Theoretical and Mathematical Physics

Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics

Bearbeitet von
Errico Presutti

1. Auflage 2008. Buch. xvi, 470 S. Hardcover
ISBN 978 3 540 73304 1
Format (B x L): 15,5 x 23,5 cm
Gewicht: 890 g

Chapter 2
Thermodynamic limit in the Ising model

Modern statistical mechanics rests on the Gibbs hypothesis that in a system in equilibrium with a reservoir at temperature T, the probability of observing a state is proportional to $e^{-E/kT}$, where E is the energy of the state, k the Boltzmann constant, and T the absolute temperature. Macroscopic equilibrium behavior (in particular thermodynamics) can then be derived via a sharp separation of scales implemented mathematically by requiring that the ratio between inter-atomic distances and macroscopic lengths vanishes. Such a scaling limit procedure is then referred to as "the thermodynamic limit."

We will discuss all that in the simpler context of the Ising model. In Sect. 2.1 we will study the Gibbs measures in bounded domains, and in Sect. 2.2 their infinite volume limits (thermodynamic limit). The limit measures obtained in this way will be characterized by the DLR property (DLR stands for Dobrushin, Lanford and Ruelle); the general structure of the DLR measures will then be examined in detail. Thermodynamic phases will be related to "extremal" DLR measures and consequently the occurrence of a phase transition to the non-uniqueness of DLR measures.

In Sects. 2.3 and 2.4 we will focus on the thermodynamic potentials of the Ising model. Using only the Boltzmann hypothesis to identify the thermodynamic entropy as the log of the number of states with given energy, we will derive the well known formula which relates the pressure to the log of the partition function; see Sect. 2.3. Formulas for all the other thermodynamic potentials will then be obtained by using the thermodynamic relations establishing a bridge between the macroscopic properties of a body and its microscopic interactions. The power of the thermodynamic formalism will become evident in Sect. 2.4, where it is applied in several different contexts. In particular, we shall see that the Gibbs assumption and DLR measures can be actually derived from the thermodynamic potentials, and we shall also show that large deviations naturally fall in the formalism.

2.1 Finite volume Gibbs measures

In this section we will study the finite volume Gibbs measures in the context of the Ising model. Ising configurations are collections of spins on the lattice $\mathbb{Z}^d$. Gibbs measures are then probabilities on the phase space of Ising configurations, they are defined by the Gibbs formula $Ce^{-E/kT}$, C a normalization constant (whose inverse is called the "partition function"), E the energy, k the Boltzmann constant and T the absolute temperature. The energy of an Ising configuration will be specified by the spin–spin interaction, the interaction of the spins with an external magnetic field and with the "external" spins, which act as boundary conditions.

Errico Presutti, *Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics*,

2.1.1 Spin configurations, phase space

The Ising spin configurations (denoted by σ) are ± 1 valued functions on $\mathbb{Z}^d$; the collection of all spin configurations is the Ising phase space $\mathcal{X} = \{-1, 1\}^{\mathbb{Z}^d}$. A spin configuration is therefore the collection

$$\sigma = \{\sigma(x),\ x \in \mathbb{Z}^d\},$$

$\sigma(x) = \pm 1$ (up or down) being the spin at site x in the configuration σ.

Analogously, spin configurations in $\Lambda \subset \mathbb{Z}^d$ are functions on Λ with values ± 1, namely elements σ_Λ of $\mathcal{X}_\Lambda = \{-1, 1\}^\Lambda$,

$$\sigma_\Lambda = \{\sigma_\Lambda(x),\ x \in \Lambda\}.$$

The restriction map from $\mathcal{X}$ to $\mathcal{X}_\Lambda$, denoted by $\restriction_\Lambda$, is defined as

$$\restriction_\Lambda (\sigma) = \sigma_\Lambda = \{\sigma(x), x \in \Lambda\}, \tag{2.1.1.1}$$

and when there is no room for doubt, we will simply write σ_Λ for $\restriction_\Lambda (\sigma)$. We will often use the expression $(\sigma_\Lambda, \sigma_\Delta)$, $\Lambda \cap \Delta = \emptyset$ to denote the element in $\mathcal{X}_{\Lambda \sqcup \Delta}$ whose restrictions to Λ and Δ are respectively σ_Λ and σ_Δ.

We regard $\mathcal{X}$ as a topological space with the product topology, namely a sequence $\sigma^{(n)} \to \sigma$ if and only if for any $x \in \mathbb{Z}^d$, $\sigma^{(n)}(x) = \sigma(x)$ for all n large enough. A countable basis of open sets is the collection of cylindrical sets. We define

Cylindrical functions and sets *A function f on $\mathcal{X}$ is cylindrical in Δ if it only depends on the restriction σ_Δ of σ to Δ. A set is cylindrical in Δ if its characteristic function is cylindrical in Δ. A function or a set is cylindrical if it is cylindrical in a bounded region Δ. Elementary cylinders are sets of the form $C_{\sigma_\Lambda} = \{\sigma' \in \mathcal{X} : \sigma'_\Lambda = \sigma_\Lambda\}$ with Λ bounded and $\sigma_\Lambda \in \mathcal{X}_\Lambda$. Their collection is denoted by $\mathcal{C}$.*

Cylindrical sets are both open and closed; cylindrical functions are evidently continuous, and vice versa any continuous function can be approximated in sup norm by cylindrical functions; see Appendix A.

2.1.2 Energy

The energy of an Ising spin system is a family $\{H_\Lambda(\sigma_\Lambda)\}$ with Λ running over all the bounded subsets of $\mathbb{Z}^d$ and $\sigma_\Lambda \in \mathcal{X}_\Lambda$. $H_\Lambda(\sigma_\Lambda)$ is the energy in Λ of the spin configuration σ_Λ, it includes all the interactions of the spins of Λ among themselves and the interaction of the spins of Λ with an external magnetic field, if present. The interaction between two disjoint, bounded regions Λ and Δ is then defined by

$$W_{\Lambda,\Delta}(\sigma_\Lambda, \sigma_\Delta) = H_{\Lambda \cup \Delta}(\sigma_\Lambda, \sigma_\Delta) - H_\Lambda(\sigma_\Lambda) - H_\Delta(\sigma_\Delta), \tag{2.1.2.1}$$

where $(\sigma_\Lambda, \sigma_\Delta) \in \mathcal{X}_{\Lambda \sqcup \Delta}$ is the configuration whose restrictions to Λ and Δ are σ_Λ and σ_Δ. The energy in Λ under the field produced by the spins in Δ, $\Lambda \cap \Delta = \emptyset$ is

$$H_{\Lambda,\Delta}(\sigma_\Lambda|\sigma_\Delta) = H_{\Lambda\sqcup\Delta}(\sigma_\Lambda, \sigma_\Delta) - H_\Delta(\sigma_\Delta), \tag{2.1.2.2}$$

and

$$H_{\Lambda,\Lambda^c}(\sigma_\Lambda|\sigma_{\Lambda^c}) = \lim_{\Delta\nearrow\Lambda^c} \big(H_{\Lambda\cup\Delta}(\sigma_\Lambda, \sigma_\Delta) - H_\Delta(\sigma_\Delta)\big), \tag{2.1.2.3}$$

if the limit exists independently of the sequence approximating Λ^c. The above notation is redundant because the region appearing as a suffix in the energy can be read off from the spin configurations, so that when no confusion arises we may drop them from the notation. In particular, $H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})$ will always stand for $H_{\Lambda,\Lambda^c}(\sigma_\Lambda|\sigma_{\Lambda^c})$.

We will restrict in the sequel to energies of the form

$$H_\Lambda(\sigma_\Lambda) = -\frac{1}{2} \sum_{x\neq y\in\Lambda} J(x,y)\sigma_\Lambda(x)\sigma_\Lambda(y) - h\sum_{x\in\Lambda}\sigma_\Lambda(x). \tag{2.1.2.4}$$

If Λ is a singleton, $\Lambda = \{x\}$,

$$H_{\{x\}}(\sigma_{\{x\}}(x)) = -h\sigma_{\{x\}}(x).$$

h is interpreted as an external magnetic field and $-h\sigma_\Lambda(x)$ is the energy of the spin at x under the sole influence of the external magnetic field; indeed such a term is the only one surviving when $\Lambda = \{x\}$. $-J(x,y)\sigma_\Lambda(x)\sigma_\Lambda(y)$ is the interaction energy between the spins at x and y as follows from (2.1.2.1) with $\Lambda = \{x\} \cup \{y\}$. Notice finally that (2.1.2.2) with $\{H_\Lambda(\sigma_\Lambda)\}$ as in (2.1.2.4) becomes

$$H_{\Lambda,\Delta}(\sigma_\Lambda|\sigma_\Delta) = H_\Lambda(\sigma_\Lambda) - \sum_{x\in\Lambda}\sum_{y\in\Delta} J(x,y)\sigma_\Lambda(x)\sigma_\Delta(y). \tag{2.1.2.5}$$

Assumptions on the interaction

- *symmetry*: $J(x,y) = J(y,x)$.
- *translational invariance*: $J(x+z, y+z) = J(x,y)$, *for all* x, y *and* z *in* $\mathbb{Z}^d$.
- *summability*: $\sum_{x\neq 0}|J(0,x)| < \infty$ (*by the summability assumption, the series in* (2.1.2.5) *is convergent*).

Under the above summability assumption, (2.1.2.3) becomes

$$H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) = H_\Lambda(\sigma_\Lambda) - \sum_{x\in\Lambda}\sum_{y\in\Lambda^c} J(x,y)\sigma_\Lambda(x)\sigma_{\Lambda^c}(y).$$

Examples

- *Classical Ising model*. The only active bonds are those connecting nearest neighbor (n.n.) sites, with ferromagnetic coupling constants all equal to $J > 0$.

- *Kac potentials*. The coupling constants $J_\gamma(x, y)$ depend on a parameter $\gamma > 0$:

$$J_\gamma(x, y) = \gamma^d J(\gamma x, \gamma y),$$

where $J(r, r') = J(r + a, r' + a) \geq 0$, for all r, r' and a in $\mathbb{R}^d$; $J(0, r)$ is continuous with compact support and normalized as a probability kernel:

$$\int_{\mathbb{R}^d} dr\, J(0, r) = 1.$$

We are interested in small γ, a regime characterized by (i) long range interactions $\approx \gamma^{-1}$, and a large, $\approx \gamma^{-d}$, connectivity of each site (i.e. the number of active bonds starting from that site); (ii) the coupling constants of the bonds are small, $\approx \gamma^d$, and (iii) the total strength of a site (i.e. the sum of all the coupling constants of bonds originating from that site) is ≈ 1.
- *Mean field models*. Here the coupling constants depend on the region Λ where the system is studied. If Λ has N sites $J(x, y) = N^{-1}$, x, y in Λ. The model shares the properties (i), (ii) and (iii) of the previous one, which was indeed conceived of as a refinement of the mean field model to correct its various unphysical features. The mean field model, though, has the great advantage of providing a simple and not too unrealistic mechanism for phase transitions.

2.1.3 Potentials

Often the energy of the system is given indirectly by assigning its potential. The potential is a family $\{U_\Delta(\sigma_\Delta)\}$, Δ running over the bounded sets and $\sigma_\Delta \in \mathcal{X}_\Delta$. Given a potential, its energy is

$$H(\sigma_\Lambda) = \sum_{\Delta \subseteq \Lambda} U_\Delta(\sigma_\Delta), \tag{2.1.3.1}$$

where σ_Δ above is the restriction to Δ of σ_Λ.

As already observed the potential associated to the energy (2.1.2.4) is made by the potential with $U_{\{x\}}(\sigma_{\{x\}}) = -h\sigma_{\{x\}}(x)$, $U_{\{x,y\}}(\sigma_{\{x,y\}}) = -J(x, y)\sigma_{\{x,y\}}(x) \times \sigma_{\{x,y\}}(y)$, while all other $U_\Delta = 0$.

The relation (2.1.3.1) can be inverted, namely, given $\{H_\Lambda(\sigma_\Lambda)\}$, we can recover uniquely $\{U_\Delta(\sigma_\Delta)\}$ in such a way that (2.1.3.1) holds. This is done iteratively starting from sets of cardinality 1 and then increasing progressively the cardinality, noticing that the potential with the set having maximal cardinality can be expressed using (2.1.3.1) in terms of the energy and of the potentials with smaller cardinality.

As stated earlier we will restrict ourselves to one and two body potentials only, but in magnetic systems also quadrupole and multipole interactions may be relevant. Theoretically the many body interactions are quite important, and an example is provided in Sect. 2.4.2 where a relation is established between DLR measures and the thermodynamic pressure as a function of the general many body potential. In general, many body potentials arise after coarse graining transformations and describe effective hamiltonians, as discussed in Chaps. 9 and 10.

