and 6.78.

Remember that $\mathscr{G}(\beta\Phi)$ denotes the set of all probability measures compatible with the Gibbsian specification $\pi^{\beta\Phi}$, and $\mathscr{G}_\theta(\beta\Phi) \stackrel{\text{def}}{=} \mathscr{G}(\beta\Phi) \cap \mathscr{M}_{1,\theta}(\Omega)$.

In the following sections, we will define a functional $\mathscr{W} : \mathscr{M}_{1,\theta}(\Omega) \mapsto \mathbb{R}$, using the densities mentioned above, and characterize the Gibbs measures of $\mathscr{G}_\theta(\beta\Phi)$ as maximizers of $\mathscr{W}(\cdot)$. Notice that, since first-order phase transitions can occur on the infinite lattice, we do not expect uniqueness of the maximizer to hold in general.

6.9.2 Specific entropy and energy density

Remember from the beginning of the chapter that the marginal of $\mu \in \mathscr{M}_1(\Omega)$ on $\Lambda \Subset \mathbb{Z}^d$ is $\mu|_\Lambda \stackrel{\text{def}}{=} \mu \circ \Pi_\Lambda^{-1} \in \mathscr{M}_1(\Omega_\Lambda)$. Since no confusion will be possible below, we will write μ_Λ instead of $\mu|_\Lambda$, to lighten the notations. One can then define the **Shannon entropy of μ in Λ** by

$$S_\Lambda(\mu) \stackrel{\text{def}}{=} S_\Lambda(\mu_\Lambda)\,,$$

where $S_\Lambda(\mu_\Lambda)$ was defined in (6.87) (bearing in mind that μ_Λ is now the marginal of μ in Λ).

Proposition 6.75 (Existence of the specific entropy)**.** *For all $\mu \in \mathscr{M}_{1,\theta}(\Omega)$,*

$$s(\mu) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|} S_{\mathsf{B}(n)}(\mu) \tag{6.92}$$

*exists and is called the **specific entropy** of μ. Moreover,*

$$s(\mu) = \inf_{\Lambda\in\mathscr{R}} \frac{S_\Lambda(\mu)}{|\Lambda|}\,, \tag{6.93}$$

*and $\mu \mapsto s(\mu)$ is **affine**: for all $\mu,\nu \in \mathscr{M}_{1,\theta}(\Omega)$ and $\alpha \in (0,1)$,*

$$s(\alpha\mu + (1-\alpha)\nu) = \alpha s(\mu) + (1-\alpha)s(\nu)\,.$$

Given $\mu,\nu \in \mathscr{M}_1(\Omega)$, let us use (6.88) to define the **relative entropy of μ with respect to ν (on Λ)**:

$$\mathsf{H}_\Lambda(\mu\,|\,\nu) \stackrel{\text{def}}{=} \mathsf{H}_\Lambda(\mu_\Lambda\,|\,\nu_\Lambda)\,.$$

Lemma 6.76. *For all $\mu,\nu \in \mathscr{M}_1(\Omega)$ and all $\Lambda \Subset \mathbb{Z}^d$,*

1. $\mathsf{H}_\Lambda(\mu\,|\,\nu) \geq 0$*, with equality if and only if* $\mu_\Lambda = \nu_\Lambda$,
2. $(\mu,\nu) \mapsto \mathsf{H}_\Lambda(\mu,\nu)$ *is convex, and*
3. *if* $\Delta \subset \Lambda$*, then* $\mathsf{H}_\Delta(\mu\,|\,\nu) \leq \mathsf{H}_\Lambda(\mu\,|\,\nu)$.

Proof. The first and second items are proved in Proposition B.66, so let us consider the third one. We can assume that $\mu_\Lambda \ll \nu_\Lambda$, otherwise the claim is trivial. Then,

$$\mathsf{H}_\Lambda(\mu\,|\,\nu) = \sum_{\omega_\Lambda} \phi\Big(\frac{\mu_\Lambda(\omega_\Lambda)}{\nu_\Lambda(\omega_\Lambda)}\Big)\nu_\Lambda(\omega_\Lambda)\,,$$

where $\phi(x) \stackrel{\text{def}}{=} x \log x$ $(x \geq 0)$ is convex. We then split ω_Λ into $\omega_\Lambda = \tau_\Delta \eta_{\Lambda\setminus\Delta}$, and consider the summation over $\eta_{\Lambda\setminus\Delta}$, for a fixed τ_Δ. Using Jensen's inequality for the distribution ν_Λ conditioned on τ_Δ, we get

$$\mathsf{H}_\Lambda(\mu\,|\,\nu) \geq \sum_{\tau_\Delta} \phi\Big(\frac{\mu_\Delta(\tau_\Delta)}{\nu_\Delta(\tau_\Delta)}\Big)\nu_\Delta(\tau_\Delta) = \mathsf{H}_\Delta(\mu\,|\,\nu)\,. \qquad \square$$

Corollary 6.77. *$\mu \mapsto -S_\Lambda(\mu)$ is convex and, when $\Lambda, \Lambda' \Subset \mathbb{Z}^d$ are disjoint,*

$$S_{\Lambda\cup\Lambda'}(\mu) \leq S_\Lambda(\mu) + S_{\Lambda'}(\mu)\,.$$

Proof. The first claim follows from (6.89) and Lemma 6.76. A straightforward computation shows that, introducing $\nu \stackrel{\text{def}}{=} \mu_\Lambda \otimes \lambda_{\Lambda^c}$ with λ_{Λ^c} the uniform product measure on Ω_{Λ^c}, one gets

$$S_\Lambda(\mu) - S_{\Lambda\cup\Lambda'}(\mu) = \mathsf{H}_{\Lambda\cup\Lambda'}(\mu|\nu) - \log|\Omega_{\Lambda'}| \geq \mathsf{H}_{\Lambda'}(\mu|\nu) - \log|\Omega_{\Lambda'}| = -S_{\Lambda'}(\mu)\,.$$

We used again Lemma 6.76 in the inequality. $\square$

Proof of Proposition 6.75: By Corollary 6.77, the set function $a(\Lambda) \stackrel{\text{def}}{=} S_\Lambda(\mu)$ is both translation invariant and subadditive (see Section B.1.3): for all pairs of disjoint paralellepipeds Λ, Λ', $a(\Lambda \cup \Lambda') \leq a(\Lambda) + a(\Lambda')$. The existence of the limit defining $s(\mu)$ is therefore guaranteed by Lemma B.6. By Corollary 6.77, $S_\Lambda(\cdot)$ is concave, which implies that $s(\cdot)$ is concave too. To verify that it is also convex, consider $\mu' \stackrel{\text{def}}{=} \alpha\mu + (1-\alpha)\nu$. Since $\log \mu'_\Lambda(\omega_\Lambda) \geq \log(\alpha\mu_\Lambda(\omega_\Lambda))$ and $\log\mu'_\Lambda(\omega_\Lambda) \geq \log((1-\alpha)\nu_\Lambda(\omega_\Lambda))$,

$$S_\Lambda(\mu') \leq \alpha S_\Lambda(\mu) + (1-\alpha)S_\Lambda(\nu) - \alpha\log(1-\alpha) - (1-\alpha)\log(1-\alpha)\,.$$

This implies that $s(\cdot)$ is also convex, proving the second claim of the proposition. $\square$

Exercise 6.29. *Show that $s(\cdot)$ is* ***upper semicontinuous****, that is,*

$$\mu_k \Rightarrow \mu \quad \textit{implies} \quad \limsup_{k\to\infty} s(\mu_k) \leq s(\mu)\,.$$

Let us now turn our attention to Gibbs measures and consider a potential $\Phi = \{\Phi_B\}_{B\Subset\mathbb{Z}^d}$. Until the end of the section, we will assume Φ to be absolutely summable and translation invariant, like the potentials considered in Section 6.7.1. Notice that translation invariance implies that Φ is in fact *uniformly* absolutely summable:

$$\sup_{i\in\mathbb{Z}^d} \sum_{\substack{B\Subset\mathbb{Z}^d:\\ B\ni i}} \|\Phi_B\|_\infty = \sum_{\substack{B\Subset\mathbb{Z}^d:\\ B\ni 0}} \|\Phi_B\|_\infty < \infty\,.$$

Proposition 6.78 (Existence of the average energy density)**.** *For all $\mu \in \mathscr{M}_{1,\theta}(\Omega)$,*

$$\lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|} \big\langle \mathscr{H}_{\mathsf{B}(n);\Phi} \big\rangle_\mu = \langle u_\Phi \rangle_\mu\,, \tag{6.94}$$

where

$$u_\Phi \stackrel{\text{def}}{=} \sum_{\substack{B\Subset\mathbb{Z}^d:\\ B\ni 0}} \frac{1}{|B|}\Phi_B\,. \tag{6.95}$$

Proof. First,

$$\Big|\mathscr{H}_{\mathsf{B}(n);\Phi} - \sum_{j\in\mathsf{B}(n)} \theta_j u_\Phi\Big| \le \sum_{j\in\mathsf{B}(n)} \sum_{\substack{B\ni j:\\ B\not\subset\mathsf{B}(n)}} \|\Phi\|_\infty \stackrel{\text{def}}{=} r_{\mathsf{B}(n);\Phi}\,. \tag{6.96}$$

The uniform absolute summability of Φ implies that $r_{\mathsf{B}(n);\Phi} = o(|\mathsf{B}(n)|)$ (see Exercise 6.30 below). By translation invariance, $\langle\theta_j u_\Phi\rangle_\mu = \langle u_\Phi\rangle_\mu$, which concludes the proof. □

Exercise 6.30. *Show that, when Φ is absolutely summable and translation invariant,*

$$\lim_{n\to\infty} \frac{r_{\mathsf{B}(n);\Phi}}{|\mathsf{B}(n)|} = 0\,. \tag{6.97}$$

The last object whose existence in the thermodynamic limit needs to be proved is the *pressure.*

Theorem 6.79. *When Φ is absolutely summable and translation invariant,*

$$\psi(\Phi) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{\beta|\mathsf{B}(n)|} \log \mathbf{Z}^\eta_{\mathsf{B}(n);\beta\Phi}$$

exists and does not depend on the boundary condition η; it is called the ***pressure****. Moreover, $\Phi \mapsto \psi(\Phi)$ is* ***convex*** *on the space of absolutely summable, translation-invariant potentials: if Φ^1, Φ^2 are two such potentials and $t \in (0,1)$, then*

$$\psi\big(t\Phi^1 + (1-t)\Phi^2\big) \le t\psi(\Phi^1) + (1-t)\psi(\Phi^2)\,.$$

We will actually see, as a byproduct of the proof, that the pressure equals

$$\beta\psi(\Phi) = s(\mu) - \beta\langle u_\Phi\rangle_\mu\,, \qquad \forall \mu \in \mathscr{G}_\theta(\beta\Phi)\,, \tag{6.98}$$

which is the analogue of the Euler relation (1.9).

We will show existence of the pressure by using the convergence proved above for the specific entropy and average energy. To start:

Lemma 6.80. *Let $\mu \in \mathscr{M}_{1,\theta}(\Omega)$ and $(\nu_n)_{n\ge1}$, $(\tilde\nu_n)_{n\ge1}$ be two arbitrary sequences in $\mathscr{M}_1(\Omega)$. If either of the sequences*

$$\Big(\frac{1}{|\mathsf{B}(n)|} \mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)})\Big)_{n\ge1}\,, \quad \Big(\frac{1}{|\mathsf{B}(n)|} \mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\tilde\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)})\Big)_{n\ge1} \tag{6.99}$$

has a limit as $n\to\infty$, then the other one also does, and the limits are equal.

Proof. By the absolute sommability of Φ, we have $\pi^{\beta\Phi}_{\mathsf{B}(n)}(\omega_{\mathsf{B}(n)}\,|\,\eta) > 0$ for all n, uniformly in $\omega_{\mathsf{B}(n)}$ and η. This guarantees that $\mu_{\mathsf{B}(n)} \ll \nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)}$ and $\mu_{\mathsf{B}(n)} \ll \tilde\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)}$. Therefore,

$$\mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)}) - \mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\tilde\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)}) = \sum_{\omega_{\mathsf{B}(n)}} \mu_{\mathsf{B}(n)}(\omega_{\mathsf{B}(n)}) \log \frac{\tilde\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)}(\omega_{\mathsf{B}(n)})}{\nu_n\pi^{\beta\Phi}_{\mathsf{B}(n)}(\omega_{\mathsf{B}(n)})}\,.$$

Since

$$\sup_{\omega_{B(n)},\eta,\tilde\eta} \left|\mathscr{H}_{B(n);\Phi}(\omega_{B(n)}\eta_{B(n)^c}) - \mathscr{H}_{B(n);\Phi}(\omega_{B(n)}\tilde\eta_{B(n)^c})\right| \le 2r_{B(n);\Phi}, \tag{6.100}$$

where $r_{B(n);\Phi}$ was defined in (6.96), we have,

$$e^{-4\beta r_{B(n);\Phi}} \le \frac{\tilde\nu_n\pi^{\beta\Phi}_{B(n)}(\omega_{B(n)})}{\nu_n\pi^{\beta\Phi}_{B(n)}(\omega_{B(n)})} \le e^{4\beta r_{B(n);\Phi}}.$$

By Exercise 6.30, $r_{B(n);\Phi} = o(|B(n)|)$, which proves the claim. □

Proof of Proposition 6.79: We use Lemma 6.80 with $\mu \in \mathscr{G}_\theta(\beta\Phi)$, $\nu_n = \mu$, and $\tilde\nu_n = \delta_\omega$, for some arbitrary fixed $\omega \in \Omega$. Then $\nu_n\pi^{\beta\Phi}_{B(n)} = \mu$, so the first sequence in (6.99) is identically equal to zero. Therefore, the second sequence must converge to zero. By writing it explicitly, the second sequence becomes

$$\frac{1}{|B(n)|}H_{B(n)}\big(\mu\,|\,\tilde\nu_n\pi^{\beta\Phi}_{B(n)}\big) = \frac{1}{|B(n)|}\log\mathbf{Z}^{\omega}_{B(n);\beta\Phi} - \frac{1}{|B(n)|}\big\{S_{B(n)}(\mu) - \beta\big\langle\mathscr{H}_{B(n);\Phi}(\cdot\,\omega_{B(n)^c})\big\rangle_\mu\big\}. \tag{6.101}$$

By Propositions 6.75 and 6.78, the second and third terms on the right-hand side are known to have limits, and the limit of the third one does not depend on ω, since, by (6.100),

$$\big\langle\mathscr{H}_{B(n);\Phi}(\cdot\,\omega_{B(n)^c})\big\rangle_\mu = \big\langle\mathscr{H}_{B(n);\Phi}\big\rangle_\mu + O(r_{B(n);\Phi}).$$

Since the whole sequence on the right-hand side of (6.101) must converge to zero, this proves the existence of the pressure and justifies (6.98). As in the proof of Lemma 3.5, convexity follows from Hölder's inequality. □

6.9.3 Variational principle for Gibbs measures

Proposition 6.81. *Let Φ be absolutely convergent and translation invariant. Let $\nu \in \mathscr{G}_\theta(\beta\Phi)$. Then, for all $\mu \in \mathscr{M}_{1,\theta}(\Omega)$, the* ***Gibbs free energy***

$$h(\mu\,|\,\Phi) \stackrel{\text{def}}{=} \lim_{n\to\infty}\frac{1}{|B(n)|}H_{B(n)}(\mu\,|\,\nu) \tag{6.102}$$

exists and does not depend on ν (only on Φ). Moreover, $h(\mu\,|\,\Phi)$ is non-negative and satisfies

$$h(\mu\,|\,\Phi) = \beta\psi(\Phi) - \{s(\mu) - \beta\langle u_\Phi\rangle_\mu\}. \tag{6.103}$$

Proof. If $\nu \in \mathscr{G}_\theta(\beta\Phi)$, then $H_{B(n)}(\mu\,|\,\nu) = H_{B(n)}(\mu\,|\,\nu\pi^{\beta\Phi}_{B(n)})$. Therefore, to show the existence of the limit (6.102), we can use Lemma 6.80 with $\nu_n = \nu$. As earlier, we choose $\tilde\nu_n = \delta_\omega$, for which we know that the limit of the second sequence in (6.99) exists and equals $\beta\psi(\Phi) - \{s(\mu) - \beta\langle u_\Phi\rangle_\mu\}$, as seen after (6.101). □

We can now formulate the infinite-volume version of (6.91).

Theorem 6.82 (Variational principle)**.** *Let $\mu \in \mathscr{M}_{1,\theta}(\Omega)$. Then*

$$\mu \in \mathscr{G}_\theta(\beta\Phi) \quad \textit{if and only if} \quad h(\mu\,|\,\Phi) = 0. \tag{6.104}$$

The variational principle stated above establishes the analogy between translation invariant Gibbs measures and the basic principles of thermostatics, as announced at the beginning of the section.

We already know from (6.98) that $\mu \in \mathscr{G}_\theta(\beta\Phi)$ implies $\mathsf{h}(\mu\,|\,\Phi) = 0$. The proof of the converse statement is trickier; it will rely on the following lemma.

Lemma 6.83. *Let $\mu, \nu \in \mathscr{M}_{1,\theta}(\Omega)$ be such that*

$$\lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|} \mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\nu) = 0\,. \tag{6.105}$$

Fix $\Delta \Subset \mathbb{Z}^d$. Then, for all $\delta > 0$ and for all k for which $\mathsf{B}(k) \supset \Delta$, there exists some finite region $\Lambda \supset \mathsf{B}(k)$ such that

$$0 \le \mathsf{H}_\Lambda(\mu\,|\,\nu) - \mathsf{H}_{\Lambda\setminus\Delta}(\mu\,|\,\nu) \le \delta\,. \tag{6.106}$$

We know from item 1 of Lemma 6.76 that, in a finite region, $\mathsf{H}_\Lambda(\mu\,|\,\nu) = 0$ implies $\mu_\Lambda = \nu_\Lambda$. Although (6.105) *does not necessarily imply that $\mu = \nu$,* (6.106) *will imply that μ and ν can be compared to each other on arbitrarily large regions Λ (see the proof of Theorem 6.82 below).* ⋄

Proof. We will use repeatedly the monotonicity of H_Λ in Λ, proved in Lemma 6.76. Fix $\delta > 0$ and k so that $\mathsf{B}(k) \supset \Delta$, and let n be such that

$$\frac{1}{|\mathsf{B}(n)|} \mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\nu) \le \frac{\delta}{2|\mathsf{B}(k)|}\,.$$

If $m = \lfloor (2n+1)/(2k+1) \rfloor$, then at least m^d adjacent disjoint translates of $\mathsf{B}(k)$ can be arranged to fit in $\mathsf{B}(n)$; we denote them by $\mathsf{B}_1(k), \dots, \mathsf{B}_{m^d}(k)$. We assume for simplicity that $\mathsf{B}_1(k) = \mathsf{B}(k)$. For each $\ell \in \{2, \dots, m^d\}$, let i_ℓ be such that $\mathsf{B}_\ell(k) = i_\ell + \mathsf{B}(k)$. Define now $\Delta(\ell) \stackrel{\text{def}}{=} i_\ell + \Delta$ and $\Lambda(\ell) \stackrel{\text{def}}{=} \mathsf{B}_1(k) \cup \dots \cup \mathsf{B}_\ell(k)$. For commodity, let $\mathsf{H}_\varnothing(\mu\,|\,\nu) \stackrel{\text{def}}{=} 0$. Since $\Lambda(\ell) \setminus \mathsf{B}_\ell(k) \subset \Lambda(\ell) \setminus \Delta(\ell)$,

$$\begin{aligned}
\frac{1}{m^d} \sum_{\ell=1}^{m^d} \bigl\{\mathsf{H}_{\Lambda(\ell)}(\mu\,|\,\nu) - \mathsf{H}_{\Lambda(\ell)\setminus\Delta(\ell)}(\mu\,|\,\nu)\bigr\} &\le \frac{1}{m^d} \sum_{\ell=1}^{m^d} \bigl\{\mathsf{H}_{\Lambda(\ell)}(\mu\,|\,\nu) - \mathsf{H}_{\Lambda(\ell)\setminus\mathsf{B}_\ell(k)}(\mu\,|\,\nu)\bigr\} \\
&= \frac{1}{m^d} \mathsf{H}_{\Lambda(m^d)}(\mu\,|\,\nu) \\
&\le \frac{1}{m^d} \mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\nu) \\
&\le \delta\,.
\end{aligned}$$

In the last line, we used the fact that $m^d \ge |\mathsf{B}(n)|/(2|\mathsf{B}(k)|)$. Since the first sum over ℓ corresponds to the arithmetic mean of a collection of non-negative numbers, at least one of them must satisfy

$$\mathsf{H}_{\Lambda(\ell)}(\mu\,|\,\nu) - \mathsf{H}_{\Lambda(\ell)\setminus\Delta(\ell)}(\mu\,|\,\nu) \le \delta\,.$$

One can thus take $\Lambda \stackrel{\text{def}}{=} \theta_{i_\ell}^{-1} \Lambda(\ell)$; translation invariance of μ and ν yields the desired result. □

Proof of Theorem 6.82: Let $\mu \in \mathscr{M}_{1,\theta}(\Omega)$ be such that $\mathsf{h}(\mu\,|\,\Phi) = 0$. We need to show that, for any $\Delta \Subset \mathbb{Z}^d$ and any local function f,

$$\mu\pi_\Delta^{\beta\Phi}(f) = \mu(f)\,. \tag{6.107}$$

By Proposition 6.81, $\mathsf{h}(\mu\,|\,\Phi) = 0$ means that we can take any $\nu \in \mathscr{G}_\theta(\beta\Phi)$ and assume that

$$\lim_{n\to\infty} \frac{1}{|\mathsf{B}(n)|}\mathsf{H}_{\mathsf{B}(n)}(\mu\,|\,\nu) = 0\,.$$

In particular, $\mu_\Lambda \ll \nu_\Lambda$ for all $\Lambda \Subset \mathbb{Z}^d$; we will denote the corresponding Radon–Nikodým derivative by $\rho_\Lambda \stackrel{\text{def}}{=} \frac{d\mu_\Lambda}{d\nu_\Lambda}$. Observe that ρ_Λ is $\mathscr{F}_\Lambda$-measurable and that, for all $\mathscr{F}_\Lambda$-measurable functions g, $\mu(g) = \nu(\rho_\Lambda g)$.

Fix $\epsilon > 0$. To start, $\mu\pi_\Delta^{\beta\Phi}(f) = \mu(\pi_\Delta^{\beta\Phi} f)$ and, since $\pi_\Delta^{\beta\Phi} f$ is quasilocal and $\mathscr{F}_{\Delta^c}$-measurable, we can find some $\mathscr{F}_{\Delta^c}$-measurable local function g_* such that $\|\pi_\Delta^{\beta\Phi} f - g_*\|_\infty \le \epsilon$. Let then k be large enough to ensure that $\mathsf{B}(k)$ contains Δ, as well as the supports of f and g_*. In this way, g_* is $\mathscr{F}_{\mathsf{B}(k)\setminus\Delta}$-measurable.