Example Potentials are often written as

$$U_{\{x_1,\dots,x_n\}}(\sigma_{\{x_1,\dots,x_n\}}) = -J(x_1,\dots,x_n)\sigma_{\{x_1,\dots,x_n\}}(x_1)\cdots\sigma_{\{x_1,\dots,x_n\}}(x_n), \qquad (2.1.3.2)$$

where $J(x_1,\dots,x_n)$ is a symmetric function on $(\mathbb{Z}^d)^n$. Then (2.1.3.1) yields

$$H(\sigma_\Lambda) = -\sum_{n=1}^{|\Lambda|}\frac{1}{n!}\sum_{x_1\neq\cdots\neq x_n\in\Lambda} J(x_1,\dots,x_n)\sigma_\Lambda(x_1)\cdots\sigma_\Lambda(x_n).$$

One may wonder why one would take only a multi-linear dependence on the spins in (2.1.3.2). This is specific of Ising spins where all functions are necessarily sums of multi-linear terms as shown below. A function f cylindrical in Δ, Δ bounded, can be written as

$$f(\sigma) = \sum_{a_\Delta\in\mathcal{X}_\Delta} f(a_\Delta)\prod_{x\in\Delta}\frac{1+a_\Delta(x)\sigma(x)}{2}, \qquad (2.1.3.3)$$

where $f(a_\Delta)$ denotes the values of $f(\sigma)$ when the restriction of σ to Δ is a_Δ. Thus any cylindrical function in Δ is a multi-linear polynomial of the variables $\{\sigma(x), x\in\Delta\}$. In particular, if C is a cylindrical set in Δ, Δ bounded,

$$\mathbf{1}_{\sigma\in C} = \sum_{a_\Delta\in C}\prod_{x\in\Delta}\frac{1+a_\Delta(x)\sigma(x)}{2}.$$

2.1.4 Finite volume Gibbs measures

In the context of the Ising model the Gibbs hypothesis says that the equilibrium state at temperature T of the Ising spin system in a bounded region Λ is described by a probability distribution on $\mathcal{X}_\Lambda$ given by the formula

$$\text{probability of observing } \sigma_\Lambda = \frac{e^{-\beta E_\Lambda(\sigma_\Lambda)}}{Z_\Lambda}, \qquad (2.1.4.1)$$

where $\beta = 1/kT$, $E_\Lambda(\sigma_\Lambda)$ is the energy of the configuration σ_Λ and Z_Λ a normalization constant. The energy E_Λ is not necessarily the same as the energy H_Λ of (2.1.2.4), as there may be interactions with the "walls" and/or with "the world" outside Λ.

Remarks

Equation (2.1.4.1) is supposed to describe the equilibrium states of the collection of spins σ_Λ in contact with a thermal reservoir at fixed temperature T which exchanges energy with the spins in Λ at a rate determined by T. The physical assumption is that the interaction is so weak that the energy levels E_Λ are not affected by the reservoir, yet it is at the same time so strong that it eventually drives the system to a final equilibrium, independently of its initial conditions. Equilibrium is eventually established throughout the system due to complex mechanisms which involve various space and time scales, hydrodynamic behaviors and all the other relevant non-equilibrium phenomena. But no matter how complicated the pattern to equilibrium is, according to Gibbs the final equilibrium has the very simple expression (2.1.4.1).

The physical meaning of representing a state as a probability and in particular the equilibrium state as in (2.1.4.1) is that if we (ideally) make repeated observations of the system after equilibrium is established, the frequency with which we observe a configuration σ_Λ is given by the Gibbs formula (2.1.4.1).

The energy $E_\Lambda(\sigma_\Lambda)$

With reference to the system defined in Sect. 2.1.2, the energy $E_\Lambda(\sigma_\Lambda)$ is not necessarily the same as the energy $H_\Lambda(\sigma_\Lambda)$ of (2.1.2.4), as there may be interactions with the "walls," namely interactions with "the world" outside Λ. They are usually quite complex, but they affect significantly only the spins close to the boundaries of Λ. We may thus suppose that

$$E_\Lambda(\sigma_\Lambda) = H_\Lambda(\sigma_\Lambda) - \sum_{x\in\Lambda} h_x^{\Lambda^c}\sigma_\Lambda(x), \tag{2.1.4.2}$$

where $h_x^{\Lambda^c}$ is an "effective" magnetic field which takes into account the interactions with Λ^c and which decays as $\mathrm{dist}(x,\Lambda^c)$ increases. If the interaction has many body potentials, more general expressions will arise, but for simplicity we will only consider here the case (2.1.4.2). The energy (2.1.4.2) refers to the ideal case where the effective magnetic field $h_x^{\Lambda^c}$ does not fluctuate, i.e. when outside Λ everything is frozen and does not change in time. We will also suppose (most of the times) that $h_x^{\Lambda^c}$ arises from the interaction with a fixed spin configuration σ_{Λ^c} so that

$$h_x^{\Lambda^c} = -\sum_{y\in\Lambda^c} J(x,y)\sigma_{\Lambda^c}(y), \quad \text{and} \quad E_\Lambda(\sigma_\Lambda) = H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}). \tag{2.1.4.3}$$

This is not a very realistic model of a wall, but it has considerable theoretical importance as we shall see.

The Gibbs measure $G_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})$

When we use the choice (2.1.4.3) we will write (2.1.4.1) as

$$G_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) = \frac{e^{-\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}}{Z_\Lambda(\sigma_{\Lambda^c})}, \quad Z_\Lambda(\sigma_{\Lambda^c}) = \sum_{\sigma'_\Lambda \in \mathcal{X}_\Lambda} e^{-\beta H_\Lambda(\sigma'_\Lambda|\sigma_{\Lambda^c})}. \tag{2.1.4.4}$$

The dependence on β and on the parameters defining the energy are not made explicit; if necessary they will be added as subscripts. The "partition function" $Z_\Lambda(\sigma_{\Lambda^c})$, which in (2.1.4.4) appears just as a normalization factor, has instead a meaning that is important physically, being directly related to the thermodynamic pressures; see Sect. 2.3.

By abuse of notation, we will use the same symbol $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ for the probability on $\mathcal{X}$ which is a finite sum of Dirac deltas, namely calling $\{\sigma'\}$ the set consisting of the singleton σ', $G_\Lambda(\{\sigma'\}|\sigma_{\Lambda^c})$ is equal to 0 for all σ' except those whose restriction to Λ^c is exactly σ_{Λ^c}; thus

$$G_\Lambda(\{\sigma'\}|\sigma_{\Lambda^c}) = \mathbf{1}_{\sigma'_{\Lambda^c}=\sigma_{\Lambda^c}} \frac{e^{-\beta H_\Lambda(\sigma'_\Lambda|\sigma_{\Lambda^c})}}{Z_\Lambda(\sigma_{\Lambda^c})}, \tag{2.1.4.5}$$

where, to make notation easier, we may just write $G_\Lambda(\sigma'|\sigma_{\Lambda^c})$ for $G_\Lambda(\{\sigma'\}|\sigma_{\Lambda^c})$. We shall switch from (2.1.4.4) to (2.1.4.5) freely, and the reader should understand from the context the meaning of $G_\Lambda(\cdot|\sigma_{\Lambda^c})$.

2.1.5 A consistency property of Gibbs measures

In this subsection we will prove that the conditional probability of a Gibbs measure is also Gibbs (hence the title of the subsection) and show that this has a nice physical meaning.

Marginal distributions

Let m_Λ be a probability on $\mathcal{X}_\Lambda$, Λ a bounded set in $\mathbb{Z}^d$, and Δ a proper subset of Λ. Then the marginal distribution [or simply the marginal] of m_Λ on $\mathcal{X}_\Delta$ is the probability $(m_\Lambda)_\Delta$ on $\mathcal{X}_\Delta$ defined by

$$(m_\Lambda)_\Delta(\sigma^*_\Delta) = \sum_{\sigma_\Lambda:\sigma_\Lambda(x)=\sigma^*_\Delta(x)\, x\in\Delta} m_\Lambda(\sigma_\Lambda) \equiv \sum_{\sigma_{\Lambda\setminus\Delta}\in\mathcal{X}_{\Lambda\setminus\Delta}} m_\Lambda(\sigma^*_\Delta, \sigma_{\Lambda\setminus\Delta}).$$

The marginal of $G_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})$ on $\mathcal{X}_{\Lambda\setminus\Delta}$ is

$$(G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma^*_{\Lambda\setminus\Delta}) = \sum_{\sigma_\Lambda:\sigma_\Lambda=\sigma^*_{\Lambda\setminus\Delta} \text{ on } \Lambda\setminus\Delta} G_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}).$$

Theorem 2.1.5.1 *With the above notation*

$$G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c}) = \frac{G_\Lambda(\sigma^*_\Delta,\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c})}{(G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma^*_{\Lambda\setminus\Delta})}. \tag{2.1.5.1}$$

Proof

$$\text{r.h.s. of (2.1.5.1)} = \frac{e^{-\beta H_\Lambda(\sigma^*_\Delta,\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c})}}{\sum_{\sigma_\Delta} e^{-\beta H_\Lambda(\sigma_\Delta,\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c})}}. \tag{2.1.5.2}$$

Equation (2.1.5.1) then follows by using the identity

$$H_\Lambda\big(\sigma'_\Delta,\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c}\big) = H_\Delta\big(\sigma'_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c}\big) + H_{\Lambda\setminus\Delta,\Lambda^c}\big(\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c}\big);$$

see (2.1.2.5) for notation, with $\sigma'_\Delta = \sigma^*_\Delta$ in the numerator of the fraction in (2.1.5.2) and $\sigma'_\Delta = \sigma_\Delta$ in the denominator. □

Remarks

Equation (2.1.5.1) has a nice physical interpretation. Suppose we make an ideal experiment where we measure N times the state of the system in equilibrium. By the Gibbs hypothesis the number of times we find in $\Lambda\setminus\Delta$ the configuration $\sigma^*_{\Lambda\setminus\Delta}$ is $\approx (G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma^*_{\Lambda\setminus\Delta})N$. We then select those experiments where we have seen $\sigma^*_{\Lambda\setminus\Delta}$ and count the relative frequency of appearance of $\sigma^*_\Delta \in \mathcal{X}_\Delta$, namely the number of times when we see both σ^*_Δ and $\sigma^*_{\Lambda\setminus\Delta}$, that is $\approx N \times G_\Lambda(\sigma^*_\Delta,\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c})$, over the number of times when we see $\sigma^*_{\Lambda\setminus\Delta}$, i.e. $\approx N \times (G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma^*_{\Lambda\setminus\Delta})$. Thus the conditional frequency in the limit $N\to\infty$ is $\frac{G_\Lambda(\sigma^*_\Delta,\sigma^*_{\Lambda\setminus\Delta}|\sigma_{\Lambda^c})}{(G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma^*_{\Lambda\setminus\Delta})}$ which is equal to the r.h.s. of (2.1.5.1), hence to $G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c})$, i.e. to the Gibbs probability in Δ as if the spins in $\Lambda\setminus\Delta$ were frozen, just like those in Λ^c. Thus the Gibbs probability $G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c})$ is not only the probability of observing σ^*_Δ when $\sigma^*_{\Lambda\setminus\Delta}$ and σ_{Λ^c} are frozen, but also the conditional probability of observing σ^*_Δ conditioned to having observed $\sigma^*_{\Lambda\setminus\Delta}$ (and with σ_{Λ^c} frozen).

Equation (2.1.5.1) can be read as meaning that *he conditional* $G_\Lambda(\cdot|\sigma_{\Lambda^c})$*-probability of* σ^*_Δ *given* $\sigma^*_{\Lambda\setminus\Delta}$ *is equal to* $G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c})$ because we have the following.

Conditional probabilities

Let m be a probability on a space Ω, $C\subset\Omega$, $m(C)>0$. Then the ratio $\frac{m(A\cap C)}{m(C)} =: m(A|C)$ is the m-conditional probability of A given C. If $\pi = (C_1,\dots,C_n)$ is a partition of Ω with $m(C_i)>0$, $i=1,\dots,n$, then

$$m(A) = \sum_{i=1}^n m(A|C_i)m(C_i). \tag{2.1.5.3}$$

Thus (2.1.5.1) is just the statement at the beginning of the subsection that the Gibbs $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ conditional probability of observing σ^*_Δ given $\sigma^*_{\Lambda\setminus\Delta}$ is equal to $G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c})$. Indeed the numerator on the r.h.s. of (2.1.5.1) is the probability of having σ^*_Δ in Δ intersected with the event $\sigma^*_{\Lambda\setminus\Delta}$ in $\Lambda\setminus\Delta$; the denominator, being the marginal probability of $\sigma^*_{\Lambda\setminus\Delta}$, is the probability of $\sigma^*_{\Lambda\setminus\Delta}$. The ratio is by definition the conditional probability of σ^*_Δ given $\sigma^*_{\Lambda\setminus\Delta}$.

Consider (2.1.5.3) with the partition $\pi=\{C_{\sigma^*_{\Lambda\setminus\Delta}},\sigma^*_{\Lambda\setminus\Delta}\in\mathcal{X}_{\Lambda\setminus\Delta}\}$, $C_{\sigma^*_{\Lambda\setminus\Delta}}$ being the set of all σ_Λ whose restriction to $\Lambda\setminus\Delta$ is $\sigma^*_{\Lambda\setminus\Delta}$. Then (2.1.5.3) reads

$$\begin{aligned}&(G_\Lambda(\cdot|\sigma_{\Lambda^c}))_\Delta(\sigma^*_\Delta)\\ &\quad=\sum_{\sigma^*_{\Lambda\setminus\Delta}\in\mathcal{X}_{\Lambda\setminus\Delta}} G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c})\,(G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma^*_{\Lambda\setminus\Delta}).\end{aligned}\tag{2.1.5.4}$$

2.1.6 Random boundary conditions

The Gibbs measures $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ cannot exhaust the set of all equilibrium measures in Λ as the correct notion should reflect the physical request that if a system is in equilibrium in a region Λ then it is also in equilibrium in all subregions Δ of Λ. However, the equilibrium state $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ observed in the subregion Δ is described by the marginal of $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ on $\mathcal{X}_\Delta$, and it is thus given by (2.1.5.4) which does not have the expression (2.1.4.1) unless $\Delta=\Lambda$.