Let $\delta \stackrel{\text{def}}{=} \frac{r\epsilon}{2}$, and take $\Lambda \supset \mathsf{B}(k)$, as in Lemma 6.83. We write

$$\begin{aligned}\mu\pi_\Delta^{\beta\Phi}(f) - \mu(f) =\,& \mu(\pi_\Delta^{\beta\Phi} f - g_*)\\ &+ \big(\mu(g_*) - \nu(\rho_{\Lambda\setminus\Delta} g_*)\big)\\ &+ \nu\big(\rho_{\Lambda\setminus\Delta}(g_* - \pi_\Delta^{\beta\Phi} f)\big)\\ &+ \nu\big(\rho_{\Lambda\setminus\Delta}(\pi_\Delta^{\beta\Phi} f - f)\big)\\ &+ \nu\big((\rho_{\Lambda\setminus\Delta} - \rho_\Lambda) f\big)\\ &+ \big(\nu(\rho_\Lambda f) - \mu(f)\big).\end{aligned}$$

We consider one by one the terms on the right-hand side of this last display. Since $\|\pi_\Delta^{\beta\Phi} f - g_*\|_\infty \le \epsilon$, the first and third terms are bounded by ϵ. The second term is zero since g_* is $\mathscr{F}_{\Lambda\setminus\Delta}$-measurable, and the sixth term is zero since f is $\mathscr{F}_\Lambda$-measurable. Now the fourth term is zero too, since $\rho_{\Lambda\setminus\Delta}$ is $\mathscr{F}_{\Delta^c}$-measurable and since $\nu \in \mathscr{G}_\theta(\beta\Phi)$ implies that

$$\nu\big(\rho_{\Lambda\setminus\Delta}(\pi_\Delta^{\beta\Phi} f)\big) = \nu\big(\pi_\Delta^{\beta\Phi}(\rho_{\Lambda\setminus\Delta} f)\big) = \nu\pi_\Delta^{\beta\Phi}(\rho_{\Lambda\setminus\Delta} f) = \nu(\rho_{\Lambda\setminus\Delta} f)\,.$$

Finally, consider the fifth term. First, notice that

$$\delta \ge \mathsf{H}_\Lambda(\mu\,|\,\nu) - \mathsf{H}_{\Lambda\setminus\Delta}(\mu\,|\,\nu) = \mu\Big(\log\frac{\rho_\Lambda}{\rho_{\Lambda\setminus\Delta}}\Big) = \nu\Big(\rho_\Lambda \log\frac{\rho_\Lambda}{\rho_{\Lambda\setminus\Delta}}\Big) = \nu\big(\rho_{\Lambda\setminus\Delta}\phi(\rho_\Lambda/\rho_{\Lambda\setminus\Delta})\big)\,,$$

where $\phi(x) \stackrel{\text{def}}{=} 1 - x + x\log x$. It can be verified that there exists $r > 0$ such that

$$\phi(x) \ge r(|x-1| - \epsilon/2) \qquad \forall x \ge 0\,. \tag{6.108}$$

Using (6.108),

$$\nu\big(\rho_{\Lambda\setminus\Delta}\phi(\rho_\Lambda/\rho_{\Lambda\setminus\Delta})\big) \ge r\nu(|\rho_\Lambda - \rho_{\Lambda\setminus\Delta}|) - \frac{r\epsilon}{2}\,.$$

By definition of δ, this implies that $\nu(|\rho_\Lambda - \rho_{\Lambda\setminus\Delta}|) \le \epsilon$. The fifth term is therefore bounded by $\epsilon\|f\|_\infty$. Altogether, $|\mu\pi_\Delta^{\beta\Phi}(f) - \mu(f)| \le (2 + \|f\|_\infty)\epsilon$, which proves (6.107) since ϵ was arbitrary. □

We have completed the program described at the beginning of the section:

Theorem 6.84. *Let Φ be an absolutely summable, translation-invariant potential. Define the affine functional $\mathscr{W} : \mathscr{M}_{1,\theta}(\Omega) \to \mathbb{R}$ by*

$$\mu \mapsto \mathscr{W}(\mu) \stackrel{\text{def}}{=} s(\mu) - \beta \langle u_\Phi \rangle_\mu .$$

Then the maximizers of $\mathscr{W}(\cdot)$ are the translation-invariant Gibbs measures compatible with $\pi^{\beta\Phi}$.

6.10 Continuous spins

As we said at the beginning of the chapter, the DLR formalism can be developed for much more general single-spin spaces. In this section, we briefly discuss some of these extensions.

As long as the single-spin space remains compact, most of the results stated and proved in the previous sections hold, although some definitions and some of the proofs given for $\Omega_0 = \{\pm 1\}$ need to be slightly adapted. In contrast, when the single-spin space is *not* compact, even the existence of Gibbs measures is not guaranteed (even for very reasonable interactions, as will be discussed in Chapter 8).

We discuss these issues briefly; for details on more general settings in which the DLR formalism can be developed, we refer the reader to [134].

6.10.1 General definitions

Let (Ω_0, d) be a separable metric space. Two guiding examples that the reader should keep in mind are $\mathbb{S}^1 = \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$ equipped with the Euclidean distance, or $\mathbb{R}$ with the usual distance $|\cdot|$.

The distance induces the Borel σ-algebra $\mathscr{B}_0$ on Ω_0, generated by the open sets. Let

$$\Omega_\Lambda \stackrel{\text{def}}{=} \Omega_0^\Lambda , \quad \Omega \stackrel{\text{def}}{=} \Omega_0^{\mathbb{Z}^d} .$$

The measurable structure on Ω_Λ is the product σ-algebra

$$\mathscr{B}_\Lambda \stackrel{\text{def}}{=} \bigotimes_{i \in \Lambda} \mathscr{B}_0 ,$$

which is the smallest σ-algebra on Ω_Λ generated by the **rectangles**, that is, the sets of the form $\times_{i \in \Lambda} A_i$, $A_i \in \mathscr{B}_0$ for all $i \in \Lambda$. The projections $\Pi_\Lambda : \Omega \to \Omega_\Lambda$ are defined as before. For all $\Lambda \Subset \mathbb{Z}^d$,

$$\mathscr{C}(\Lambda) \stackrel{\text{def}}{=} \Pi_\Lambda^{-1}(\mathscr{B}_\Lambda)$$

denotes the σ-algebra of **cylinders with base in Λ** and, for $S \subset \mathbb{Z}^d$ (possibly infinite),

$$\mathscr{C}_S \stackrel{\text{def}}{=} \bigcup_{\Lambda \Subset S} \mathscr{C}(\Lambda)$$

is the **algebra of cylinders with base in S**, $\mathscr{F}_S \stackrel{\text{def}}{=} \sigma(\mathscr{C}_S)$. As we did earlier, we let $\mathscr{C} \stackrel{\text{def}}{=} \mathscr{C}_{\mathbb{Z}^d}$, $\mathscr{F} \stackrel{\text{def}}{=} \mathscr{F}_{\mathbb{Z}^d}$, and denote the set of probability measures on $(\Omega, \mathscr{F})$ by $\mathscr{M}_1(\Omega)$.

Rather than consider general specifications, we focus only on *Gibbsian* specifications. Let therefore $\Phi = \{\Phi_B\}_{B \Subset \mathbb{Z}^d}$ be a collection of maps $\Phi_B : \Omega \to \mathbb{R}$, where each Φ_B is $\mathscr{F}_B$-measurable, and absolutely summable, as in (6.25). We present two important examples.

- **The** $\mathsf{O}(N)$ **model.** This model has single-spin space $\Omega_0 = \{x \in \mathbb{R}^N : \|x\|_2 = 1\}$ and its potential can be taken as

$$\Phi_B(\omega) = \begin{cases} -\beta\omega_i \cdot \omega_j & \text{if } B = \{i, j\},\ i \sim j\,, \\ 0 & \text{otherwise,} \end{cases} \tag{6.109}$$

where $x \cdot y$ denotes the scalar product of $x, y \in \mathbb{R}^N$. The case $N = 2$ corresponds to the ***XY* model,** $N = 3$ corresponds to the **Heisenberg model.** These models, and their generalizations, will be discussed in Chapters 9 and 10.

- **The Gaussian Free Field.** Here, as already mentioned, $\Omega_0 = \mathbb{R}$ and

$$\Phi_B(\omega) = \begin{cases} \beta(\omega_i - \omega_j)^2 & \text{if } B = \{i, j\}, i \sim j, \\ \lambda\omega_i^2 & \text{if } B = \{i\}, \\ 0 & \text{otherwise.} \end{cases}$$

This model will be discussed in Chapter 8, and some generalizations in Chapter 9.

Let $\mathscr{H}_{\Lambda;\Phi}$ denote the Hamiltonian associated to Φ, in a region $\Lambda \Subset \mathbb{Z}^d$, defined as in (6.24). For spins taking values in $\{\pm 1\}$, a Gibbsian specification associated to a Hamiltonian was defined pointwise in (6.30) through the numbers $\pi_\Lambda^\Phi(\tau_\Lambda \,|\, \omega)$. Due to the a priori continuous nature of Ω_0 (as in the case $\Omega_0 = \mathbb{S}^1$), the finite-volume Gibbs distribution must be defined differently, since even configurations in a finite volume will usually have zero probability.

Assume therefore we are given a measure λ_0 on $(\Omega_0, \mathscr{B}_0)$, called the **reference measure**. λ_0 need not necessarily be a probability measure. In the case of $\mathbb{S}^1$ and $\mathbb{R}$, the most natural choice for λ_0 is the Lebesgue measure [2]. The product measure on $(\Omega_\Lambda, \mathscr{B}_\Lambda)$, usually denoted $\bigotimes_{i\in\Lambda} \lambda_0$ but which we will here abbreviate by λ_0^Λ, is defined by setting, for all rectangles $\times_{i\in\Lambda} A_i$,

$$\lambda_0^\Lambda(\times_{i\in\Lambda} A_i) \stackrel{\text{def}}{=} \prod_{i\in\Lambda} \lambda_0(A_i)\,.$$

We then define the **Gibbsian specification** $\pi^\Phi = \{\pi_\Lambda^\Phi\}_{\Lambda \Subset \mathbb{Z}^d}$ by setting, for all $A \in \mathscr{F}$ and all boundary conditions $\eta \in \Omega$ (compare with (6.30)),

$$\pi_\Lambda^\Phi(A \,|\, \eta) \stackrel{\text{def}}{=} \frac{1}{\mathbf{Z}_\Lambda^\eta} \int_{\Omega_\Lambda} \mathbf{1}_A(\omega_\Lambda \eta_{\Lambda^c})\, e^{-\mathscr{H}_{\Lambda;\Phi}(\omega_\Lambda \eta_{\Lambda^c})}\, \lambda_0^\Lambda(\mathrm{d}\omega_\Lambda)\,, \tag{6.110}$$

where

$$\mathbf{Z}_{\Lambda;\Phi}^\eta \stackrel{\text{def}}{=} \int_{\Omega_\Lambda} e^{-\mathscr{H}_{\Lambda;\Phi}(\omega_\Lambda \eta_{\Lambda^c})}\, \lambda_0^\Lambda(\mathrm{d}\omega_\Lambda)\,.$$

For convenience, we will sometimes use the following abbreviation:

$$\pi_\Lambda^\Phi(\mathrm{d}\omega \,|\, \eta) = \frac{e^{-\mathscr{H}_{\Lambda;\Phi}(\omega)}}{\mathbf{Z}_{\Lambda;\Phi}^\eta} \lambda_0^\Lambda \otimes \delta_\eta(\mathrm{d}\omega)\,, \tag{6.111}$$

[2] In the case $\Omega_0 = \{-1, 1\}$, the (implicitly used) reference measure λ_0 was simply the counting measure $\lambda_0 = \delta_{-1} + \delta_1$.

in which δ_η denotes the Dirac mass on Ω_{Λ^c} concentrated at η_{Λ^c}. The expectation of a function $f:\Omega\to\mathbb{R}$ with respect to $\pi^\Phi_\Lambda(\cdot\,|\,\eta)$ thus becomes, after integrating out over δ_η,

$$\pi^\Phi_\Lambda f(\eta) = \int_{\Omega_\Lambda} \frac{e^{-\mathscr{H}_{\Lambda;\Phi}(\omega_\Lambda\eta_{\Lambda^c})}}{\mathbf{Z}^\eta_{\Lambda;\Phi}} f(\omega_\Lambda\eta_{\Lambda^c})\lambda_0^\Lambda(\mathrm{d}\omega_\Lambda)\,.$$

All the notions introduced earlier, in particular the notion of consistency of kernels, of Gibbsian specification π^Φ and of the set of probability measures compatible with a specification π^Φ, $\mathscr{G}(\Phi)$, extend immediately to this more general setting.

The topological notions related to Ω have immediate generalizations. Let $\omega\in\Omega$. A sequence $(\omega^{(n)})_{n\geq 1}\subset\Omega$ **converges to** ω if, for all $j\in\mathbb{Z}^d$,

$$d(\omega^{(n)}_j,\omega_j)\to 0 \quad \text{as } n\to\infty\,.$$

In the present general context, a sequence $\mu_n\in\mathscr{M}_1(\Omega)$ is said to **converge to** $\mu\in\mathscr{M}_1(\Omega)$, which we write $\mu_n\Rightarrow\mu$, if [2]

$$\mu_n(f)\to\mu(f) \quad \text{for all bounded local functions } f\,.$$

6.10.2 DLR formalism for compact spin space

When Ω_0 is compact, most of the important results of this chapter have immediate analogues. The starting point is that (as in the case $\Omega_0=\{\pm1\}$, see Proposition 6.20) the notions of convergence for configurations and measures make Ω and $\mathscr{M}_1(\Omega)$ sequentially compact. Proceeding exactly as in the proof of Theorem 6.26, the compactness of $\mathscr{M}_1(\Omega)$ and the Feller property allow to show that there exists at least one Gibbs measure compatible with π^Φ: $\mathscr{G}(\Phi)\neq\varnothing$. Although some proofs need to be slightly adapted, all the main results presented on the structure of $\mathscr{G}(\Phi)$ when $\Omega_0=\{\pm1\}$ remain true when Ω_0 is a compact metric space. In particular, $\mathscr{G}(\Phi)$ is convex and its extremal elements enjoy the same properties as before:

Theorem 6.85. *(Compact spin space) Let Φ be an absolutely summable potential. Let $\mu\in\mathscr{G}(\Phi)$. The following conditions are equivalent characterizations of extremality.*

1. *μ is extremal.*
2. *μ is **trivial on** $\mathscr{T}_\infty$: if $A\in\mathscr{T}_\infty$, then $\mu(A)$ is either 1 or 0.*
3. *All $\mathscr{T}_\infty$-measurable functions are μ-almost surely constant.*
4. *μ has **short-range correlations**: for all $A\in\mathscr{F}$ (or, equivalently, for all $A\in\mathscr{C}$),*

$$\lim_{\Lambda\uparrow\mathbb{Z}^d}\sup_{B\in\mathscr{F}_{\Lambda^c}}\bigl|\mu(A\cap B)-\mu(A)\mu(B)\bigr| = 0\,. \qquad (6.112)$$

Extremal elements can also be constructed using limits:

Theorem 6.86. *(Compact spin space) Let $\mu\in\mathrm{ex}\mathscr{G}(\Phi)$. Then, for μ-almost all ω,*

$$\pi^\Phi_{\mathrm{B}(n)}(\cdot\,|\,\omega)\Rightarrow\mu\,.$$

As in the case of finite single-spin space, to each $\mu \in \mathscr{G}(\Phi)$ corresponds a unique probability distribution λ_μ on $\mathscr{M}_1(\Omega)$, concentrated on the extremal measures of $\mathscr{G}(\Phi)$, leading to the following extremal decomposition:

Theorem 6.87. *(Compact spin space) For all $\mu \in \mathscr{G}(\Phi)$,*

$$\forall B \in \mathscr{F}, \quad \mu(B) = \int_{\mathrm{ex}\mathscr{G}(\pi)} \nu(B)\lambda_\mu(\mathrm{d}\nu). \tag{6.113}$$

Moreover, λ_μ is the unique measure on $\mathscr{M}_1(\Omega)$ for which such a representation holds.

Finally, uniqueness results similar to Theorems 6.31 and 6.40 hold.

6.10.3 Symmetries

Let $(\mathsf{G},\cdot)$ be a group acting on Ω_0 (this action can then be extended to Ω in the natural way as explained in Section 6.6). The notion of G-invariant potential is the same as before, but, in order to state the main result about symmetries, we will need to assume that the reference measure is also invariant under G, that is, $\tau_{\mathsf{g}}\lambda_0 = \lambda_0$ for all $\mathsf{g} \in \mathsf{G}$.

To illustrate this, let us return to the two examples introduced above.

- For the $\mathsf{O}(N)$ models, the potential is invariant under the action of the orthogonal group $\mathsf{O}(N)$, which acts on Ω_0 via its representation as the set of all $N \times N$ orthogonal matrices. Observe that the reference measure is then $\mathsf{O}(N)$-invariant since the determinant of each such matrix is ± 1.
- For the Gaussian Free Field, the potential is invariant under the action of the group $(\mathbb{R},+)$ (that is, under the addition of the same arbitrary real number to all spins of the configuration). Since the reference measure is the Lebesgue measure, its invariance is clear.

One then gets:

Theorem 6.88. *Let G be an internal transformation group under which the reference measure is invariant. Let π be a G-invariant specification. Then, $\mathscr{G}(\pi)$ is preserved by G:*

$$\mu \in \mathscr{G}(\pi) \Rightarrow \tau_{\mathsf{g}}\mu \in \mathscr{G}(\pi) \quad \forall \mathsf{g} \in \mathsf{G}.$$

6.11 A criterion for non-uniqueness

The results on uniqueness and on the extremal decomposition mentioned earlier hold for of a very general class of specifications. Unfortunately, it is much harder to establish *non-uniqueness* in a general setting and one usually has to resort to more model-specific methods. Two such approaches will be presented in Chapters 7 (the Pirogov–Sinai theory) and 10 (reflection positivity).

In the present section, we derive a criterion relating non-uniqueness to the non-differentiability of a suitably-defined pressure [10]. This provides a vast generalization of the corresponding discussion in Section 3.2.2. This criterion will be used in Chapter 10 to establish non-uniqueness in models with continuous spin. For simplicity, we will restrict our attention to translation-invariant potentials of finite range and assume that Ω_0 is compact.

Let Φ be a translation invariant and finite-range potential, and g be any local function. For each $\Lambda \Subset \mathbb{Z}^d$, let $\Lambda(g) \stackrel{\text{def}}{=} \{i \in \mathbb{Z}^d : \text{supp}(g \circ \theta_i) \cap \Lambda \neq \varnothing\}$. Then define, for each $\omega \in \Omega$,

$$\psi_\Lambda^\omega(\lambda) \stackrel{\text{def}}{=} \frac{1}{|\Lambda(g)|} \log \Big\langle \exp\Big\{\lambda \sum_{j \in \Lambda(g)} g \circ \theta_j \Big\}\Big\rangle_{\Lambda;\Phi}^\omega, \tag{6.114}$$

where $\langle \cdot \rangle_{\Lambda;\Phi}^\omega$ denotes expectation with respect to the kernel $\pi_\Lambda^\Phi(\cdot \,|\, \omega)$.

Lemma 6.89. *For any sequence $\Lambda_n \Uparrow \mathbb{Z}^d$ and any sequence of boundary conditions $(\omega_n)_{n\geq 1}$, the limit*

$$\psi(\lambda) \stackrel{\text{def}}{=} \lim_{n\to\infty} \psi_{\Lambda_n}^{\omega_n}(\lambda) \tag{6.115}$$

exists and is independent of the choice of $(\Lambda_n)_{n\geq 1}$ and $(\omega_n)_{n\geq 1}$. Moreover, $\lambda \mapsto \psi(\lambda)$ is convex.

Although it could be adapted to the present situation, the existence of the pressure proved in Theorem 6.79 does not apply since, here, we do not assume that Ω_0 contains finitely many elements.

Proof. Notice that $|\Lambda_n(g)|/|\Lambda_n| \to 0$. Using the fact that Φ has finite range and that g is local, one can repeat the same steps as in the proof of Theorem 3.6. Using the Hölder Inequality as we did in the proof of Lemma 3.5, we deduce that $\lambda \mapsto \psi_\Lambda^\omega(\lambda)$ is convex. □

Exercise 6.31. *Complete the details of the proof of Lemma 6.89.*

Remark 6.90. In (6.114), the expectation with respect to $\pi_\Lambda^\Phi(\cdot \,|\, \omega)$ can be substituted by the expectation with respect to any Gibbs measure $\mu \in \mathscr{G}(\Phi)$. Namely, let

$$\psi_\Lambda^\mu(\lambda) \stackrel{\text{def}}{=} \frac{1}{|\Lambda(g)|} \log \Big\langle \exp\Big\{\lambda \sum_{j \in \Lambda(g)} g \circ \theta_j \Big\}\Big\rangle_\mu . \tag{6.116}$$

Observe that, since $\langle f \rangle_\mu = \langle \langle f \rangle_{\Lambda;\Phi}^{\cdot} \rangle_\mu$, there exist ω', ω'' (depending on Λ, Φ, etc.) such that

$$\Big\langle \exp\Big\{\lambda \sum_{j \in \Lambda(g)} g \circ \theta_j \Big\}\Big\rangle_{\Lambda;\Phi}^{\omega'} \leq \Big\langle \exp\Big\{\lambda \sum_{j \in \Lambda(g)} g \circ \theta_j \Big\}\Big\rangle_\mu \leq \Big\langle \exp\Big\{\lambda \sum_{j \in \Lambda(g)} g \circ \theta_j \Big\}\Big\rangle_{\Lambda;\Phi}^{\omega''} .$$

(The existence of ω', ω'' follows from the fact that, for any local function f, the function $\omega \mapsto \langle f \rangle_{\Lambda;\Phi}^\omega$ is continuous and bounded and therefore attains its bounds.) Using Lemma 6.89, it follows that $\psi(\lambda) = \lim_{n\to\infty} \psi_{\Lambda_n}^\mu(\lambda)$. ◇

Remember that convexity guarantees that ψ possesses one-sided derivatives with respect to λ. As we did for the Ising model in Theorem 3.34 and Proposition 3.29, we can relate these derivatives to the expectation of g.

Proposition 6.91. *For all $\mu \in \mathscr{G}_\theta(\Phi)$,*

$$\frac{\partial \psi}{\partial \lambda^-}\Big|_{\lambda=0} \leq \langle g \rangle_\mu \leq \frac{\partial \psi}{\partial \lambda^+}\Big|_{\lambda=0} . \tag{6.117}$$

Moreover, there exist $\mu^+, \mu^- \in \mathscr{G}_\theta(\Phi)$, such that

$$\frac{\partial \psi}{\partial \lambda^+}\Big|_{\lambda=0} = \langle g \rangle_{\mu^+} , \qquad \frac{\partial \psi}{\partial \lambda^-}\Big|_{\lambda=0} = \langle g \rangle_{\mu^-} . \tag{6.118}$$

In practice, this result is used to obtain non-uniqueness (namely, the existence of distinct measures μ^+, μ^-) by finding a local function g for which ψ is not differentiable at $\lambda = 0$. In the Ising model, that function was $g = \sigma_0$.