We thus need to relax the statement that the equilibrium measures have the form $G_\Lambda(\cdot|\sigma_{\Lambda^c})$. Since by (2.1.5.4) the marginal of $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ on $\mathcal{X}_\Delta$ is a convex combination of measures $G_\Delta(\sigma_\Delta|\sigma_{\Delta^c})$, we certainly want to include among the equilibrium measures in Δ probabilities of the form

$$\mu(\sigma^*_\Delta)=\sum_{\sigma^*_{\Lambda\setminus\Delta}\in\mathcal{X}_{\Lambda\setminus\Delta}} G_\Delta(\sigma^*_\Delta|\sigma^*_{\Lambda\setminus\Delta},\sigma_{\Lambda^c})\,m(\sigma^*_{\Lambda\setminus\Delta}),\tag{2.1.6.1}$$

with m a probability on $\mathcal{X}_{\Lambda\setminus\Delta}$. Referring to Appendix A for definitions and notation relative to measures on Ising spaces we formalize the above considerations as follows.

Definition A probability μ on $\mathcal{X}$ is a "*Gibbs measure in Λ with random boundary conditions*" if ($\mu(f)$ is the integral of f)

$$\mu(f)=\int_{\mathcal{X}}\sum_{\sigma'\in\mathcal{X}} G_\Lambda\big(\sigma'|\sigma_{\Lambda^c}\big)f(\sigma')\,\nu(d\sigma),$$

$$\text{for any bounded, measurable } f,\tag{2.1.6.2}$$

where ν is a Borel probability on $\mathcal{X}$ and measurability means Borel measurability. $\mathcal{G}_\Lambda$ denotes the set of Gibbs measures in Λ with random boundary conditions and it is identified with the set of equilibrium measures in Λ.

The integral on the r.h.s. of (2.1.6.2) is well defined because

Lemma 2.1.6.1 *For any bounded, measurable function f,*

$$b_f(\sigma) := G_\Lambda\big(f|\sigma_{\Lambda^c}\big) \equiv \sum_{\sigma'\in\mathcal{X}} G_\Lambda\big(\sigma'|\sigma_{\Lambda^c}\big) f(\sigma') \tag{2.1.6.3}$$

is a bounded, measurable function of σ, which is continuous if f is.

Proof Obviously $|b_f(\sigma)| \le \|f\|_\infty$. Writing $(\sigma'_\Lambda, \sigma_{\Lambda^c})$ for the configuration whose restrictions to Λ and Λ^c are σ'_Λ and σ_{Λ^c}, we have

$$b_f(\sigma) = \sum_{\sigma'_\Lambda} G_\Lambda\big((\sigma'_\Lambda, \sigma_{\Lambda^c})\big|\sigma_{\Lambda^c}\big) f\big((\sigma'_\Lambda, \sigma_{\Lambda^c})\big).$$

Since the sum is finite, it is sufficient to prove that each term is measurable. $\sigma \to f((\sigma'_\Lambda, \sigma_{\Lambda^c}))$ is measurable by assumption, while

$$\big|G_\Lambda\big(\sigma'_\Lambda\big|\sigma_{\Lambda^c}\big) - G_\Lambda\big(\sigma'_\Lambda\big|\sigma^N_{\Lambda^c}\big)\big| \le \epsilon(N), \tag{2.1.6.4}$$

where σ^N is obtained from σ by replacing $\sigma(x)$ by $+1$ whenever $|x| > N$ and $\epsilon(N)$ is a function of N which vanishes as $N \to \infty$. (2.1.6.4) follows from the summability assumption on the interaction.

Since $\sigma \to G_\Lambda(\sigma'_\Lambda|\sigma^N_{\Lambda^c})$ is continuous (being a cylindrical function), the above shows that $G_\Lambda(\sigma'_\Lambda|\sigma_{\Lambda^c})$ is approximated in sup norm by cylindrical functions and it is therefore itself a continuous function of σ (see Theorem A.1). We have proved that $b_f(\sigma)$ is measurable and continuous as well if f is continuous; hence Lemma 2.1.6.1 is proved. □

We conclude the subsection with another lemma which will play an important role in the sequel.

Lemma 2.1.6.2 *$\mu \in \mathcal{G}_\Lambda$ if and only if for any bounded, measurable f*

$$\mu(f) = \int_{\mathcal{X}} \sum_{\sigma'} G_\Lambda\big(\sigma'|\sigma_{\Lambda^c}\big) f(\sigma')\, \mu(d\sigma). \tag{2.1.6.5}$$

Proof μ is obviously in $\mathcal{G}_\Lambda$ if (2.1.6.5) holds. If μ is given by (2.1.6.2), then for any f $\mu(f) = \nu(b_f)$, b_f as in (2.1.6.3). We denote by $\mathcal{B}_{\Lambda^c}$ the σ-algebra generated by the cylindrical sets in Λ^c. If f is $\mathcal{B}_{\Lambda^c}$ measurable by (2.1.6.3) $b_f(\sigma) = f(\sigma)$, and since in general $\mu(f) = \nu(b_f)$ we have $\mu(f) = \nu(f)$ for all $\mathcal{B}_{\Lambda^c}$ measurable functions f. Take now any (bounded, measurable) f, then $\mu(f) = \nu(b_f)$, by (2.1.6.2), and since b_f is $\mathcal{B}_{\Lambda^c}$ measurable, $\mu(f) = \nu(b_f) = \mu(b_f)$, which is (2.1.6.5). □

Remarks Let ν in (2.1.6.2) be the measure supported by the single configuration σ, then $\mu = G_\Lambda(\cdot|\sigma_{\Lambda^c})$, thus $G_\Lambda(\cdot|\sigma_{\Lambda^c}) \in \mathcal{G}_\Lambda$ as implicit in the whole discussion so far.

2.1.7 Structure of finite volume Gibbs measures

With the identification of the equilibrium measures as Gibbs measures with random boundary conditions, the original Gibbs measures $G_\Lambda(\cdot|\sigma_{\Lambda^c})$ do not only belong to $\mathcal{G}_\Lambda$ as stated in the last remark of Sect. 2.1.6, but, as a consequence of (2.1.5.4), they are also in any $\mathcal{G}_\Delta$, $\Delta \subset \Lambda$. Indeed, we have the following.

Theorem 2.1.7.1 *$\mathcal{G}_\Lambda$ is a convex, weakly compact set and*

$$\mathcal{G}_\Lambda \subset \mathcal{G}_\Delta, \quad \textit{for any } \Delta \subset \Lambda.$$

In particular, $G_\Lambda(\cdot|\sigma_{\Lambda^c}) \in \mathcal{G}_\Delta$ for any $\Delta \subset \Lambda$.

Proof It directly follows from the representation (2.1.6.2) that $\mathcal{G}_\Lambda$ is a convex set. We will next prove that $\mathcal{G}_\Lambda$ is closed (and since the space $M(\mathcal{X})$ of all probabilities on $\mathcal{X}$ is compact, see Sect. A.4 of Appendix A, it will also follow that $\mathcal{G}_\Lambda$ is compact).

Let $\mu_n \in \mathcal{G}_\Lambda$, $n \geq 1$, and $\mu_n \to \mu$ weakly, i.e. for all continuous f ($f \in C(\mathcal{X})$), $\mu_n(f) \to \mu(f)$. We need to prove that $\mu \in \mathcal{G}_\Lambda$. Let $f \in C(\mathcal{X})$. By (2.1.6.5), $\mu_n(f) = \mu_n(b_f)$, b_f as in (2.1.6.3). We have proved in Lemma 2.1.6 that $b_f \in C(\mathcal{X})$, i.e. it is continuous, since we are supposing that $f \in C(\mathcal{X})$. By letting $n \to \infty$, $\mu(f) = \mu(b_f)$, which, recalling the expression (2.1.6.3) for b_f proves (2.1.6.2) for continuous f. On the other hand if the integrals of all continuous functions are equal, then the two measures are the same and (2.1.6.2) holds also for all bounded, measurable f, so that $\mu \in \mathcal{G}_\Lambda$ and the latter is weakly closed.

By (2.1.5.1), $G_\Lambda(f|\sigma_{\Lambda^c})$, f a bounded measurable function, is equal to

$$\sum_{\sigma'_{\Lambda\setminus\Delta}} \sum_{\sigma_\Delta} f(\sigma_\Delta, \sigma'_{\Lambda\setminus\Delta}, \sigma_{\Lambda^c}) G_\Delta(\sigma_\Delta|\sigma'_{\Lambda\setminus\Delta}, \sigma_{\Lambda^c}) (G_\Lambda(\cdot|\sigma_{\Lambda^c}))_{\Lambda\setminus\Delta}(\sigma'_{\Lambda\setminus\Delta}),$$

which can be rewritten as

$$\sum_{\sigma'\in\mathcal{X}} \left\{ \sum_{\sigma_\Delta} f(\sigma_\Delta, \sigma'_{\Lambda\setminus\Delta}, \sigma_{\Lambda^c}) G_\Delta(\sigma_\Delta|\sigma'_{\Lambda\setminus\Delta}, \sigma_{\Lambda^c}) \right\} G_\Lambda(\sigma'|\sigma_{\Lambda^c}),$$

so that $G_\Lambda(\cdot|\sigma_{\Lambda^c}) \in \mathcal{G}_\Delta$.

Let now $\mu \in \mathcal{G}_\Lambda$; then $\mu(f) = \int \{\sum_{\sigma'} G_\Lambda(\sigma'|\sigma_{\Lambda^c}) f(\sigma')\}\, \mu(d\sigma)$, hence

$$\mu(f) = \int \left\{ \sum_{\sigma'} G_\Lambda(\sigma'|\sigma_{\Lambda^c}) \sum_{\sigma''} G_\Delta(\sigma''|\sigma'_{\Delta^c}) f(\sigma'') \right\} \mu(d\sigma).$$

Then

$$\mu(f) = \int \sum_{\sigma''} f(\sigma'') G_\Delta(\sigma''|\sigma'_{\Delta^c})\, \nu(d\sigma'), \tag{2.1.7.1}$$

where ν is such that $\nu(g) = \int \sum_{\sigma'} G_\Lambda(\sigma'|\sigma_{\Lambda^c}) g(\sigma')\, \mu(d\sigma)$. Since ν is a probability, the last expression in (2.1.7.1) is by (2.1.6.2) in $\mathcal{G}_\Delta$ and the theorem is proved. □

2.2 Thermodynamic limit and DLR measures

We will derive a macroscopic theory from the Gibbs hypothesis by separating bulk from surface effects. Bulk properties are those which refer to the behavior of the spins [in our Ising model] which are far from the boundaries; if the region occupied by the system is not too weird, indeed most spins will be far from the boundaries. To make this quantitative, we must specify the meaning of "most" and "far from the boundaries." There is here an evident degree of arbitrariness, which has to be lifted if we want a mathematical theory with precise statements. Modern statistical mechanics defines the bulk properties as those which emerge in an infinite volume limit, which is usually referred to as "the thermodynamic limit." In this limit in fact the bulk thermodynamics of the system is singled out.

As we will see in the next two sections the notion is well defined for intensive thermodynamic potentials like pressure and free energy density, for which, under quite general assumptions on the interaction, the infinite volume limit exists and is essentially independent of the sequence of regions and boundary conditions used in the limit procedure. The situation is different when we look at the full equilibrium states, the object of our present analysis, where, as we will argue below, we cannot expect in general the existence of a limit independent of the sequence of regions and boundary conditions.

The modern rigorous theory of equilibrium states is founded on an hypothesis which avoids [or better, it seems to avoid] the thermodynamic limit procedure by extending the original Gibbs hypothesis to one formulated directly in infinite systems, the so called DLR condition. DLR stands for Dobrushin, Lanford and Ruelle, who are the founders of the theory. The condition translates the physically obvious notion that if a system is globally in equilibrium, then it is also locally in equilibrium. According to what we have argued so far, a state is in equilibrium in a bounded region Λ if it is in $\mathcal{G}_\Lambda$; thus DLR define the set of all equilibrium measures $\mathcal{G}$ as

$$\mathcal{G} := \bigcap_{\Lambda \text{ bounded in } \mathbb{Z}^d} \mathcal{G}_\Lambda. \tag{2.2.0.1}$$

The definition (2.2.0.1) immediately raises three questions: the existence of DLR measures, the meaning of possible non-uniqueness, and the way the notion is related to the thermodynamic limit procedure discussed earlier.

We shall prove that $\mathcal{G}$ is non-empty, weakly compact and convex. Convexity allows one to distinguish in $\mathcal{G}$ extremal elements and mixtures of extremal elements.

We will see that the extremal measures in $\mathcal{G}$ are obtained as infinite volume limits, $\Delta_n \to \mathbb{Z}^d$, of Gibbs measures $G_{\Delta_n}(\cdot|\sigma_{\Delta_n^c})$, where σ is a fixed configuration and $\sigma_{\Delta_n^c}$ is the restriction of σ to Δ_n^c. They will be interpreted as pure phases. The non-uniqueness of the limit then means that $\mathcal{G}$ has several pure phases (selected by the appropriate boundary conditions) and we have a phase transition. Thus phase transitions are related to a persistent diversity among Gibbs states in large domains Λ when the boundary conditions are varied: from such a perspective, phase transitions means "sensitive dependence" on the boundary conditions. As we vary the spins at the boundary, we cause a chain reaction which propagates inside Λ, affecting eventually all the spins; thus a volume effect is produced by a comparatively small surface change. The context when phase transitions occur must therefore be critical and phase transitions rare. Indeed, thermodynamics tells us that they occur on surfaces [of the phase diagram] with positive codimension; in this sense they are "exceptional." In the next section we will see that a formulation of this property (that unfortunately is very weak) is true in general.

Outline of the main results As already mentioned, the set $\mathcal{G}$ of DLR measures (at fixed inverse temperature β) is a non-empty, convex, weakly compact set, just as $\mathcal{G}_\Lambda$ and indeed $\mathcal{G}$ is structurally similar to $\mathcal{G}_\Lambda$, and we may in fact think of $\mathcal{G}$ as $\mathcal{G}_\Lambda$ with $\Lambda = \mathbb{Z}^d$, as we are going to argue. Recall that for finite Λ the extremal elements of $\mathcal{G}_\Lambda$ are obtained by taking any $\sigma \in \mathcal{X}$ and constructing the measure $G_\Lambda(\cdot|\sigma_{\Lambda^c})$. Any element in $\mathcal{G}_\Lambda$ can then be written as $\int G_\Lambda(\cdot|\sigma_{\Lambda^c})p(d\sigma)$, p a probability on $\mathcal{X}$; it is namely a convex combination of extremal states, and any such integral defines an element of $\mathcal{G}_\Lambda$. The extremal elements of $\mathcal{G}$ are obtained in a similar way. Fix arbitrarily an increasing sequence $\{\Delta_n\}$ of regions invading $\mathbb{Z}^d$; take any configuration σ, not as before in the whole $\mathcal{X}$, but only in a suitable set $\mathcal{X}_{gg}$ (which depends on $\{\Delta_n\}$, "gg" for very good); take as before the measure $G_{\Delta_n}(\cdot|\sigma_{\Delta_n^c})$ and let $\Delta_n \to \mathbb{Z}^d$. The weak limit, which is proved to exist in $\mathcal{X}_{gg}$, defines an extremal measure G_σ in $\mathcal{G}$. It is also true that any element in $\mathcal{G}$ is an integral $\int G_\sigma(\cdot)p(d\sigma)$ with p a probability with support on $\mathcal{X}_{gg}$.

To continue with the analogy between $\mathcal{G}_\Lambda$ and $\mathcal{G}$, observe that the measures $\{G_\Lambda(\cdot|\sigma_{\Lambda^c}),\, \sigma \in \mathcal{X}\}$ define a natural partition π_Λ of $\mathcal{X}$: σ' and σ'' are in the same atom of π_Λ if and only if $G_\Lambda(\sigma_\Lambda|\sigma'_{\Lambda^c}) = G_\Lambda(\sigma_\Lambda|\sigma''_{\Lambda^c})$ for all $\sigma_\Lambda \in \mathcal{X}_\Lambda$. If we change σ' only inside Λ, the new configuration σ'' is trivially in the same atom as σ', and the partition is thus called measurable on Λ^c. Analogously the measures $\{G_\sigma,\, \sigma \in \mathcal{X}_{gg}\}$ are extremal in $\mathcal{G}$ and define a partition π_∞ of $\mathcal{X}_{gg}$ by the equivalence relation $\sigma' \sim \sigma''$ if and only if $G_{\sigma'} = G_{\sigma''}$. The atoms Ω_σ are such that if $\sigma' \in \Omega_\sigma$ then any modification of σ in a bounded set gives rise to a new configuration which however is in the same Ω_σ, for this reason the partition π_∞ is said to be "measurable at infinity." Moreover, let Ω_σ, $\sigma \in \mathcal{X}_{gg}$ the atom of π_∞ containing σ, then $G_\sigma(\Omega_\sigma) = 1$, and distinct extremal measures have therefore disjoint support and the decomposition of an element of $\mathcal{G}$ as $\int G_\sigma(\cdot)p(d\sigma)$ is unique.

We will conclude the section by discussing the group of space translations and its action on $\mathcal{G}$; see Sects. 2.2.6 and 2.2.7. We shall introduce the notion of "ergodic DLR measures" and prove that any translation invariant DLR measure is an integral over ergodic DLR measures, an ergodic decomposition.

2.2.1 DLR measures

We fix hereafter $\beta > 0$ and drop it from the notation when no ambiguity may arise; all measures in the sequel are meant as Gibbs measures at the inverse temperature β. With such an understanding we define:

Definition 1 A probability measure μ on $\mathcal{X}$ is an *equilibrium measure* if μ belongs to $\mathcal{G}_\Lambda$ for any bounded Λ in $\mathbb{Z}^d$. The set of all equilibrium measures is denoted by $\mathcal{G}$ with $\mathcal{G}$ given in (2.2.0.1).

Theorem 2.2.1.1 *The set $\mathcal{G}$ is non-empty, convex and weakly compact. Moreover, if Δ_n is any sequence of increasing regions which invades $\mathbb{Z}^d$, $\mathcal{G}_{\Delta_n}$ is non-increasing and*

$$\mathcal{G} = \bigcap_{\Delta_n} \mathcal{G}_{\Delta_n}. \tag{2.2.1.1}$$

Proof (2.2.1.1) follows from (2.2.0.1) because $\mathcal{G}_\Lambda \subset \mathcal{G}_\Delta$ if $\Delta \subset \Lambda$ (by Theorem 2.1.7.1). The l.h.s. of (2.2.1.1) is a non-empty, convex, weakly compact set, because all $\mathcal{G}_{\Delta_n}$ are convex and weakly compact sets (again by Theorem 2.1.7.1), each one containing the successive one. □

Thus Definition 1 is non-empty and equilibrium measures do indeed exist. The elements of $\mathcal{G}$ have a nice interpretation in terms of conditional probabilities which, as we will see, is the key to their analysis. The definition and properties of conditional probabilities are given in Appendix A, see Sect. A.6; all proofs hereafter strongly depend on Appendix A. Call $\mathcal{B}_{\Lambda^c}$, Λ a finite subset of $\mathbb{Z}^d$, the minimal σ algebra which contains all the cylinders in Λ^c.

Definition 2 [DLR measures] A probability measure μ on $\mathcal{X}$ is called *DLR* if for any bounded $\Lambda \subset \mathbb{Z}^d$, $\{\mathcal{X}, G_\Lambda(\cdot|\sigma_{\Lambda^c})\}$ is a version of the conditional probability of μ given $\mathcal{B}_{\Lambda^c}$; see Sect. A.6.

Theorem 2.2.1.2 *μ is DLR if and only if $\mu \in \mathcal{G}$.*

Proof Suppose $\mu \in \mathcal{G}$ and let Λ be a bounded set. Then $\mu \in \mathcal{G}_\Lambda$ and by (2.1.6.5) with $f = \mathbf{1}_{A\cap B}$, $A \in \mathcal{B}$, $B \in \mathcal{B}_{\Lambda^c}$,

$$\mu(A \cap B) = \int_{\mathcal{X}} \sum_{\sigma'} G_\Lambda\big(\sigma'|\sigma_{\Lambda^c}\big)\mathbf{1}_{\sigma'\in A}\mathbf{1}_{\sigma'\in B}\,\mu(d\sigma).$$

Since $B \in \mathcal{B}_{\Lambda^c}$ and $G_\Lambda(\sigma'|\sigma_{\Lambda^c}) = 0$ unless $\sigma' = \sigma$ on Λ^c, $\mathbf{1}_{\sigma'\in B} = \mathbf{1}_{\sigma\in B}$,

$$\mu(A \cap B) = \int_B \Big\{\sum_{\sigma'} G_\Lambda\big(\sigma'|\sigma_{\Lambda^c}\big)\mathbf{1}_{\sigma'\in A}\Big\}\,\mu(d\sigma) = \int_B G_\Lambda\big(A|\sigma_{\Lambda^c}\big)\,\mu(d\sigma),$$

which, by (A.6.1), proves that $\{\mathcal{X}, G_\Lambda(\cdot|\sigma_{\Lambda^c})\}$ is a version of the conditional probability of μ given $\mathcal{B}_{\Lambda^c}$. By the arbitrariness of Λ, μ is DLR.

Vice versa, suppose that μ is DLR. Then by (A.6.1),

$$\mu(A) = \int_{\mathcal{X}} \sum_{\sigma'} G_\Lambda\big(\sigma'|\sigma_{\Lambda^c}\big)\mathbf{1}_{\sigma'\in A}\, \mu(d\sigma),$$

which proves (2.1.6.3) for functions f which are characteristic functions of Borel sets. By a density argument (details are omitted) the equality extends to all bounded, Borel measurable functions. Thus $\mu \in \mathcal{G}_\Lambda$ and by the arbitrariness of Λ, $\mu \in \mathcal{G}$. $\square$

2.2.2 Thermodynamic limits of Gibbs measures

In this subsection we will establish a first relation between DLR measures and thermodynamic limits of finite volume Gibbs measures. Let $\{\Delta_n\}$ be an increasing sequence of finite regions which invades the whole space; the construction below will depend on the choice of $\{\Delta_n\}$, but, as we shall see, the final conclusions about the elements of $\mathcal{G}$ are structural and independent of $\{\Delta_n\}$. Recalling from Sect. 2.1.3 that $\mathcal{C}$ denotes the family of all elementary cylindrical sets, we introduce "the good set"

$$\mathcal{X}_g = \Big\{\sigma \in \mathcal{X} : \lim_{n\to\infty} G_{\Delta_n}\big(C|\sigma_{\Delta_n^c}\big) \text{ exists for all } C \in \mathcal{C}\Big\}$$

(we shall later introduce a very good set $\mathcal{X}_{gg}$). A priori $\mathcal{X}_g$ may be empty, but the following theorem excludes such a possibility:

Theorem 2.2.2.1 *$\mathcal{X}_g$ is a non-empty Borel set and $\mu(\mathcal{X}_g) = 1$ for any $\mu \in \mathcal{G}$. Moreover, for any $\sigma \in \mathcal{X}_g$ there is a unique measure $G_\sigma(\cdot)$ such that*

$$G_\sigma(f) = \lim_{n\to\infty} G_{\Delta_n}\big(f|\sigma_{\Delta_n^c}\big) \quad \textit{for all continuous functions } f. \tag{2.2.2.1}$$

All G_σ, $\sigma \in \mathcal{X}_g$, are in $\mathcal{G}$.

Proof Let $\mu \in \mathcal{G}$; recall that we have proved in Theorem 2.2.1.1 that $\mathcal{G} \neq \emptyset$. Then μ is DLR and the pair $\{\mathcal{X}, G_{\Delta_n}(\cdot|\sigma_{\Delta_n^c})\}$ is a version of the conditional probability of μ given the σ-algebra $\{\mathcal{B}_{\Delta_n^c}\}$. We can then apply Theorem A.11 with $\Sigma_n = \mathcal{B}_{\Delta_n^c}$, $\mathcal{X}_n = \mathcal{X}$ and $\mu(\cdot|\Sigma_n)(\sigma) = G_{\Delta_n}(\cdot|\sigma_{\Delta_n^c})$. Then the set $\mathcal{X}'$ in (A.10.1) is our set $\mathcal{X}_g$; hence by Theorem A.11, $\mathcal{X}_g \in \mathcal{B}$ and $\mu(\mathcal{X}_g) = 1$ (thus $\mathcal{X}_g \neq \emptyset$). Theorem A.11 also states that for any $\sigma \in \mathcal{X}_g$ there is a unique probability $\mu(\cdot|\mathcal{B}_\infty)(\sigma)$, $\mu(\cdot|\Sigma)(\sigma)$ in the notation of Theorem A.11 (that we identify with G_σ of (2.2.2.1)), such that

$$\lim_{n\to\infty} G_{\Delta_n}\big(C|\sigma_{\Delta_n^c}\big) = \mu\big(C|\mathcal{B}_\infty\big)(\sigma).$$

By Theorem A.4 it follows that, for any $\sigma \in \mathcal{X}_g$, $G_{\Delta_n}(f|\sigma_{\Delta_n^c}) = \mu(f|\mathcal{B}_\infty)(\sigma)$ for all continuous functions.

Finally, the statement $G_\sigma \in \mathcal{G}$ is an immediate consequence of (2.2.2.1), as the latter states that G_σ is the weak limit of a sequence which is definitively in $\mathcal{G}_\Lambda$ (as soon as n is such that $\Lambda \subset \Delta_n$). Since $\mathcal{G}_\Lambda$ is weakly closed, $G_\sigma \in \mathcal{G}_\Lambda$, and by the arbitrariness of Λ it is in the intersection of all $\mathcal{G}_\Lambda$; hence it is in $\mathcal{G}$. □

The set $\mathcal{X}_g$ is good but not very good! For any $\sigma \in \mathcal{X}_g$, call

$$\Omega_\sigma = \left\{\sigma' \in \mathcal{X}_g : G_{\sigma'} = G_\sigma\right\}, \qquad \mathcal{X}_{gg} = \left\{\sigma \in \mathcal{X}_g : G_\sigma(\Omega_\sigma) = 1\right\}. \tag{2.2.2.2}$$

$\mathcal{X}_{gg}$ is the "very good" set we are looking for and which has been described in the beginning of the section, in the paragraph "Outline of main results."

Theorem 2.2.2.2 *$\mathcal{X}_{gg}$ is a non-empty Borel set and for any $\sigma \in \mathcal{X}_g$ either $\Omega_\sigma \in \mathcal{X}_{gg}$ or $\Omega_\sigma \cap \mathcal{X}_{gg} = \emptyset$. For any $\mu \in \mathcal{G}$, $\mu(\mathcal{X}_{gg}) = 1$ and for any bounded measurable function f, $G_\sigma(f)$, $\sigma \in \mathcal{X}_{gg}$, is a measurable function and*

$$\mu(f) = \int_{\mathcal{X}_{gg}} G_\sigma(f)\,\mu(d\sigma). \tag{2.2.2.3}$$

Conversely, for any probability ν on $\mathcal{X}_{gg}$, the measure μ defined by

$$\mu(f) = \int_{\mathcal{X}_{gg}} G_\sigma(f)\,\nu(d\sigma) \tag{2.2.2.4}$$

is in $\mathcal{G}$.