Proof. We first use the fact that $\psi = \lim_{n\to\infty} \psi^{\mu}_{\Lambda_n}(\lambda)$ (see Remark 6.90). Since ψ^{μ}_{Λ} is convex, it follows from (B.9) that, for all $\lambda > 0$,

$$\frac{\psi^{\mu}_{\Lambda}(\lambda) - \psi^{\mu}_{\Lambda}(0)}{\lambda} \geq \frac{\partial \psi^{\mu}_{\Lambda}}{\partial \lambda^+}\Big|_{\lambda=0} = \frac{1}{|\Lambda(g)|}\Big\langle \sum_{j\in\Lambda(g)} g\circ\theta_j \Big\rangle_{\mu} = \langle g\rangle_{\mu},$$

where the last identity is a consequence of the translation invariance of μ. Taking $\Lambda \Uparrow \mathbb{Z}^d$ followed by $\lambda \downarrow 0$, we get the upper bound in (6.117). The lower bound is obtained similarly.

Let us turn to the second claim. We now use the fact that $\psi = \lim_{n\to\infty} \psi^{\omega}_{\Lambda_n}$, for any ω. We *fix* $\lambda > 0$, and use again convexity: for all small $\epsilon > 0$,

$$\begin{aligned}\frac{\psi^{\omega}_{\Lambda}(\lambda) - \psi^{\omega}_{\Lambda}(\lambda-\epsilon)}{\epsilon} \leq \frac{\partial \psi^{\omega}_{\Lambda}}{\partial \lambda^-}\Big|_{\lambda} &= \frac{1}{|\Lambda(g)|}\frac{\Big\langle \big(\sum_{i\in\Lambda(g)} g\circ\theta_i\big) e^{\lambda\sum_{i\in\Lambda(g)} g\circ\theta_i}\Big\rangle^{\omega}_{\Lambda;\Phi}}{\big\langle e^{\lambda\sum_{i\in\Lambda(g)} g\circ\theta_i}\big\rangle^{\omega}_{\Lambda;\Phi}} \\ &= \frac{1}{|\Lambda(g)|}\Big\langle \sum_{i\in\Lambda(g)} (g\circ\theta_i)\Big\rangle^{\omega}_{\Lambda;\Phi^{\lambda}},\end{aligned} \tag{6.119}$$

where the (translation-invariant) potential $\Phi^{\lambda} = \{\Phi^{\lambda}_B\}_{B\Subset\mathbb{Z}^d}$ is defined by

$$\Phi^{\lambda}_B \overset{\text{def}}{=} \begin{cases} \Phi_B + \lambda g\circ\theta_i & \text{if } B = \theta_i(\operatorname{supp} g), \\ \Phi_B & \text{otherwise.}\end{cases}$$

Let now $\mu^{\lambda} \in \mathscr{G}(\Phi^{\lambda})$ be translation invariant (this is always possible by adapting the construction of Exercise 6.22). Integrating both sides of (6.119) with respect to μ^{λ},

$$\Big\langle \frac{\psi^{\cdot}_{\Lambda}(\lambda) - \psi^{\cdot}_{\Lambda}(\lambda-\epsilon)}{\epsilon}\Big\rangle_{\mu^{\lambda}} \leq \frac{1}{|\Lambda(g)|}\Big\langle \sum_{i\in\Lambda(g)} g\circ\theta_i\Big\rangle_{\mu^{\lambda}} = \langle g\rangle_{\mu^{\lambda}}.$$

Notice that $\|\psi^{\cdot}_{\Lambda}(\lambda)\|_{\infty} \leq |\lambda|\|g\|_{\infty} < \infty$. Taking $\Lambda \Uparrow \mathbb{Z}^d$, followed by $\epsilon \downarrow 0$, we get

$$\langle g\rangle_{\mu^{\lambda}} \geq \frac{\partial\psi}{\partial\lambda^-}\Big|_{\lambda} \geq \frac{\partial\psi}{\partial\lambda^+}\Big|_{\lambda=0}, \qquad \forall \lambda > 0.$$

In the last step, we used item 3 of Theorem B.12. Consider now any sequence $(\lambda_k)_{k\geq 1}$ decreasing to 0. By compactness (Theorem 6.24 applies here too), there exists a subsequence $(\lambda_{k_m})_{m\geq 1}$ and a probability measure μ^+ such that $\mu^{\lambda_{k_m}} \Rightarrow \mu^+$ as $m \to \infty$. Clearly, μ^+ is also translation invariant and, by Exercise 6.32 below, $\mu^+ \in \mathscr{G}(\Phi)$. Since g is local, $\langle g\rangle_{\mu^{\lambda_k}} \to \langle g\rangle_{\mu^+}$. Applying (6.117) to μ^+, we conclude that $\langle g\rangle_{\mu^+} = \frac{\partial\psi}{\partial\lambda^+}\big|_{\lambda=0}$. □

Remark 6.92. It can be shown that the measures μ^+, μ^- in the second claim are in fact ergodic with respect to lattice translations. ⋄

Exercise 6.32. *Let $(\Phi^k)_{k\geq 1}$, Φ be translation-invariant potentials of range at most r and such that $\|\Phi^k_B - \Phi_B\|_{\infty} \to 0$ when $k \to \infty$, for all $B \Subset \mathbb{Z}^d$. Let $\mu^k \in \mathscr{G}(\Phi^k)$ and μ be a probability measure such that $\mu^k \Rightarrow \mu$. Then $\mu \in \mathscr{G}(\Phi)$.* Hint: *use a trick like* (6.35).

Exercise 6.33. *Consider the Ising model on $\mathbb{Z}^d$. Prove that $\beta \mapsto \psi^{\text{Ising}}(\beta, h)$ is differentiable whenever $|\mathscr{G}(\beta, h)| = 1$.* Hint: *To prove differentiability at β_0, combine Theorem 3.25 and Proposition 6.91 with $\lambda = \beta - \beta_0$ and $g = \frac{1}{2d}\sum_{i\sim 0}\sigma_0\sigma_i$.*

Remark 6.93. As a matter of fact, it can be proved [11] that the pressure of the Ising model on $\mathbb{Z}^d$ is differentiable with respect to β for *any* values of $\beta \geq 0$ and $h \in \mathbb{R}$, not only in the uniqueness regime. ⋄

Exercise 6.34. *Consider a one-dimensional model with a finite-range potential Φ, and let $\psi(\lambda)$ denote the pressure defined in* (6.115)*, with $g \stackrel{\text{def}}{=} \sigma_0$. Use Proposition 6.91, combined with Theorem 6.40, to show that ψ is differentiable at $\lambda = 0$.*

Remark 6.94. It can be shown that in one-dimension, the pressure of a model with finite-range interactions is always *real-analytic* in its parameters [12]. ⋄

6.12 Some proofs

6.12.1 Proofs related to the construction of probability measures

The existence results of this chapter rely on the sequential compactness of Ω. This implies in particular the following property, actually equivalent to compactness:

Lemma 6.95. *Let $(C_n)_{n\geq 1} \subset \mathscr{C}$ be a decreasing ($C_{n+1} \subset C_n$) sequence of cylinders such that $\bigcap_n C_n = \varnothing$. Then $C_n = \varnothing$ for all large enough n.*

Proof. Let each cylinder C_n be of the form $C_n = \Pi^{-1}_{\Lambda(n)}(A_n)$, where $\Lambda(n) \Subset \mathbb{Z}^d$ and $A_n \in \mathscr{P}(\Omega_{\Lambda(n)})$. With no loss of generality, we can assume that $\Lambda(n) \subset \Lambda(n+1)$ (remember the hint of Exercise 6.2). Assume that $C_n \neq \varnothing$ for all n, and let $\omega^{(n)} \in C_n$. Since $C_m \subset C_n$ for all $m > n$,

$$\Pi_{\Lambda(n)}(\omega^{(m)}) \in A_n \quad \text{for all } m > n\,.$$

By compactness of Ω, there exists a configuration ω^* and a subsequence $(\omega^{(n_k)})_{k\geq 1}$ such that $\omega^{(n_k)} \to \omega^*$. Of course,

$$\Pi_{\Lambda(n)}(\omega^*) \in A_n \quad \text{for all } n\,,$$

which implies $\omega^* \in C_n$ for all n. Therefore, $\bigcap_n C_n \neq \varnothing$. □

Theorem 6.96. *A finitely additive set function $\mu : \mathscr{C} \to \mathbb{R}_{\geq 0}$ with $\mu(\Omega) < \infty$ always has a unique extension to $\mathscr{F}$.*

Proof. To use Carathéodory's Extension Theorem B.33, we must verify that if μ is finitely additive on $\mathscr{C}$, then it is also σ-additive on $\mathscr{C}$, in the sense that if $(C_n)_{n\geq 1} \subset \mathscr{C}$ is a sequence of pairwise disjoint cylinders such that $C = \bigcup_{n\geq 1} C_n \in \mathscr{C}$, then $\mu(C) = \sum_{n\geq 1}\mu(C_n)$. For this, it suffices to write $\bigcup_{n\geq 1} C_n = A_N \cup B_N$, where $A_N = \bigcup_{n=1}^N C_n \in \mathscr{C}$, $B_N = \bigcup_{n>N} C_n \in \mathscr{C}$. Notice that, as $N \to \infty$, $A_N \uparrow C$, and $B_N \downarrow \varnothing$. Since $B_{N+1} \subset B_N$, Lemma 6.95 implies that there exists some N_0 such that $B_N = \varnothing$ when $N > N_0$. This also implies that $C_n = \varnothing$ for all $n > N_0$, and so $\mu(C) = \mu(A_{N_0}) = \sum_{n=1}^{N_0}\mu(C_n) = \sum_{n\geq 1}\mu(C_n)$. □

6.12.2 Proof of Theorem 6.5

We will prove Theorem 6.5 using the following classical result:

Theorem 6.97 (Riesz–Markov–Kakutani Representation Theorem on $\Omega = \{\pm 1\}^{\mathbb{Z}^d}$**).** *Let $L : C(\Omega) \to \mathbb{R}$ be a positive normalized linear functional, that is:*

1. *If $f \geq 0$, then $L(f) \geq 0$.*
2. *For all $f, g \in C(\Omega)$, $\alpha, \beta \in \mathbb{R}$, $L(\alpha f + \beta g) = \alpha L(f) + \beta L(g)$.*
3. *$L(1) = 1$.*

Then, there exists a unique measure $\mu \in \mathscr{M}_1(\Omega)$ such that

$$L(f) = \int f \, \mathrm{d}\mu, \quad \text{for all } f \in C(\Omega).$$

This result holds in a much broader setting; its proof can be found in many textbooks. For the sake of concreteness, we give an elementary proof that makes use of the simple structure of $\Omega = \{-1, 1\}^{\mathbb{Z}^d}$.