Proof It immediately follows from the definition (2.2.2.2) that for any $\sigma \in \mathcal{X}_g$ either $\Omega_\sigma \in \mathcal{X}_{gg}$ or $\Omega_\sigma \cap \mathcal{X}_{gg} = \emptyset$. Suppose $\mu \in \mathcal{G}$. By Theorem A.11 and the identification $G_\sigma = \mu(\cdot|\mathcal{B}_\infty)(\sigma)$, the pair $(\mathcal{X}_g, G_\sigma)$ is a version of the conditional probability of μ with respect to the σ-algebra $\mathcal{B}_\infty$, which is defined as the minimal σ-algebra which contains the sets $B \in \mathcal{B}$ which are in $\mathcal{B}_{\Delta_n^c}$ for all n (for this reason we will also write $G_\sigma = \mu(\cdot|\mathcal{B}_\infty)(\sigma)$, $\sigma \in \mathcal{X}_g$). By Theorem A.8 $(\mathcal{X}_{gg}, G_\sigma)$ is also a version of the conditional probability given $\mathcal{B}_\infty$, hence $\mathcal{X}_{gg}$ is a Borel set, $\mu(\mathcal{X}_{gg}) = 1$ and (2.2.2.3) holds. Since $\mu(\mathcal{X}_{gg}) = 1$, $\mathcal{X}_{gg}$ is non-empty. Equation (2.2.2.3) then follows from $(\mathcal{X}_{gg}, G_\sigma)$ being a conditional probability.

Let μ be as in (2.2.2.4). Since $G_\sigma \in \mathcal{G}$, $G_\sigma \in \mathcal{G}_\Lambda$, Λ bounded,

$$\begin{aligned}\mu(f) &= \int_{\mathcal{X}_{gg}} \left\{\int_{\mathcal{X}} G_\Lambda(f|\sigma'_{\Lambda^c}) G_\sigma(d\sigma')\right\} \nu(d\sigma) \\ &= \int_{\mathcal{X}} G_\Lambda(f|\sigma'_{\Lambda^c})\mu(d\sigma').\end{aligned} \tag{2.2.2.5}$$

Equation (2.2.2.5) shows that $\mu \in \mathcal{G}_\Lambda$ and by the arbitrariness of Λ that $\mu \in \mathcal{G}$. □

We will next improve the above analysis by establishing fine and detailed properties of the DLR measures. The analysis is more technical and heavily relies on the theory of conditional probabilities in Appendix A. In a first reading one may go directly to Sect. 2.2.6 (however, the proof of Theorem 2.2.7.2 will use some of the following results).

The sets $\mathcal{X}_g$ and $\mathcal{X}_{gg}$ depend on the choice of the sequence $\{\Delta_n\}$ used in the definition of G_σ. If we take another sequence $\{\Delta'_n\}$ we will have in general new $\mathcal{X}'_g$ and $\mathcal{X}'_{gg}$ and new measures $G'_{\sigma'}, \sigma' \in \mathcal{X}'_g$. The two families are however strictly related to each other.

Theorem 2.2.2.3 *The set $\mathcal{X}_0 := \mathcal{X}_{gg} \cap \mathcal{X}'_{gg}$ is non-empty, $\Omega_\sigma \cap \mathcal{X}_0 \neq \emptyset$ for all $\sigma \in \mathcal{X}_{gg}$ and $\Omega'_\sigma \cap \mathcal{X}_0 \neq \emptyset$ for all $\sigma \in \mathcal{X}'_{gg}$. Moreover, $\mu(\mathcal{X}_0) = 1$ for any $\mu \in \mathcal{G}$; for any $\sigma \in \mathcal{X}_0$, $G_\sigma = G'_\sigma$, $\Omega_\sigma \cap \mathcal{X}_0 = \Omega'_\sigma \cap \mathcal{X}_0$ and $G_\sigma(\Omega_\sigma \cap \mathcal{X}_0) = 1$.*

Proof By (2.2.2.3) with $\mu = G_\sigma$ and $f = \mathbf{1}_{\Omega_\sigma}$,

$$G_\sigma(\Omega_\sigma) = \int_{\mathcal{X}'_{gg}} G'_{\sigma'}(\Omega_\sigma)\, G_\sigma(d\sigma').$$

Since $G_\sigma(\Omega_\sigma) = 1$ and $G'_{\sigma'}(\Omega_\sigma) \leq 1$,

$$G_\sigma\big(\{\sigma' \in \Omega_\sigma \cap \mathcal{X}'_{gg} : G'_{\sigma'}(\Omega_\sigma) = 1\}\big) = 1.$$

Then $\Omega_\sigma \cap \mathcal{X}'_{gg} \neq \emptyset$ and if $\sigma' \in \Omega_\sigma \cap \mathcal{X}'_{gg}$, $G_\sigma = G'_{\sigma'}$. If $\Omega'_{\sigma'} \cap \Omega'_{\sigma''} = \emptyset$, then $G'_{\sigma'} \neq G'_{\sigma''}$; hence $G_\sigma \neq G'_{\sigma''}$ so that $\Omega_\sigma \cap \Omega'_{\sigma''} = \emptyset$. □

2.2.3 Pure states and extremal DLR measures

$\mathcal{G}$ is a convex set and, being weakly compact, its extremal points, whose collection is denoted by $\mathcal{G}_{\text{extr}}$, are also in $\mathcal{G}$. By the Krein–Millman theorem, see I.3.10 in the book by Naimark [175], elements in $\mathcal{G} \setminus \mathcal{G}_{\text{extr}}$ are convex combinations (in general integrals) over the extremal elements and are therefore called "mixture states," while the extremal elements are "pure states." We will see at the end of this subsection that pure states can be identified as the pure phases of the system. In the next theorem we will characterize the extremal DLR measures as the measures $G_\sigma, \sigma \in \mathcal{X}_{gg}$; then the decomposition into extremal measures is just (2.2.2.3).

Theorem 2.2.3.1 *The following two statements are equivalent*:

- $\mu \in \mathcal{G}$ *and* $\mu(\Omega_\sigma) = 1$ *for some* $\sigma \in \mathcal{X}_{gg}$.
- $\mu = G_\sigma, \sigma \in \mathcal{X}_{gg}$. (2.2.3.1)

The set $\mathcal{G}_{\text{extr}}$ of extremal elements of $\mathcal{G}$ is $\mathcal{G}_{\text{extr}} = \{G_\sigma, \sigma \in \mathcal{X}_{gg}\}$ so that (2.2.2.3) is a decomposition of μ into extremal states.

Proof If $\mu = G_\sigma$ for some $\sigma \in \mathcal{X}_{gg}$, then $\mu \in \mathcal{G}$ and $\mu(\Omega_\sigma) = 1$ by the definition of $\mathcal{X}_{gg}$. Suppose conversely that $\mu \in \mathcal{G}$ and $\mu(\Omega_\sigma) = 1$, then by (2.2.2.3) $\mu(f) = \int_{\Omega_\sigma} G_{\sigma'}(f)\, \mu(d\sigma') = G_\sigma(f) \int_{\Omega_\sigma} \mu(d\sigma') = G_\sigma(f)$, hence (2.2.3.1).

Let us next prove the statements about $\mathcal{G}_{\text{extr}}$. Suppose first that $\mu \in \mathcal{G}_{\text{extr}}$. Arguing by contradiction, we will show that if $\mu \neq G_\sigma(\cdot)$ for all $\sigma \in \mathcal{X}_{gg}$, then the integral decomposition in (2.2.2.3) can be reduced to a convex combination of two distinct measures in $\mathcal{G}$. We first observe that $\sigma \to G_\sigma$ on $\mathcal{X}_{gg}$ cannot be μ-a.s. constant, otherwise, by (2.2.2.3), it would be equal to μ. Then there must be f and $b \in \mathbb{R}$ such that

$$\mu\left(\left\{\sigma \in \mathcal{X}_{gg} : G_\sigma(f) \le b\right\}\right) = \alpha, \quad \alpha \neq 0, 1. \tag{2.2.3.2}$$

Calling B the set in curly brackets, we define the probabilities

$$\begin{aligned} \mu'(\cdot) &= \alpha^{-1} \int_B G_\sigma(\cdot)\, \mu(d\sigma), \\ \mu''(\cdot) &= (1-\alpha)^{-1} \int_{B^c} G_\sigma(\cdot)\, \mu(d\sigma). \end{aligned} \tag{2.2.3.3}$$

By (2.2.2.4), both μ' and μ'' are in $\mathcal{G}$; moreover, $\mu' \neq \mu''$ because, by construction, $\mu'(f) \le b$ and $\mu''(f) > b$. Also by construction $\mu = \alpha\mu' + (1-\alpha)\mu''$, hence the desired contradiction, which proves that $\mu = G_\sigma$ for some $\sigma \in \mathcal{X}_{gg}$.

Conversely, suppose that for $\sigma \in \mathcal{X}_{gg}$, $G_\sigma \notin \mathcal{G}_{\text{extr}}$. Then there are $\alpha \in [0,1]$, μ' and μ'' in $\mathcal{G}$ for which $G_\sigma = \alpha\mu' + (1-\alpha)\mu''$. By applying (2.2.2.3) to $\mu = G_\sigma$, we get, for any measurable set A,

$$G_\sigma(A) = \int_{\mathcal{X}_{gg}} G_{\sigma'}(A)\, [\alpha\mu' + (1-\alpha)\mu''](d\sigma'). \tag{2.2.3.4}$$

Since for all $\sigma' \in \mathcal{X}_{gg}$, $G_{\sigma'}(\Omega_\sigma) = \mathbf{1}_{\sigma' \in \Omega_\sigma}$, (2.2.3.4) with $A = \Omega_\sigma$ yields

$$\alpha\mu'(\Omega_\sigma) + (1-\alpha)\mu''(\Omega_\sigma) = 1.$$

If $\alpha \in (0,1)$, this implies $\mu'(\Omega_\sigma) = \mu''(\Omega_\sigma) = 1$, so that, by (2.2.3.1) which has already been proved, $\mu' = \mu'' = G_\sigma$. □

We will next argue that extremal measures have the physical interpretation of "pure phases." The idea is to start from the particular case that $\mathcal{G}$ is a singleton, in which case there is no doubt that its only element is a pure phase. We will prove for the singleton properties that in the general case are satisfied by the extremal measures of $\mathcal{G}$, thus interpreting the latter as pure phases.

Theorem 2.2.3.2 *If $\mathcal{G} = \{\mu\}$ is a singleton, then $\mathcal{X}_g = \mathcal{X}_{gg} = \mathcal{X}$, and, given any increasing sequence $\Lambda_n \to \mathbb{Z}^d$, for any continuous function f*

$$\begin{aligned} &\lim_{\Lambda_n \to \mathbb{Z}^d} G_{\Lambda_n}(f|\sigma^c_{\Lambda_n}) = \mu(f), \quad \sigma \in \mathcal{X}; \\ &\lim_{\Lambda_n \to \mathbb{Z}^d} \int |G_{\Lambda_n}(f|\sigma^c_{\Lambda_n}) - \mu(f)|\mu(d\sigma) = 0. \end{aligned} \tag{2.2.3.5}$$

Proof By compactness $G_{\Lambda_n}(\cdot|\sigma^c_{\Lambda_n})$ converges weakly by subsequences, and since $\mathcal{G}_\Delta$ is weakly compact any limit point is in $\mathcal{G}_\Delta$ and, by the arbitrariness of Δ it is in $\mathcal{G}$. Since $\mathcal{G}$ is a singleton the limit point is μ and the sequence actually converges. We have thus proved the first limit in (2.2.3.6); the second one follows from the first one by the Lebesgue dominated convergence theorem. □

If $\mathcal{G}$ is not a singleton, we have

Theorem 2.2.3.3 *If $\mu \in \mathcal{G}_{\text{extr}}$, given any increasing sequence $\Lambda_n \to \mathbb{Z}^d$ then for any continuous function f, $\mu(\{\sigma : \lim_{\Lambda_n \to \mathbb{Z}^d} G_{\Lambda_n}(f|\sigma^c_{\Lambda_n}) = \mu(f)\}) = 1$ and*

$$\lim_{\Lambda_n \to \mathbb{Z}^d} \int |G_{\Lambda_n}(f|\sigma^c_{\Lambda_n}) - \mu(f)|\mu(d\sigma) = 0. \tag{2.2.3.6}$$

Proof Let $\mathcal{X}_{gg}$ be the set associated to the sequence Λ_n. Then by Theorem 2.2.3.1 there is $\sigma \in \mathcal{X}_{gg}$ such that $\mu = G_\sigma$; hence

$$\mu\Big(\Big\{\sigma : \lim_{\Lambda_n \to \mathbb{Z}^d} G_{\Lambda_n}(f|\sigma^c_{\Lambda_n}) = \mu(f)\Big\}\Big) \geq G_\sigma(\Omega_\sigma) = 1,$$

which proves the first statement in the theorem. Actually by taking countably many intersections, the set where there is convergence can be chosen once for all f. The limit in (2.2.3.6) then follows by the Lebesgue dominated convergence theorem. □

Extremal measures as pure phases

By comparing Theorems 2.2.3.2 and 2.2.3.3 we see that the extremal measures enjoy the same properties as the unique measure when $\mathcal{G}$ is a singleton, provided we replace “everywhere” by “almost everywhere” (i.e. with probability 1). That is, if we view the phase space from the viewpoint of an extremal measure μ then we see the same homogeneous behavior as when $\mathcal{G}$ is a singleton. For such reasons we will call the states in $\mathcal{G}_{\text{extr}}$ pure phases.

To exemplify the discussion let us refer to the $d = 2$ ferromagnetic Ising model with n.n. interactions and no magnetic field. We will prove in the next chapter that at small temperatures there exists an extremal measure, the so called plus measure, where with probability 1 all configurations have a positive magnetization. There is also another extremal measure, the minus measure, where with probability 1 all configurations have negative magnetization. There are then mixture states where in a fraction of configurations we observe a positive and in the other one a negative magnetization. Mixtures in this example are convex combinations of the two pure states.