Proof of Theorem 6.97: We use some of the notions developed in Section 6.4. Since $-\|f\|_\infty \leq f \leq \|f\|_\infty$, linearity and positivity of L yield $|L(f)| \leq \|f\|_\infty$. We have already seen that, for each cylinder $C \in \mathscr{C}$, $\mathbf{1}_C \in C(\Omega)$. Let then

$$\mu(C) \stackrel{\text{def}}{=} L(\mathbf{1}_C).$$

Observe that $0 \leq \mu(C) \leq 1$, and that if $C_1, C_2 \in \mathscr{C}$ are disjoint, then $\mu(C_1 \cup C_2) = \mu(C_1) + \mu(C_2)$. By Theorem 6.96, μ extends uniquely to a measure on $(\Omega, \mathscr{F})$. To show that $\mu(f) = L(f)$ for all $f \in C(\Omega)$, let, for each n, f_n be a finite linear combination of the form $\sum_i a_i \mathbf{1}_{C_i}$, $C_i \in \mathscr{C}$, such that $\|f_n - f\|_\infty \to 0$. Then $\mu(f_n) = L(f_n)$ for all n, and therefore

$$|\mu(f) - L(f)| \leq |\mu(f) - \mu(f_n)| + |L(f_n) - L(f)| \leq 2\|f_n - f\|_\infty \to 0$$

as $n \to \infty$. □

We can now prove Theorem 6.5. Since a state $\langle \cdot \rangle$ is defined only on local functions, we must first extend it to continuous functions. Let $f \in C(\Omega)$ and let $(f_n)_{n \geq 1}$ be a sequence of local functions converging to f: $\|f_n - f\|_\infty \to 0$. Define $\langle f \rangle \stackrel{\text{def}}{=} \lim_n \langle f_n \rangle$. This definition does not depend on the choice of the sequence $(f_n)_{n \geq 1}$. Namely, if $(g_n)_{n \geq 1}$ is another such sequence, then $|\langle f_n \rangle - \langle g_n \rangle| \leq \|f_n - g_n\|_\infty \leq \|f_n - f\|_\infty + \|g_n - f\|_\infty \to 0$. The linear map $\langle \cdot \rangle : C(\Omega) \to \mathbb{R}$ then satisfies all the hypotheses of Theorem 6.97, which proves the result. □

6.12.3 Proof of Theorem 6.6

We first define a probability measure on $\mathscr{C}$ and then extend it to $\mathscr{F}$ using Carathéodory's Extension Theorem (Theorem B.33). Consider a cylinder $C \in \mathscr{C}(\Lambda)$. Then C can be written in the form $C = \Pi_\Lambda^{-1}(A)$ where $A \in \mathscr{P}(\Omega_\Lambda)$. Let then

$$\mu(C) \stackrel{\text{def}}{=} \mu_\Lambda(A).$$

The consistency condition (6.4) guarantees that this number is well defined. Namely, if C can also be written as $C = \pi_{\Lambda'}^{-1}(A')$, where $A' \in \mathscr{P}(\Omega_{\Lambda'})$, we must show that $\mu_\Lambda(A) = \mu_{\Lambda'}(A')$. But (remember Exercise 6.2), if Δ is large enough to contain both Λ and Λ', one can write $C = \pi_\Delta^{-1}(B)$, for some $B \in \mathscr{P}(\Omega_\Delta)$. But then $A = \Pi_\Lambda(\Pi_\Delta^{-1}(B))$, and so

$$\mu_\Lambda(A) = \mu_\Lambda\big(\Pi_\Lambda(\Pi_\Delta^{-1}(B))\big) = \mu_\Lambda(\Pi_\Lambda^\Delta(A)) = \mu_\Delta(B)\,.$$

The same with A' gives $\mu_\Lambda(A) = \mu_{\Lambda'}(A') = \mu_\Delta(B)$.

One then verifies that μ, defined as above, defines a probability measure on cylinders. For instance, if $C_1, C_2 \in \mathscr{C}$ are disjoint, then one can find some $\Delta \Subset \mathbb{Z}^d$ such that $C_1 = \Pi_\Delta^{-1}(A_1)$, $C_2 = \Pi_\Delta^{-1}(A_2)$, where $A_1, A_2 \in \mathscr{P}(\Omega_\Delta)$ are also disjoint. Then,

$$\begin{aligned}\mu(C_1 \cup C_2) = \mu(\Pi_\Delta^{-1}(A_1 \cup A_2)) &= \mu_\Delta(A_1 \cup A_2)\\ &= \mu_\Delta(A_1) + \mu_\Delta(A_2)\\ &= \mu(\Pi_\Delta^{-1}(A_1)) + \mu(\Pi_\Delta^{-1}(A_2)) = \mu(C_1) + \mu(C_2)\,.\end{aligned}$$

By Theorem 6.96, μ extends uniquely to a probability measure on $\mathscr{F}$, and (6.4) holds by construction.

Remark 6.98. Kolmogorov's Extension Theorem holds in more general settings, in particular for much more general single-spin spaces. ⋄

6.12.4 Proof of Theorem 6.24

Let $\{C_1, C_2, \dots\}$ be an enumeration of all the cylinders of $\mathscr{C}$ (Exercise 6.2). First, we can extract from the sequence $(\mu_n(C_1))_{n\geq 1} \subset [0,1]$ a convergent subsequence $(\mu_{n_{1,j}}(C_1))_{j\geq 1}$ such that

$$\mu(C_1) \stackrel{\text{def}}{=} \lim_{j\to\infty} \mu_{n_{1,j}}(C_1) \quad \text{exists.}$$

Then, we extract from $(\mu_{n_{1,j}}(C_2))_{j\geq 1} \in [0,1]$ a convergent subsequence $(\mu_{n_{2,j}}(C_2))_{j\geq 1}$ such that

$$\mu(C_2) \stackrel{\text{def}}{=} \lim_{j\to\infty} \mu_{n_{2,j}}(C_2) \quad \text{exists.}$$

This process continues until we have, for each $k \geq 1$, a subsequence $(n_{k,j})_{j\geq 1}$ such that

$$\mu(C_k) \stackrel{\text{def}}{=} \lim_{j\to\infty} \mu_{n_{k,j}}(C_k)\,.$$

By considering the diagonal sequence $(n_{j,j})_{j\geq 1}$, we have that $\mu_{n_{j,j}}(C) \to \mu(C)$ for all $C \in \mathscr{C}$. Proceeding as in the proof of Theorem 6.97, using Lemma 6.95, we can verify that μ is a probability measure on $\mathscr{C}$ and use again Theorem 6.96 to extend it to a measure μ on $\mathscr{F}$. Obviously, $\mu_{n_{j,j}} \Rightarrow \mu$.

6.12.5 Proof of Proposition 6.39

To lighten the notations, we will omit $\beta\Phi$ most of the time. Let S_* denote the support of f. Assume that Λ is sufficiently large to contain S_*, and let $\Lambda' \stackrel{\text{def}}{=} \Lambda \setminus S_*$.

Writing $\eta_\Lambda = \eta_{S_*}\eta_{\Lambda'}$, we have, by definition,

$$\begin{aligned}\pi_\Lambda f(\omega) &= \frac{1}{\mathbf{Z}_\Lambda^\omega}\sum_{\eta_\Lambda} f(\eta_{S_*})e^{-\beta\mathscr{H}_\Lambda(\eta_\Lambda\omega_{\Lambda^c})}\\ &= \frac{1}{\mathbf{Z}_\Lambda^\omega}\sum_{\eta_{S_*}} F_\Lambda^\omega(\eta_{S_*})\sum_{\eta_{\Lambda'}} e^{-\beta\mathscr{H}_{\Lambda'}(\eta_{\Lambda'}\eta_{S_*}\omega_{\Lambda^c})}\\ &= \sum_{\eta_{S_*}} F_\Lambda^\omega(\eta_{S_*})\frac{\mathbf{Z}_{\Lambda'}^{\omega'}}{\mathbf{Z}_\Lambda^\omega}\,,\end{aligned} \tag{6.120}$$

where we have abbreviated $\eta_{S_*}\omega_{\Lambda^c}$ by ω', and defined

$$F_\Lambda^\omega(\eta_{S_*}) \stackrel{\text{def}}{=} f(\eta_{S_*})\exp\Big\{-\beta\sum_{\substack{B\cap S_*\neq\varnothing\\ B\cap\Lambda'=\varnothing}}\Phi_B(\eta_{S_*}\omega_{\Lambda^c})\Big\}.$$

First, observe that

$$\lim_{\Lambda\Uparrow\mathbb{Z}^d} F_\Lambda^\omega(\eta_{S_*}) = f(\eta_{S_*})\exp\Big\{-\beta\sum_{B\subset S_*}\Phi_B(\eta_{S_*})\Big\}, \tag{6.121}$$

the latter expression being independent of Λ and ω. Indeed, by the absolute summability of the potential Φ,

$$\forall i,\ \lim_{r\to\infty}\sum_{\substack{B\ni i\\ \operatorname{diam}(B)>r}}\|\Phi_B\|_\infty = 0 \qquad \text{so that} \qquad \lim_{\Lambda\Uparrow\mathbb{Z}^d}\sum_{\substack{B\cap S_*\neq\varnothing\\ B\cap\Lambda^c\neq\varnothing}}\|\Phi_B\|_\infty = 0\,.$$

$F_\Lambda^\omega(\eta_*)$ thus becomes independent of ω in the limit $\Lambda\Uparrow\mathbb{Z}^d$. We will now prove that the same is true of the ratio appearing in (6.120). In order to do this, we will show that, when β is small, the ratio can be controlled using convergent cluster expansions, leading to crucial cancellations. We discuss explicitly only the case of $\mathbf{Z}_\Lambda^\omega$, the analysis being the same for $\mathbf{Z}_{\Lambda'}^{\omega'}$.

An application of the "$+1-1$ trick" (see Exercise 3.22) yields

$$e^{-\beta\mathscr{H}_\Lambda} = \prod_{B\cap\Lambda\neq\varnothing} e^{-\beta\Phi_B} = \sum_{\mathscr{B}}\prod_{B\in\mathscr{B}}\big(e^{-\beta\Phi_B}-1\big),$$

where the sum is over all finite collections $\mathscr{B}$ of finite sets B such that $B\cap\Lambda\neq\varnothing$. Of course, we can assume that the only sets B used are those for which $\Phi_B\not\equiv 0$ (this will be done implicitly from now on). We associate to each collection $\mathscr{B}$ a graph, as follows. To each $B\in\mathscr{B}$ is associated an abstract vertex x. We add an edge between two vertices x, x' if and only if they are associated to sets B, B' for which $B\cap B'\neq\varnothing$. The resulting graph is then decomposed into maximal connected components. To each such component, say with vertices $\{x_1,\dots,x_k\}$, corresponds a collection $\gamma = \{B_1,\dots,B_k\}$, called a **polymer**. The **support of γ** is defined by $\overline{\gamma} \stackrel{\text{def}}{=} B_1\cup\dots\cup B_k$. In $\mathbf{Z}_\Lambda^\omega$, one can interchange the summations over $\omega_\Lambda\in\Omega_\Lambda$ and $\mathscr{B}$ and obtain

$$\mathbf{Z}_\Lambda^\omega = 2^{|\Lambda|}\sum_\Gamma\prod_{\gamma\in\Gamma}\mathsf{w}(\gamma)\,,$$

where the sum is over families Γ such that $\overline{\gamma}\cap\overline{\gamma}'=\varnothing$ whenever $\gamma=\{B_1,\dots,B_k\}$ and $\gamma'=\{B_1',\dots,B_{k'}'\}$ are two distinct collections in Γ. The **weight** of γ is defined by

$$\mathsf{w}(\gamma) \stackrel{\text{def}}{=} 2^{-|\overline{\gamma}\cap\Lambda|}\sum_{\eta_{\overline{\gamma}\cap\Lambda}}\prod_{B\in\gamma}\big(e^{-\beta\Phi_B(\eta_{\overline{\gamma}\cap\Lambda}\omega_{\Lambda^c})}-1\big).$$

To avoid too heavy notations, we have not indicated the possible dependence of these weights on ω and Λ. Observe that, when $\overline{\gamma} \subset \Lambda$, $\mathsf{w}(\gamma)$ does not depend on ω. The following bound always holds:

$$|\mathsf{w}(\gamma)| \le \prod_{B\in\gamma} \|e^{-\beta\Phi_B} - 1\|_\infty . \tag{6.122}$$

We now show that, when β is small, the polymers and their weights satisfy condition (5.10) that guarantees convergence of the cluster expansion for $\log \mathbf{Z}_\Lambda^\omega$. We will use the function $a(\gamma) \stackrel{\text{def}}{=} |\overline{\gamma}|$.