2.2.4 Simplicial structure of DLR measures

The are two opposite types of convex sets, as far as the decomposition into their extremal points is concerned, which in the plane are visualized as circles and triangles.

In the former the decomposition into extremal elements is highly non-unique, just the opposite of what happens in the latter. $\mathcal{G}$ belongs to the latter category; we are going to show that the decomposition of an element $\mu \in \mathcal{G}$ into extremal DLR states is unique.

We want to write any $\mu \in \mathcal{G}$ as an integral over $\mathcal{G}_{\text{extr}}$. Thus the first step requires us to introduce a measurable space (Ω, Σ), i.e. a space Ω and a σ-algebra Σ, with Ω in one to one correspondence with $\mathcal{G}_{\text{extr}}$. By (2.2.3.1) we can take for Ω the space whose points ω are the sets Ω_σ, $\sigma \in \mathcal{X}_{gg}$. Call ϕ the map from $\mathcal{X}_{gg}$ onto Ω, which associates to any $\sigma \in \mathcal{X}_{gg}$ the element $\Omega_\sigma \in \Omega$. We then define Σ as the σ-algebra made of all sets $A \subset \Omega$ such that $\phi^{-1}(A)$ is measurable in $\mathcal{X}$. We also define for any $\mu \in \mathcal{G}$ a measure p_μ on (Ω, Σ) by setting $p_\mu(A) = \mu(\phi^{-1}(A))$, $A \in \Sigma$, so that (Ω, Σ, p_μ) is isomorphic to the restriction of μ to the σ-algebra $\phi^{-1}(\Sigma)$.

Theorem 2.2.4.1 *Let $\mu \in \mathcal{G}$; then there is a unique measure p on (Ω, Σ) such that for any bounded measurable function f on $\mathcal{X}$,*

$$\mu(f) = \int_\Omega G_\omega(f)\, p(d\omega),$$

where, by abuse of notation, $G_\omega \equiv G_\sigma$, $\sigma \in \phi^{-1}(\omega)$; recall that G_σ is the same for all $\sigma \in \phi^{-1}(\omega)$. The measure p is equal to p_μ.

Proof By Theorem 2.2.2.2, $\sigma \to G_\sigma(f)$ is measurable and constant on each Ω_σ; hence it is measurable on the σ-algebra $\phi^{-1}(\Sigma)$. Calling μ' the restriction of μ to $\phi^{-1}(\Sigma)$, we thus conclude from (2.2.2.3) that

$$\mu(f) = \int_{\mathcal{X}_{gg}} G_\sigma(f)\, \mu'(d\sigma) = \int_\Omega G_\omega(f)\, p_\mu(d\omega).$$

To prove uniqueness, let (Ω, Σ, q) be a probability space and

$$\mu(f) = \int_\Omega G_\omega(f)\, q(d\omega).$$

We claim that $q = p_\mu$. For $B \in \Sigma$ we have

$$\mu(\phi^{-1}(B)) = \int G_\omega(\phi^{-1}(B))\, q(d\omega) = \int_B G_\omega(\phi^{-1}(B))\, p_\mu(d\omega). \tag{2.2.4.1}$$

Since $G_{\phi^{-1}(\omega)}(\phi^{-1}(B)) = \mathbf{1}_{\omega \in B}$, (2.2.4.1) yields $q(B) = p_\mu(B)$. □

2.2.5 *Sigma algebra at infinity*

In the course of the proof of Theorem 2.2.2 we have introduced the σ-algebra $\mathcal{B}_\infty$ whose definition is made explicit here.

Definition The σ-algebra $\mathcal{B}_\infty = \bigcap_{\Lambda \text{ bounded}} \mathcal{B}_{\Lambda^c}$ is called the *σ-algebra at infinity*, or *tail field*.

Sets in $\mathcal{B}_\infty$ have the following property: if $A \in \mathcal{B}_\infty$ and $\sigma \in A$, then any local modification σ' of σ is also in A. Given a measure μ, a set A is μ-modulo 0 in $\mathcal{B}_\infty$ if there is a set $N \in \mathcal{B}_\infty$ such that $\mu(N) = 0$ and $A \cap N^c \in \mathcal{B}_\infty$. Functions which are μ-modulo 0 measurable at infinity are defined analogously.

The next theorem shows that measure theoretically the sets Ω_σ are the smallest ones among those measurable at infinity. Indeed, we shall see that if $\mu \in \mathcal{G}$, then any function f which is μ-modulo 0 measurable at infinity is μ almost surely constant on each set Ω_σ.

Theorem 2.2.5.1 *If $\mu \in \mathcal{G}$ and f is μ-almost surely measurable at infinity, then*

$$\mu\big(\{\sigma \in \mathcal{X}_{gg} : f(\sigma') = G_\sigma(f),\ \textit{for } G_\sigma\textit{-almost all } \sigma' \in \Omega_\sigma\}\big) = 1. \tag{2.2.5.1}$$

In particular for any $\mu \in \mathcal{G}_{\text{extr}}$ if A is μ-modulo 0 *in $\mathcal{B}_\infty$, then $\mu(A)$ is either* 0 *or* 1*, namely $\mathcal{B}_\infty = \{\emptyset, \mathcal{X}\}$, μ-modulo* 0.

Proof By assumption there is $N \in \mathcal{B}_\infty$ so that $\mu(N) = 0$ and $f((\sigma'_\Lambda, \sigma_{\Lambda^c})) = f((\sigma_\Lambda, \sigma_{\Lambda^c}))$, for all bounded Λ, all $\sigma \notin N$ and all σ'_Λ. By (2.2.2.3), $\mu(\{\sigma : G_\sigma(N) = 0\} = 1$, and (2.2.5.1) will be proved by showing that if $\sigma : G_\sigma(N) = 0$, then $f(\sigma') = G_\sigma(f)$ for G_σ-almost all $\sigma' \in \Omega_\sigma$.

Let $\sigma : G_\sigma(N) = 0$. Arguing by contradiction, similarly to the proof of Theorem 2.2.3 (see (2.2.3.2)–(2.2.3.4)), suppose that there are $\alpha \neq 0, 1$ and b such that $G_\sigma(B) = \alpha$, $B := \{\sigma' \in \Omega_\sigma \sqcap N^c : f(\sigma') \le b\}$. Let

$$\mu' = \alpha^{-1} \mathbf{1}_B G_\sigma, \qquad \mu'' = (1-\alpha)^{-1} \mathbf{1}_{B^c} G_\sigma,$$

so that $G_\sigma = \alpha\mu' + (1-\alpha)\mu''$ and $\mu' \neq \mu''$. The contradiction will arise once we show that μ' and μ'' are in $\mathcal{G}$, because by Theorem 2.2.3, $G_\sigma \in \mathcal{G}_{\text{extr}}$. Let us prove that $\mu' \in \mathcal{G}_\Lambda$, Λ bounded, namely that (2.1.6.5) holds with $\mu \to \mu'$. We have for any bounded measurable function h,

$$\begin{aligned}
\int G_\Lambda(h|\sigma' r_{\Lambda^c})\, \mu'(d\sigma') &= \alpha^{-1} \int_{\Omega_\sigma \cap N^c} \mathbf{1}_{f(\sigma') \le b} G_\Lambda(h|\sigma'_{\Lambda^c})\, G_\sigma(d\sigma') \\
&= \alpha^{-1} \int_{\Omega_\sigma \cap N^c} G_\Lambda(h \mathbf{1}_{f \le b}|\sigma'_{\Lambda^c})\, G_\sigma(d\sigma') \\
&= \alpha^{-1} \int_{\Omega_\sigma \cap N^c} h \mathbf{1}_{f \le b}\, G_\sigma(d\sigma') = \mu'(h),
\end{aligned}$$

where we have used that $\{f(\sigma') \le b\}$ does not depend on σ_Λ, and it is therefore a constant with respect to the measure $G_\Lambda(\cdot|\sigma'_{\Lambda^c}())$. Thus $\mu' \in \mathcal{G}_\Lambda$ and by the arbitrariness of Λ, $\mu' \in \mathcal{G}$; the same proof shows that $\mu'' \in \mathcal{G}$ hence the desired contradiction. $\square$

2.2.6 Pure phases, phase transitions

Here we discuss another aspect of phase transitions, related to the notion of intensive variables, which will lead to a different definition of pure phases, alternative to the one used so far, where pure phases have been associated to extremal DLR measures. In the new notion pure phases still correspond to extremal DLR measures but with extremality referring to the set of translational invariant DLR measures. The new notion is related to that of intensive variables. From a physical point of view in fact the intensive variables are expected to have sharp values in a pure phase, while in mixtures they fluctuate, as their values will depend on which pure state of the mixture the system is actually in. The notion of an intensive variable is related to space translations, which have been so far ignored, and which are going to enter strongly in the game.

We start with some simple basic considerations. Observe that, while the probability of a spin configuration with energy E is just given by the Gibbs formula (and it is therefore proportional to $e^{-\beta E}$), the probability that the system has an energy E is instead not at all as simple. In fact, counting the number of states with a given energy is in general a very complex task. Let us denote their number by $e^{S_\Lambda(E)}$, Λ the bounded domain where the system is confined. $S_\Lambda(E)$ is thus an entropy, the entropy for the given values of Λ and E; see Sect. 2.3 where the notion will become central. The energy distribution is then proportional to $\exp\{-\beta E + S_\Lambda(E)\}$. Both energy and entropy are extensive quantities; the corresponding intensive quantities are $e = E/|\Lambda|$ and $s_\Lambda(e) = S_\Lambda(E)/|\Lambda|$, respectively the energy and the entropy densities. We thus get for the energy distribution $\exp\{-|\Lambda|(\beta e - s_\Lambda(e))\}$. Supposing that $s_\Lambda(e)$ has the limit $s(e)$ as $\Lambda \to \mathbb{Z}^d$, we may then conclude that the energy distribution is concentrated for large $|\Lambda|$ around the minimizers of the free energy $e - \beta^{-1}s(e)$. When there is a unique minimizer, the distribution is uni-modal and most of the mass of the distribution is in its neighborhood: the fluctuations of the energy density are thus small and disappear in the thermodynamic limit. On the other hand, when the minimizer is not unique (the distribution is then called multi-modal), the energy density has macroscopic fluctuations which survive in the limit and the system exhibits a phase transition. But phase transitions may also come from the loss of uni-modality of other intensive variables rather than the energy, as in the classical Ising model, where the relevant order parameter is the magnetization density.

As we shall see, the theory of Gibbs measures encodes the above ideas in an elegant formulation which involves the action of the group of space translations on $\mathcal{G}$. Referring to Appendix A for our notation, for any $i \in \mathbb{Z}^d$ we denote by τ_i the map on $\mathbb{Z}^d$ defined by $\tau_i(x) = x + i$ and by $\tau_i(\sigma)$, $\tau_i(f)$ and $\tau_i(\mu)$ its dual actions, respectively, on spin configurations, on functions of the spin configurations and on probabilities on the space of spin configurations. Explicitly $\tau_i(\sigma)(x) = \sigma(x - i)$, $\tau_i(f)(\sigma) = f(\tau_{-i}(\sigma))$, $\tau_i(\mu)(f) = \mu(\tau_{-i}(f))$.

Theorem 2.2.6.1 *The set of DLR measure is invariant under translations,*

$$\tau_i(\mathcal{G}) = \mathcal{G}, \quad \textit{for any } i \in \mathbb{Z}^d, \tag{2.2.6.1}$$

and its subset $\mathcal{G}^0$ of translational invariant measures is a non-empty, convex, compact set.

Proof For any bounded Δ, $\tau_i(\mathcal{G}_\Delta) = \mathcal{G}_{\tau_i(\Delta)}$, see Appendix A after (A.1.10), hence (2.2.6.1). Let Δ_n be an increasing sequence of cubes which invades $\mathbb{Z}^d$, $\mu \in \mathcal{G}$ and

$$\mu^{(n)} = \frac{1}{|\Delta_n|} \sum_{i \in \Delta_n} \tau_i(\mu). \tag{2.2.6.2}$$

For all this, $\mu^{(n)} \in \mathcal{G}$ and let ν be a weak limit point of $\{\mu^{(n)}\}$, whose existence follows from compactness. $\nu = \tau_i(\nu)$ as ν is the limit of Cesaro's averages, and $\nu \in \mathcal{G}$ because $\mathcal{G}$ is weakly closed; thus $\mathcal{G}^0 \neq \emptyset$. Convexity and compactness also easily follow. $\square$

2.2.7 Ergodic decomposition

In the first part of this subsection we will recall the ergodic decomposition of a translational invariant measure μ into its ergodic components, and in the second part we shall use it to determine the structure of the translational invariant DLR measures. Recall that a translational invariant measure μ is ergodic if and only if $\mu(A)$ is either 0 or 1 for any measurable, translational invariant set A.

The ergodic decomposition consists in writing μ as an integral over measures which are translational invariant and ergodic. We state the results here without proofs, as they are well known from the literature; however, for the sake of completeness, a proof is given in Sect. A.9 of Appendix A.

Let Δ_n be an increasing sequence of cubes which invades $\mathbb{Z}^d$. Call

$$A^{(n)} f(\sigma) = \frac{1}{|\Delta_n|} \sum_{i \in \Delta_n} \tau_i f(\sigma),$$

and denote by $\mathcal{C}$ the set of all elementary cylinders; finally, define, in analogy with (2.2.2.2),

$$\mathcal{X}_g^0 = \Big\{\sigma : \lim_{n\to\infty} A^{(n)} \mathbf{1}_C(\sigma) \text{ exists for all } C \in \mathcal{C}\Big\}.$$

Analogously to Theorem 2.2.2.1, $\mathcal{X}_g^0$ is non-empty and for any $\sigma \in \mathcal{X}_g^0$ there is a translational invariant measure $\mathcal{A}_\sigma$ such that

$$\mathcal{A}_\sigma(f) = \lim_{n\to\infty} A^{(n)} f(\sigma) \quad \text{for all continuous functions } f. \tag{2.2.7.1}$$

The proof of the above statements uses the Birkhoff theorem. Fix arbitrarily a translational invariant measure, for instance a Bernoulli measure ν (Bernoulli measures are measures such that all spins are independent and identically distributed, hence they are translational invariant). Then by the Birkhoff theorem $\nu(\mathcal{X}_g^0) = 1$ and hence $\mathcal{X}_g^0$ is non-empty. The analogue of (2.2.2.2) is

$$\Omega_\sigma^0 = \{\sigma' \in \mathcal{X}_g^0 : \mathcal{A}_{\sigma'} = \mathcal{A}_\sigma\}, \qquad \mathcal{X}_{gg}^0 = \{\sigma \in \mathcal{X}_g^0 : \mathcal{A}_\sigma(\Omega_\sigma^0) = 1\},$$

and the analogue of Theorem 2.2.2.2 holds as well.

Theorem 2.2.7.1 *$\mathcal{X}_{gg}^0$ is non-empty and for any translational invariant measure μ, $\mu(\mathcal{X}_{gg}^0) = 1$. The measures $\mathcal{A}_\sigma$, $\sigma \in \mathcal{X}_{gg}^0$, are ergodic with respect to space translations. Moreover for any bounded measurable function f, $\mathcal{A}_\sigma(f)$ is a measurable function and for any probability ν on $\mathcal{X}_{gg}^0$, the measure μ defined by*

$$\mu(f) = \int_{\mathcal{X}_{gg}^0} \mathcal{A}_\sigma(f)\, \nu(d\sigma)$$

is translational invariant. Vice versa, if μ is translational invariant then

$$\mu(f) = \int_{\mathcal{X}_{gg}^0} \mathcal{A}_\sigma(f)\, \mu(d\sigma).$$

Finally, a translational invariant measure μ is ergodic if and only if $\mu = \mathcal{A}_\sigma$ for some $\sigma \in \mathcal{X}_{gg}^0$.

So far everything was general with no relations with the Gibbs measures. Let us now specify μ as a translational invariant DLR measure, namely $\mu \in \mathcal{G}^0$. If μ is ergodic, then for μ a.a. σ, $\mu = \mathcal{A}_\sigma$, which shows that $\mathcal{A}_\sigma \in \mathcal{G}^0$, μ almost surely. The above conclusion extends to any $\mu \in \mathcal{G}^0$.

Theorem 2.2.7.2 *If $\mu \in \mathcal{G}^0$, $\mu(\mathcal{X}_{gg}^0) = 1$ and for μ almost all $\sigma \in \mathcal{X}_{gg}^0$, $\mathcal{A}_\sigma \in \mathcal{G}^0$. Namely any translational invariant DLR measure μ is supported by configurations σ such that their ergodic averages $\mathcal{A}_\sigma$ defined in (2.2.7.1) exist and are ergodic DLR measures (so that the set of ergodic DLR measures is non-empty). As a consequence*

$$\mu(f) = \int_{\mathcal{X}_{gg}^0 \cap \mathcal{G}^0} \mathcal{A}_\sigma(f)\, \mu(d\sigma). \tag{2.2.7.2}$$

Proof The set $\mathcal{X}_g^0$ is $\mathcal{B}_\infty$ measurable because the value of $\mathcal{A}_\sigma(\cdot)$ does not change after a local modification of σ. On the other hand, by the Birkhoff theorem, $\mu(\mathcal{X}_g^0) = 1$ and by (2.2.2.3), $\mu(\{\sigma \in \mathcal{X}_{gg} : G_\sigma(\mathcal{X}_g^0) = 1\}) = 1$. Calling K the intersection of $\{\sigma \in \mathcal{X}_{gg} : G_\sigma(\mathcal{X}_g^0) = 1\}$ and $\{\sigma \in \mathcal{X}_{gg} \cap \mathcal{X}_g^0 :$ for any $C \in \mathcal{C}$, $\mathcal{A}_{\sigma'}(\mathbf{1}_C) = G_\sigma(A_{\sigma''}(\mathbf{1}_C))$ for G_σ a.a. $\sigma'\}$, by (2.2.5.1) $\mu(K) = 1$. By (2.2.7.1) and the Lebesgue dominated convergence theorem, for $\sigma \in K$

$$\int_{\Omega_\sigma} \mathcal{A}_{\sigma''}(f)\, G_\sigma(d\sigma'') = \lim_{n\to\infty} \int_{\Omega_\sigma} A_{\sigma''}^{(n)}(f)\, G_\sigma(d\sigma''),$$

since $\int_{\Omega_\sigma} A^{(n)}_{\sigma''}(f)\, G_\sigma(d\sigma'') = \frac{1}{|\Delta_n|}\sum_{i\in\Delta_n} \tau_i(G_\sigma)(f) =: \nu^{(n)}(f)$, $\nu^{(n)} \in \mathcal{G}$, being a convex combination of Gibbs measures ($G_\sigma \in \mathcal{G}$ and $\tau_i(G_\sigma)$ as well by (2.2.6.1)). Thus, by definition of K, if $\sigma \in K$, for G_σ a.a. $\sigma' \in \Omega_\sigma$, $\nu^{(n)} \to \mathcal{A}_{\sigma'}$ on the cylinders and, by a density argument (see Theorem A.4)

$$\mathcal{A}_{\sigma'}(f) = \lim_{n\to\infty} \nu^{(n)}(f), \quad \text{for all continuous functions } f.$$

Then $\mathcal{A}_{\sigma'}$ is weak limit of Gibbs measures and hence $\mathcal{A}_{\sigma'} \in \mathcal{G}^0$. Since this holds for G_σ a.a. σ' and for all $\sigma \in K$, $\mathcal{A}_{\sigma'} \in \mathcal{G}^0$ for μ a.a. σ'. □

Let us next see how the theorem fits with what we have said before about phase transitions. Intensive variables are defined in physics as spatial averages of local observables, represented, for instance, by continuous functions. We then recognize in $\mathcal{A}_\sigma(f)$ the value in the state σ of the intensive variable associated to the continuous function f. Particular cases occur when $f(\sigma) = \sigma(0)$; the intensive variable then has the meaning of the magnetization density of the system. The energy density is also an intensive variable, with

$$u(\sigma) = -\frac{1}{2}\sum_{y\neq 0} J(0,y)\,\sigma(0)\sigma(y) - h\sigma(0). \tag{2.2.7.3}$$

The relevant quantities for phase transitions, according to what we have said before, are the probability distributions of the intensive variables $\mathcal{A}_{\sigma'}(f)$. If a measure is ergodic, by Theorem 2.2.7.1, it coincides with a measure $\mathcal{A}_\sigma$ and it is supported by a set Ω^0_σ. Then any intensive variable $\mathcal{A}_{\sigma'}(f)$ is constant in the support of the ergodic measure, i.e. when $\sigma' \in \Omega_\sigma$. By Theorems 2.2.7.2 and 2.2.7.1, any translational invariant Gibbs measure can be decomposed into Gibbs measures which are ergodic, hence in each one of these all intensive variables do not fluctuate. The absence of fluctuations of the intensive variables characterizes the pure phases; hence the ergodic Gibbs measures represent pure phases. We have thus proved that any translational invariant Gibbs measure is a mixture of pure phases with weights which are uniquely determined. The existence of intensive variables which have non-trivial fluctuations is then the indication that there is a phase transition and the variables which fluctuate can be used as order parameters to classify the transition.

We denote by $\mathcal{G}^0_{\mathrm{extr}}$ the extremal points of $\mathcal{G}^0$, which are ergodic DLR measures and represent the pure phases of the system, described above. If the cardinality of $\mathcal{G}^0_{\mathrm{extr}}$ is larger than 1, there are several pure phases and there is a phase transition. However, a phase transition can occur also if $|\mathcal{G}^0| = 1$, due to a break of the translational symmetry. It may in fact be that the unique element $\mu \in \mathcal{G}^0$ can be non-trivially decomposed into non-translational invariant states, if the cardinality of $\mathcal{G}_e$ is larger than 1. Such effects may be related to the appearance of crystalline structures and states of the solid phase, but also to the existence of states describing coexistence of phases, like the Dobrushin states in $d \geq 3$ ferromagnetic Ising systems at low temperatures.

2.3 Boltzmann hypothesis, entropy and pressure

In this section we completely change perspective, now focusing on thermodynamics rather than Gibbs and DLR measures. In a sense we go back to the origins as the derivation of thermodynamics was historically the first goal of statistical mechanics with the aim of establishing a quantitative link between the macroscopic thermodynamic potentials and the microscopic inter-molecular interactions.

The analysis in this section is entirely based on the postulate that "entropy is proportional to the log of the number of states," the famous Boltzmann hypothesis. We will study the Ising model starting from such an assumption and derive expressions for the thermodynamic potentials in terms of the Ising hamiltonian, in particular we will relate thermodynamic pressure to partition functions. The Gibbs assumption on the structure of the equilibrium states is not needed, nonetheless our proofs will use extensively the DLR theory but only as a technical tool.

2.3.1 An example from information theory

A simple example of what we are going to do is borrowed from information theory. Consider a channel which transmits messages with a finite alphabet Ω; *we want to compute its capacity by counting how many messages can be emitted by a source which is "sending f with frequency ϕ."* By this we mean the following: f is a real valued function on Ω, $\phi \in (\min f, \max f)$, and the "normalized" number of messages we want to count is

$$\lim_{\delta \to 0} \lim_{N \to \infty} \frac{\log K_\delta(N)}{N}, \tag{2.3.1.1}$$

where

$$K_\delta(N) = \mathrm{card}\left\{ (\omega_1, \dots, \omega_N) \in \Omega^N : A_N := \left| \frac{1}{N} \sum_{i=1}^N f(\omega_i) - \phi \right| \le \delta \right\}. \tag{2.3.1.2}$$

Instead of going into combinatorics and Stirling formulas, it is more instructive for the applications to statistical mechanics to use a probabilistic approach. We start from the identity

$$K_\delta(N) = \sum_{\omega_1, \dots, \omega_N} \mathbf{1}_{A_N \le \delta} \frac{p(\omega_1) \cdots p(\omega_N)}{p(\omega_1) \cdots p(\omega_N)}, \tag{2.3.1.3}$$

and the whole trick is to choose properly the probability $p(\omega)$ on Ω. Calling $Z_b = \sum_{\omega \in \Omega} e^{bf(\omega)}$, we will see that the "right choice" is

$$p(\omega) = \frac{e^{bf(\omega)}}{Z_b}, \quad b \text{ such that } \sum_{\omega \in \Omega} p(\omega) f(\omega) = \phi. \tag{2.3.1.4}$$

Existence [and uniqueness] of b follows from the fact that $\sum_{\omega\in\Omega} p(\omega)f(\omega)$ is an increasing function of b which converges to $\min f$ and $\max f$ as $b \to \mp\infty$. With the choice (2.3.1.4) for $p(\cdot)$, we get from (2.3.1.3)

$$K_\delta(N) = \sum_{\omega_1,\dots,\omega_N} \mathbf{1}_{A_N\le\delta}[p(\omega_1)\cdots p(\omega_N)]e^{-\sum(bf(\omega_i)-\log Z_b)}. \tag{2.3.1.5}$$

It is now clear why (2.3.1.4) is a good choice: the sum $\sum_{i=1}^N bf(\omega_i)$ in the exponent is, by (2.3.1.2), approximately the same as the one fixed by the condition $A_N \le \delta$, while the second equality in (2.3.1.4) ensures that with probability converging to 1 the condition $A_N \le \delta$ is satisfied. Indeed, call $S := \log Z_b - b\phi$ and P_N the product probability $p(\omega_1)\cdots p(\omega_N)$, then

$$P_N(A_N \le \delta)e^{(S-b\delta)N} \le K_\delta(N) \le e^{(S+b\delta)N}. \tag{2.3.1.6}$$

By the law of large numbers, for any $\delta > 0$ $\lim_{N\to\infty} P_N(A_N \le \delta) = 1$ so that

$$\lim_{\delta\to 0}\lim_{N\to\infty} \frac{\log K_\delta(N)}{N} = S, \tag{2.3.1.7}$$

and by explicit computation

$$S = \log Z_b - b\phi = -\sum_{\omega\in\Omega} p(\omega)\log p(\omega) =: S(p). \tag{2.3.1.8}$$

In conclusion, the capacity of the channel equals the "information entropy" $S(p)$ of the associated Gibbs measure p of (2.3.1.4). In our applications $\Omega = \{-1, 1\}$, $\mathbb{N}_+$ is replaced by $\mathbb{Z}^d$, b by the inverse temperature β and f by an "energy function" u, which, however, unlike f, does not depend on a single spin. Independence will then fail and the above law of large numbers for Bernoulli measures will be replaced by ergodic theorems for DLR measures. The analysis will then identify the thermodynamic entropy with the above information entropy.

2.3.2 *Boltzmann hypothesis*

We will consider in the sequel the Ising hamiltonian

$$H_\Lambda(\sigma_\Lambda) = -\frac{1}{2}\sum_{x\ne y\in\Lambda} J(x,y)\sigma_\Lambda(x)\sigma_\Lambda(y) - h\sum_{x\in\Lambda}\sigma_\Lambda(x), \tag{2.3.2.1}$$

supposing that the $J(x, y)$ are translational invariant and summable; see Sect. 2.1.2. To formulate the Boltzmann hypothesis we first introduce the notion of "number of states with given energy."