Lemma 6.99. *Let $\epsilon \ge 0$. Assume that*

$$\alpha \stackrel{\text{def}}{=} \sup_{i\in\mathbb{Z}^d} \sum_{B\ni i} \|e^{-\beta\Phi_B} - 1\|_\infty e^{(3+\epsilon)|B|} \le 1 . \tag{6.123}$$

Then, for all γ_0, uniformly in ω and Λ,

$$\sum_{\gamma:\overline{\gamma}\cap\gamma_0\neq\varnothing} |\mathsf{w}(\gamma)| e^{(1+\epsilon)|\overline{\gamma}|} \le |\overline{\gamma}_0| . \tag{6.124}$$

Exercise 6.35. *Show that, for any $\epsilon \ge 0$, any potential Φ satisfying* (6.50) *also satisfies* (6.123) *once β is small enough.*

Proof of Lemma 6.99: Let $b(\gamma)$ denote the number of sets B_i contained in γ. Let, for all $n \ge 1$,

$$\xi(n) \stackrel{\text{def}}{=} \max_{i\in\Lambda} \sum_{\substack{\gamma:\overline{\gamma}\ni i\\ b(\gamma)\le n}} |\mathsf{w}(\gamma)| e^{(1+\epsilon)|\overline{\gamma}|} . \tag{6.125}$$

We will show that, when (6.123) is satisfied,

$$\xi(n) \le \alpha , \qquad \forall n \ge 1, \tag{6.126}$$

which of course implies (6.124) after letting $n \to \infty$.

Let us first consider the case $n = 1$. In this case, γ contains a single set B and so $|\mathsf{w}(\gamma)| \le \|e^{-\beta\Phi_B} - 1\|_\infty$. This gives

$$\xi(1) \le \max_{i\in\Lambda} \sum_{B\ni i} \|e^{-\beta\Phi_B(\cdot)} - 1\|_\infty e^{(1+\epsilon)|B|} \le \alpha .$$

Let us then assume that (6.126) holds for n, and let us verify that it also holds for $n+1$. Since $\overline{\gamma} \ni i$, each polymer γ appearing in the sum for $n+1$ can be decomposed (not necessarily in a unique manner) as follows: $\gamma = \{B_0\} \cup \gamma^{(1)} \cup \dots \cup \gamma^{(k)}$, where $B_0 \ni i$ and the $\gamma^{(j)}$s are polymers with disjoint support such that $b(\gamma^{(j)}) \le n$ and $\overline{\gamma}^{(j)} \cap B_0 \neq \varnothing$. Since the $\overline{\gamma}^{(j)}$s are disjoint,

$$|\mathsf{w}(\gamma)| \le 2^{|B_0|} \|e^{-\beta\Phi_{B_0}} - 1\|_\infty \prod_{j=1}^k |\mathsf{w}(\gamma^{(j)})| .$$

We have $|\overline{\gamma}| \le |B_0| + \sum_{j=1}^k |\overline{\gamma}^{(j)}|$ and therefore, for a fixed B_0, we can sum over the polymers $\gamma^{(j)}$ and use the induction hypothesis, obtaining a contribution bounded by

$$\sum_{k\geq 0}\frac{1}{k!}\sum_{\substack{\gamma^{(1)}:\\ \overline{\gamma}^{(1)}\cap B_0\neq\varnothing\\ b(\gamma^{(1)})\leq n}}\cdots\sum_{\substack{\gamma^{(k)}:\\ \overline{\gamma}^{(k)}\cap B_0\neq\varnothing\\ b(\gamma^{(k)})\leq n}}\prod_{j=1}^{k}|\mathrm{w}(\gamma^{(j)})|e^{(1+\epsilon)|\overline{\gamma}^{(j)}|}\leq\sum_{k\geq 0}\frac{1}{k!}(|B_0|\xi(n))^k\leq e^{\alpha|B_0|},$$

and we are left with

$$\xi(n+1)\leq\sum_{B_0\ni 0}2^{|B_0|}\|e^{-\beta\Phi_{B_0}}-1\|_\infty e^{(1+\epsilon)|B_0|}e^{\alpha|B_0|}\leq\alpha.$$

In the last inequality, we used $\alpha\leq 1$ and the definition of α. □

Proof of Proposition 6.39: Let $\epsilon>0$ and let β_1 be such that (6.123) holds for all $\beta\leq\beta_1$ (Exercise 6.35). We study the ratio in (6.120) by using convergent cluster expansions for its numerator and denominator. We use the terminology of Section 5.6. We denote by χ_Λ the set of clusters appearing in the expansion of $\log\mathbf{Z}_\Lambda^\omega$; the latter are made of polymers $\gamma=\{B_1,\dots,B_k\}$ for which $B_i\cap\Lambda\neq\varnothing$ for all i. The weight of $X\in\chi_\Lambda$ is denoted $\Psi_{\Lambda,\omega}(X)$ (see (5.20)); it is built using the weights $\mathrm{w}(\gamma)$, which can depend on ω if γ has a support that intersects Λ^{c}. Similarly, we denote by $\chi_{\Lambda'}$ the set of clusters appearing in the expansion of $\log\mathbf{Z}_{\Lambda'}^{\omega'}$; the latter are made of polymers $\gamma=\{B_1,\dots,B_k\}$ for which $B_i\cap\Lambda'\neq\varnothing$. The weight of $X\in\chi_{\Lambda'}$ is denoted $\Psi_{\Lambda',\omega'}(X)$. Let us denote the support of X by $\overline{X}\stackrel{\text{def}}{=}\bigcup_{\gamma\in X}\overline{\gamma}$. Taking $\beta\leq\beta_1$ guarantees in particular that

$$\sum_{\gamma:\overline{\gamma}\cap\gamma_0\neq\varnothing}|\mathrm{w}(\gamma)|e^{|\overline{\gamma}|}\leq|\overline{\gamma}_0|,$$

so we can expand that ratio using an absolutely convergent cluster expansion for each partition function:

$$\frac{\mathbf{Z}_{\Lambda'}^{\omega'}}{\mathbf{Z}_\Lambda^\omega}=2^{-|S_*|}\frac{\exp\Bigl\{\sum_{X\in\chi_{\Lambda'}}\Psi_{\Lambda',\omega'}(X)\Bigr\}}{\exp\Bigl\{\sum_{X\in\chi_\Lambda}\Psi_{\Lambda,\omega}(X)\Bigr\}}=2^{-|S_*|}\frac{\exp\Bigl\{\sum_{\substack{X\in\chi_{\Lambda'}\\ \overline{X}\cap S_*\neq\varnothing}}\Psi_{\Lambda',\omega'}(X)\Bigr\}}{\exp\Bigl\{\sum_{\substack{X\in\chi_\Lambda\\ \overline{X}\cap S_*\neq\varnothing}}\Psi_{\Lambda,\omega}(X)\Bigr\}}.$$

The second identity is due to the fact that each cluster $X\in\chi_{\Lambda'}$ in the numerator whose support does not intersect S_* also appears, with the same weight, in the denominator as a cluster $X\in\chi_\Lambda$. Their contributions thus cancel out. Among the remaining clusters, there are those that intersect Λ^{c}. These yield no contribution in the thermodynamic limit. Indeed, considering the denominator for example,

$$\sum_{\substack{X\in\chi_\Lambda:\\ \overline{X}\cap S_*\neq\varnothing\\ \overline{X}\cap\Lambda^{\mathrm{c}}\neq\varnothing}}\bigl|\Psi_{\Lambda,\omega}(X)\bigr|\leq|S_*|\max_{i\in S_*}\sum_{\substack{X:\overline{X}\ni i\\ \mathrm{diam}(\overline{X})\geq d(S_*,\Lambda^{\mathrm{c}})}}\bigl|\Psi_{\Lambda,\omega}(X)\bigr|,\tag{6.127}$$

and this last sum converges to zero when $\Lambda\uparrow\mathbb{Z}^d$. Indeed, we know (see (5.29)) that

$$\sum_{X:\overline{X}\ni i}\bigl|\Psi_{\Lambda,\omega}(X)\bigr|=\sum_{N\geq 1}\sum_{\substack{X:\overline{X}\ni i\\ \mathrm{diam}(\overline{X})=N}}\bigl|\Psi_{\Lambda,\omega}(X)\bigr|$$

is convergent. Therefore, the second sum above goes to zero when $N\to\infty$, allowing us to conclude that the contribution of the clusters intersecting Λ^{c} vanishes when $\Lambda\uparrow\mathbb{Z}^d$.

Revised version, August 22 2017
To be published by Cambridge University Press (2017)

www.unige.ch/math/folks/velenik/smbook

We are thus left with the clusters X which are strictly contained in Λ and intersect S_*. The weights of these do not depend on ω_{Λ^c} anymore (for that reason, the corresponding subscripts will be removed from their weights), but those which appear in the numerator have weights that still depend on η_{S_*} and their weights will be written, for simplicity, as $\Psi_{\eta_{S_*}}$. We get

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{\mathbf{Z}^{\omega'}_{\Lambda'}}{\mathbf{Z}^{\omega}_{\Lambda}} = \exp\Big\{ \sum_{\substack{X:\overline{X}\not\subset S_* \\ \overline{X}\cap S_*\neq\varnothing}} \Psi_{\eta_{S_*}}(X) - \sum_{X:\overline{X}\cap S_*\neq\varnothing} \Psi(X) \Big\}.$$

Combined with (6.120) and (6.121), this completes the proof of the first claim.

Let us then see what more can be done when Φ has finite range: $r(\Phi) < \infty$. In this case, $F^{\omega}_{\Lambda}(\eta_{S_*})$ becomes *equal* to its limit as soon as Λ is large enough. Moreover, each cluster $X = \{\gamma_1, \dots, \gamma_n\}$ in the second sum of the right-hand side of (6.127) satisfies $\sum_{i=1}^{n} |\overline{\gamma}_i| \geq d(S_*, \Lambda^c)/r(\Phi)$. We can therefore write

$$\big|\Psi_{\Lambda,\omega}(X)\big| \leq e^{-\epsilon d(S_*,\Lambda^c)/r(\Phi)} \big|\Psi^{\epsilon}_{\Lambda,\omega}(X)\big|,$$

where $\Psi^{\epsilon}_{\Lambda,\omega}(X)$ is defined as $\Psi_{\Lambda,\omega}(X)$, with $\mathrm{w}(\gamma)$ replaced by $\mathrm{w}(\gamma)e^{\epsilon|\overline{\gamma}|}$. Since this modified weight $\mathrm{w}(\gamma)e^{\epsilon|\overline{\gamma}|}$ also satisfies the condition ensuring the convergence of the cluster expansion (see (6.124)),

$$\sum_{\substack{X:\overline{X}\ni i \\ \mathrm{diam}(\overline{X})\geq d(S_*,\Lambda^c)}} \big|\Psi_{\Lambda,\omega}(X)\big| \leq e^{-\epsilon d(S_*,\Lambda^c)/r(\Phi)} \sum_{X:\overline{X}\ni i} \big|\Psi^{\epsilon}_{\Lambda,\omega}(X)\big|.$$

This last series is convergent as before. Gathering these bounds leads to (6.53). □

6.13 Bibliographical references

The notion of Gibbs measure was introduced independently by Dobrushin [88] and Lanford and Ruelle [204]. It has since then been firmly established as the proper probabilistic description of large classical systems of particles in equilibrium.

The standard reference to this subject is the well-known book by Georgii [134]. Although our aim is to be more introductory, large parts of the present chapter have benefited from that book, and the interested reader can consult the latter for additional information and generalizations. We also strongly encourage the reader to have a look at its section on Bibliographical Notes, the latter containing a large amount of information, presented in a very readable fashion.

Texts containing some introductory material on Gibbs measures include, for example, the books by Prum [282], Olivieri and Vares [258], Bovier [37] and Rassoul-Agha and Seppäläinen [283], the monograph by Preston [278], the lecture notes by Fernández [101] and by Le Ny [213]. The paper by van Enter, Fernández, and Sokal [343] contains a very nice introduction, mostly without proofs, with a strong emphasis on the physical motivations behind the relevant mathematical concepts. The books by Israel [176] and Simon [308] also provide a general presentation of the subject, but their point of view is more functional-analytic than probabilistic.

Uniqueness. Dobrushin's uniqueness theorem, Theorem 6.31, was first proved in [88], but our presentation follows [108]; note that additional information can be extracted using the same strategy, such as exponential decay of correlations.

It can be shown that Dobrushin's condition of weak dependence cannot be improved in general [309, 178].

The one-dimensional uniqueness criterion given in Theorem 6.40 was originally proved in [49].

The approach in the proof of Theorem 6.38 is folklore.

Extremal decomposition. The integral decomposition (6.74) is usually derived from abstract functional-analytic arguments. Here, we follow the measure-theoretic approach exposed in [134], itself based on an approach of Dynkin [97].

Variational principle. The exposition in Section 6.9 is inspired by [134, Chapter 15]. For a more general version of the variational principle, see Pfister's lecture notes [274]. Israel's book [176] develops the whole theory of Gibbs measures from the point of view of the variational principle and is a beautiful example of the kind of results that can be obtained within this framework.

6.14 Complements and further reading

6.14.1 The equivalence of ensembles

The variational principle allowed us to determine which translation-invariant infinite-volume measures are Gibbs measures. In this section, we explain, at a heuristic level, how the same approach might be used to prove a general version of the *equivalence of ensembles*, which we already mentioned in Chapter 1 and in Section 4.7.1.

For simplicity, we avoid the use of boundary conditions. Consider a finite region $\Lambda \Subset \mathbb{Z}^d$ (for example a box), and let Ω_Λ be as before. To stay simple, assume that the Hamiltonian is just a function $\mathscr{H}_\Lambda : \Omega_\Lambda \to \mathbb{R}$.

In Chapter 1, we introduced several probability distributions on Ω_Λ that were good candidates for the description of a system at equilibrium. The first was the microcanonical distribution $\nu^{\mathrm{Mic}}_{\Lambda;U}$, defined as the uniform distribution on the energy shell $\Omega_{\Lambda;U} \stackrel{\mathrm{def}}{=} \{\omega \in \Omega_\Lambda : \mathscr{H}_\Lambda(\omega) = U\}$. The second one was the canonical Gibbs distribution at inverse temperature β defined as $\mu_{\Lambda;\beta} \stackrel{\mathrm{def}}{=} e^{-\beta\mathscr{H}_\Lambda}/\mathbf{Z}_{\Lambda;\beta}$.

Obviously , these two distributions *differ* in finite volume. In view of the equivalence between these different descriptions in thermodynamics, one might however hope that these distributions yield similar predictions for large systems, or even become "identical" in the thermodynamic limit, at least when U and β are related in a suitable way. Properly stated, this is actually true and can be proved using the theory of large deviations.

In this section, we give a hint as to how this can be shown, but since a full proof lies beyond the scope of this book, we will only motivate the result by a heuristic argument. The interested reader can find precise statements and detailed proofs in the papers of Lewis, Pfister and Sullivan [222, 223], or Georgii [133] and Deuschel, Stroock and Zessin [78]; a pedagogical account can be found in Pfister's lecture notes [274].

One way of trying to obtain the equivalence of $\nu^{\mathrm{Mic}}_{\Lambda;U}$ and $\mu_{\Lambda;\beta}$ in the thermodynamic limit is to proceed as in Proposition 6.81 and Theorem 6.82, and to find conditions under which

$$\frac{1}{|\Lambda|}\mathsf{H}_\Lambda(\nu^{\mathrm{Mic}}_{\Lambda;U} \,|\, \mu_{\Lambda;\beta}) \to 0\,, \quad \text{when } \Lambda \uparrow \mathbb{Z}^d\,. \tag{6.128}$$

Although the setting is not the same as the one of Section 6.9.3 (in particular, the distributions under consideration are defined in finite volume and are thus not translation invariant), the variational principle at least makes it plausible that when this limit is zero, the thermodynamic limit of $\nu^{\mathrm{Mic}}_{\Lambda;U}$ is an infinite-volume Gibbs measure. (Let us however emphasize that the proofs mentioned above do *not* proceed via (6.128); their approach is however similar in spirit.)

Remember that, for a finite system, a close relation between $\nu^{\mathrm{Mic}}_{\Lambda;U}$ and $\mu_{\Lambda;\beta}$ was established when it was shown, in Section 1.3, that if β is chosen properly as $\beta = \beta(U)$, then $\langle \mathscr{H}_\Lambda \rangle_{\mu_\beta} = U$ and $\mu_{\Lambda;\beta}$ has a maximal Shannon Entropy among all distributions with this property. A new look can be given at this relation, in the light of the variational principle and the thermodynamic limit. Namely, observe that

$$\begin{aligned}\frac{1}{|\Lambda|}\mathsf{H}_\Lambda(\nu^{\mathrm{Mic}}_{\Lambda;U}\,|\,\mu_{\Lambda;\beta}) &= -\frac{1}{|\Lambda|}S_\Lambda(\nu^{\mathrm{Mic}}_{\Lambda;U}) + \beta\Big\langle \frac{\mathscr{H}_\Lambda}{|\Lambda|}\Big\rangle^{\mathrm{Mic}}_{\Lambda;U} + \frac{1}{|\Lambda|}\log \mathbf{Z}_{\Lambda;\beta}\\ &= -\frac{1}{|\Lambda|}\log|\Omega_{\Lambda;U}| + \beta\frac{U}{|\Lambda|} + \frac{1}{|\Lambda|}\log\mathbf{Z}_{\Lambda;\beta}\,.\end{aligned}$$

In view of this expression, it is clear how (6.128) can be guaranteed. As was done for the variational principle in infinite volume, it is necessary to work with *densities*. So let us consider $\Lambda \uparrow \mathbb{Z}^d$, and assume that U also grows with the system, in such a way that $\frac{U}{|\Lambda|} \to u \in (h_{\min}, h_{\max})$, where $h_{\min} \stackrel{\mathrm{def}}{=} \inf_\Lambda \inf_{\omega_\Lambda} \frac{\mathscr{H}_\Lambda(\omega_\Lambda)}{|\Lambda|}$, and $h_{\max} \stackrel{\mathrm{def}}{=} \sup_\Lambda \sup_{\omega_\Lambda} \frac{\mathscr{H}_\Lambda(\omega_\Lambda)}{|\Lambda|}$.

As we explained in (1.37),

$$\lim \frac{1}{V}\log\mathbf{Z}_{\Lambda;\beta} = -\inf_{\tilde u}\{\beta\tilde u - \mathsf{s}_{\mathrm{Boltz}}(\tilde u)\}\,,$$

where $\mathsf{s}_{\mathrm{Boltz}}$ is the Boltzmann entropy density

$$\mathsf{s}_{\mathrm{Boltz}}(u) \stackrel{\mathrm{def}}{=} \lim \frac{1}{|\Lambda|}\log|\Omega_{\Lambda;U}|\,.$$

This shows that

$$\lim \frac{1}{|\Lambda|}\mathsf{H}_\Lambda(\nu^{\mathrm{Mic}}_{\Lambda;U}\,|\,\mu_{\Lambda;\beta}) = \beta u - \mathsf{s}_{\mathrm{Boltz}}(u) - \inf_{\tilde u}\{\beta\tilde u - \mathsf{s}_{\mathrm{Boltz}}(\tilde u)\}\,. \tag{6.129}$$

Now, the infimum above is realized for a particular value $\tilde u = \tilde u(\beta)$. If β is chosen in such a way that $\tilde u(\beta) = u$, we see that the right-hand side of (6.129) vanishes as desired. To see when this is possible, an analysis is required, along the same lines as what was done in Chapter 4 to prove the equivalence of the canonical and grand canonical ensembles at the level of thermodynamic potentials.

We thus conclude that *if equivalence of ensembles holds at the level of the thermodynamic potentials, then it should also hold at the level of measures.* As mentioned above, this conclusion can be made rigorous.

6.14.2 Pathologies of transformations and weaker notions of Gibbsianness.

The notion of Gibbs measure presented in this chapter, although efficient for the description of infinite systems in equilibrium, is not as robust as one might expect: the image of a Gibbs measure under natural transformations $T : \Omega \to \Omega$ can cease to be Gibbsian. An example of such a transformation has been mentioned in Section 3.10.11, when motivating the renormalization group.

Consider for example the two-dimensional Ising model at low temperature. Let $\mathscr{L} \stackrel{\text{def}}{=} \{(i,0) \in \mathbb{Z}^2 : i \in \mathbb{Z}\}$ and consider the projection $\Pi_{\mathscr{L}} : \omega = (\omega_i)_{i\in\mathbb{Z}^2} \mapsto (\omega_j)_{j\in\mathscr{L}}$. The image of $\mu^+_{\beta,0}$ under $\Pi_{\mathscr{L}}$ is a measure ν^+_β on $\{\pm1\}^{\mathbb{Z}}$. It was shown by Schonmann [295] that ν^+_β *is not a Gibbs measure*: there exists no absolutely summable potential Φ so that ν^+_β is compatible with the Gibbsian specification associated to Φ.

Before that, from a more general point of view, it had already been observed by Griffiths and Pearce [147], and Israel [177], that the same kind of phenomenon occurs when implementing rigorously certain renormalization group transformations. This is an important observation inasmuch as the renormalization group is often presented in the physics literature as a map defined on the space of all interactions (or Hamiltonians) (see the brief discussion in Section 3.10.11). What this shows is that such a map, which can always be defined on the set of probability measures, does not induce, in general, a map on the space of (physically reasonable) interactions.

More recently, there has been interest in whether the evolution of a Gibbs measure at temperature T under a stochastic dynamics corresponding to another temperature T' remains Gibbsian (which would again mean that one could follow the dynamics on the space of interactions). The observation is that the Gibbsian character can be quickly lost, depending on the values of T and T', see [341] for example.

A general discussion of this type of issues can be found in [343].

These so-called *pathologies* have led to the search for weaker notions of Gibbs measures, which would encompass the one presented in this chapter but would remain stable under transformations such as the one described above. This research had originally been initiated by Dobrushin, and is nowadays known as *Dobrushin's restoration program*. A summary of the latter can be found in the review of van Enter, Maes and Shlosman [342]. Other careful presentations of the subject are [101] and [213].

6.14.3 Gibbs measures and the thermodynamic formalism.

The ideas and techniques of equilibrium statistical mechanics have been useful in the theory of dynamical systems. For instance, Gibbs measures were introduced in ergodic theory by Sinai [311]. Moreover, the characterization of Gibbs measures via the variational principle of Section 6.9 is well suited for the definition of Gibbs measures in other settings. In symbolic dynamics, for instance, an invariant probability measure is said to be an *equilibrium measure* if it satisfies the variational principle. The monograph [40] by Bowen is considered as a pioneering contributions to this field. See also the books by Ruelle [291] or Keller [187], as well as Sarig's lecture notes [294].