Definition [Number of states with given energy density] For any bounded set $\Lambda \subset \mathbb{Z}^d$, $\delta > 0$ and $E \in \mathbb{R}$, we define

$$N_{E,\Lambda,\delta} = \text{card}\left\{\sigma_\Lambda : \left| H_\Lambda(\sigma_\Lambda) - |\Lambda| E \right| \le \delta |\Lambda| \right\}. \tag{2.3.2.2}$$

With the above definition we have relaxed the notion of number of states with given energy by introducing the accuracy parameter δ. This is technically convenient but also natural in a lattice model where the finite volume hamiltonian has finitely many values. We will eventually let $\delta \to 0$, but only after $|\Lambda| \to \infty$. Notice that E in (2.3.2.2) has the meaning of the energy density as $H_\Lambda(\sigma_\Lambda)/|\Lambda|$ is close to E (by δ). Recalling that the Boltzmann hypothesis relates the entropy to the log of the number of states, we next introduce

$$S_{E,\Lambda,\delta} = \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|}, \tag{2.3.2.3}$$

having divided by $|\Lambda|$ because we want the entropy per unit volume. $S_{E,\Lambda,\delta}$ cannot be a candidate for the Boltzmann entropy, as it is not a function of E, depending also on Λ and δ. Thus we need to let $\Lambda \to \mathbb{Z}^d$ and we want this to happen in "a regular way"; namely, we want the volume of a neighborhood of the boundary to be much smaller than the whole volume:

Definition 2.3.2.1 (van Hove sequences) A sequence Λ_n is "*van Hove*" if it is an increasing sequence of bounded regions which invades the whole $\mathbb{Z}^d$ (for any $x \in \mathbb{Z}^d$ there is n so that $x \in \Lambda_n$) and verifies the following property. Given any cube $\Delta \subset \mathbb{Z}^d$ and any partition of $\mathbb{Z}^d$ into translates of Δ, call Λ'_n the union of those cubes of the partition which are contained in Λ_n and Λ''_n of those which have non-empty intersection with Λ_n. Then

$$\lim_{n\to\infty} \frac{|\Lambda'_n|}{|\Lambda_n|} = \lim_{n\to\infty} \frac{|\Lambda''_n|}{|\Lambda_n|} = 1. \tag{2.3.2.4}$$

By default in the sequel $\Lambda \to \mathbb{Z}^d$ is meant to be taken in the van Hove sense.

We would then like to have a theorem which says that there exists $S(E)$ such that for any van Hove sequence Λ_n

$$\lim_{\delta\to 0}\lim_{n\to\infty} S_{E,\Lambda_n,\delta} = S(E). \tag{2.3.2.5}$$

With such a result we can then reasonably formulate the Boltzmann hypothesis in our Ising model by saying that *"$S(E)$ as defined by (2.3.2.5) is the thermodynamic entropy of a system whose microscopic interactions are described by (2.3.2.1)."* The definition poses consistency problems as $S(E)$ should verify the properties that entropy has in thermodynamics and a great success of the theory is that all this can indeed be rigorously established. Using the thermodynamic relations we can obtain from $S(E)$ other thermodynamic potentials, for instance the inverse temperature β

as a function of E is equal to the derivative $dS(E)/dE$. As we will see, there are also formulas for computing the pressure P_β as a function of β, once $S(E)$ as a function of E is known. We will prove that the pressure P_β obtained from $S(\cdot)$ in this way has a simple expression in terms of the Ising partition function. Let Λ_n be any van Hove sequence; then

$$P_\beta = \lim_{n\to\infty} \frac{\log Z_{\beta,\Lambda_n}}{\beta|\Lambda_n|}, \quad Z_{\beta,\Lambda} = \sum_{\sigma_\Lambda\in\mathcal{X}_\Lambda} e^{-\beta H_\Lambda(\sigma_\Lambda)}. \tag{2.3.2.6}$$

For practical and numerical purposes the expression (2.3.2.6) is much easier to handle than (2.3.2.5); thus, while from an axiomatic viewpoint entropy is the starting point from where all the other thermodynamic potentials are derived, often in the literature (2.3.2.6) is taken as a definition of pressure and all thermodynamics can then be derived.

The equality between the pressure computed from the entropy (2.3.2.5) via thermodynamic relations and the pressure computed via (2.3.2.6) goes under the name of "equivalence of ensembles." The ensembles in the statement are the "microcanonical ensemble," which denotes the phase space reduced to the subset (2.3.2.2) and the "grand-canonical ensemble" which is the full phase space $\mathcal{X}_\Lambda$ where the partition function is computed. This reminds one of the minimization of a function in $\mathbb{R}^n$ where using Lagrange multipliers a constraint can be dropped and the problem reduced to one in the whole $\mathbb{R}^n$. The analogy is made more evident in the following outline of the proof of (2.3.2.5); the real proof is postponed to the next subsections, while (2.3.2.6) is proved in Sect. 2.3.3. We start from the trivial identity

$$\begin{aligned} N_{E,\Lambda,\delta} &= \sum_{\sigma_\Lambda\in\mathcal{X}_\Lambda} \mathbf{1}_{|H_\Lambda(\sigma_\Lambda)-|\Lambda|E|\le\delta|\Lambda|} \\ &= \sum_{\sigma_\Lambda\in\mathcal{X}_\Lambda} \mathbf{1}_{|H_\Lambda(\sigma_\Lambda)-|\Lambda|E|\le\delta|\Lambda|} \frac{e^{-\beta H_\Lambda(\sigma_\Lambda)}}{Z_{\beta,\Lambda}} \{Z_{\beta,\Lambda} e^{\beta H_\Lambda(\sigma_\Lambda)}\}. \end{aligned}$$

We proceed by writing upper and lower bounds, which will hopefully coincide in the limit, and bound the last term $e^{\beta H_\Lambda(\sigma_\Lambda)}$ in the curly brackets by $e^{\beta(E\pm\delta)|\Lambda|}$: we then find for $\frac{\log N_{E,\Lambda,\delta}}{|\Lambda|}$ the bounds

$$\begin{aligned} &\le \frac{\log Z_{\beta,\Lambda}}{|\Lambda|} + \beta(E+\delta) \\ &\ge \frac{\log Z_{\beta,\Lambda}}{|\Lambda|} + \beta(E-\delta) \\ &\quad + \frac{1}{|\Lambda|} \log G_{\beta,\Lambda}\big(\{|H_\Lambda(\sigma_\Lambda) - |\Lambda|E| \le \delta|\Lambda|\}\big), \end{aligned} \tag{2.3.2.7}$$

where $G_{\beta,\Lambda}(\sigma_\Lambda) = \frac{e^{-\beta H_\Lambda(\sigma_\Lambda)}}{Z_{\beta,\Lambda}}$ is the Gibbs measure (with zero boundary conditions). We now face two problems: the first one, the easier one, solved in the next subsection, is to prove that the limit on the r.h.s. of (2.3.2.6) exists, thus defining a function

P_β which only afterwards will be identified with the thermodynamic pressure. This yields

$$\lim_{\delta\to 0}\limsup_{\Lambda\to\mathbb{Z}^d}\frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \quad\le\quad \inf_{\beta>0}(\beta P_\beta+\beta E). \tag{2.3.2.8}$$

The second and more serious problem concerns the lower bound, a problem that can be avoided if we suppose that there is β^* so that

$$\lim_{\delta\to 0}\lim_{\Lambda\to\mathbb{Z}^d}\frac{1}{|\Lambda|}\log G_{\beta^*,\Lambda}\left(\left\{\left|\frac{H_\Lambda(\sigma_\Lambda)}{|\Lambda|}-E\right|\le\delta\right\}\right)=0. \tag{2.3.2.9}$$

Then

$$\lim_{\delta\to 0}\liminf_{\Lambda\to\mathbb{Z}^d}\frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \quad\ge\quad \beta^* P_{\beta^*}+\beta^* E, \tag{2.3.2.10}$$

which together with (2.3.2.8) yields the result that the limit in (2.3.2.5) exists and is equal to

$$S(E)=\beta^* P_{\beta^*}+\beta^* E=\inf_{\beta>0}(\beta P_\beta+\beta E), \tag{2.3.2.11}$$

which is the well known thermodynamic formula for the entropy in terms of the pressure.

The problem is that (2.3.2.9) is not true in general. As we will see it holds for E in a set $\mathcal{E}_{\rm erg}$: for any $E\in\mathcal{E}_{\rm erg}$ there is a special value of β for which (2.3.2.9) holds. We will discuss later how to proceed when $E\notin\mathcal{E}_{\rm erg}$. The terminology hints that the crucial estimate (2.3.2.9) is related to an ergodic theorem. Let $\tau_z u(\sigma)$ be the translate by $z\in\mathbb{Z}^d$ of the function $u(\sigma)$ defined in (2.2.7.3). Then

$$\mathcal{A}_{\sigma,\Lambda}(u):=\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\tau_x u(\sigma) \tag{2.3.2.12}$$

is the "ergodic average" of u in Λ computed at σ. We claim that $\mathcal{A}_{\sigma,\Lambda}(u)$ is "close to" $H_\Lambda(\sigma_\Lambda)$, the latter to be thought of as a function of $\sigma\in\mathcal{X}$ (with σ_Λ the restriction of σ to Λ) and thus also denoted by $H_\Lambda(\sigma)$. Indeed

$$H_\Lambda(\sigma)\equiv H_\Lambda(\sigma_\Lambda)=|\Lambda|\mathcal{A}_{\sigma,\Lambda}(u)+\frac{1}{2}\sum_{x\in\Lambda}\sum_{y\in\Lambda^c}J(x,y)\sigma(x)\sigma(y). \tag{2.3.2.13}$$

Analogously, for any $\sigma_\Lambda\in\mathcal{X}_\Lambda$ and $\sigma_{\Lambda^c}\in\mathcal{X}_{\Lambda^c}$, denoting $\sigma=(\sigma_\Lambda,\sigma_{\Lambda^c})$

$$H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})=|\Lambda|\mathcal{A}_{\sigma,\Lambda}(u)-\frac{1}{2}\sum_{x\in\Lambda,y\in\Lambda^c}J(x,y)\sigma_\Lambda(x)\sigma_{\Lambda^c}(y). \tag{2.3.2.14}$$

We will see in the next subsection that if $\Lambda\to\mathbb{Z}^d$ in the van Hove sense, then

$$\lim_{\Lambda\to\mathbb{Z}^d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{y\in\Lambda^c}|J(x,y)|=0, \tag{2.3.2.15}$$

so that the last term in (2.3.2.13)–(2.3.2.14) is a "negligible error" (once divided by $|\Lambda|$) and $H_\Lambda(\sigma)$ is "essentially" the ergodic average of u in Λ.

If μ is an ergodic measure on $\mathcal{X}$, $\mathcal{A}_{\sigma,\Lambda}(u) \to \mu(u)$ ($\mu(u)$ being the expectation of u) for μ almost all σ. Thus, if $\mu(u) = E$,

$$\lim_{\delta\to 0}\lim_{\Lambda\to\mathbb{Z}^d}\mu\left(\left\{\left|\frac{H_\Lambda(\sigma_\Lambda)}{|\Lambda|} - E\right| \le \delta\right\}\right) = 1. \tag{2.3.2.16}$$

As we shall see using (2.3.2.15) we can modify (2.3.2.7) so that instead of $G_{\beta,\Lambda}$ we can put any DLR measure at inverse temperature β with an error such that its log, divided by $|\Lambda|$ vanishes as $|\Lambda| \to \infty$. Thus (2.3.2.9) holds for all $E \in \mathcal{E}_{\text{erg}}$ where the latter is the set of all $\mu(u)$ as μ varies over all the ergodic DLR measures at $\beta > 0$ and β varies in the whole $\mathbb{R}_+$.

If all DLR measures were ergodic we will be in business, but as we have seen in the previous section this is not true in case of phase transitions, when in fact there are distinct ergodic DLR measures at the same β so that their convex combinations are DLR at β but not ergodic. In such a case the approach which starts from (2.3.2.7) must be aborted and we have to go one step back. Suppose the energy density E belongs to an interval $[E', E'']$, $E = aE' + (1-a)E''$, $a \in (0,1)$, whose endpoints are both in $\mathcal{E}_{\text{erg}}$, namely are expectations of $u(\sigma)$ relative to two ergodic DLR measures at the same inverse temperature β, then the previous argument can be reproduced in the following way. Suppose for simplicity Λ to be a cube and suppose that it can be split into two rectangles Λ' and Λ'' such that $a = |\Lambda'|/|\Lambda|$ (approximate equality is however sufficient). By (2.3.2.15), if $|\Lambda|$ is large enough the set $\{|H_{\Lambda'}(\sigma_{\Lambda'}) - |\Lambda'|E'| \le \frac{\delta}{4}|\Lambda'|\} \cap \{|H_{\Lambda''}(\sigma_{\Lambda''}) - |\Lambda''|E''| \le \frac{\delta}{4}|\Lambda''|\}$ is contained in $\{|H_\Lambda(\sigma_\Lambda) - |\Lambda|E| \le \delta|\Lambda|\}$. Then the lower bound in (2.3.2.7) can be replaced by

$$\begin{aligned}\frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \ge \frac{\log Z_{\beta,\Lambda}}{|\Lambda|} &+ \beta(E-\delta)\\ &+ \frac{1}{|\Lambda|}\log G_{\beta,\Lambda}\left(\left\{|H_\Lambda(\sigma_{\Lambda'}) - |\Lambda'|E'| \le \frac{\delta}{4}|\Lambda'|\right\}\right.\\ &\left.\cap\left\{|H_\Lambda(\sigma_{\Lambda''}) - |\Lambda''|E''| \le \frac{\delta}{4}|\Lambda''|\right\}\right).\end{aligned} \tag{2.3.2.17}$$

Using again (2.3.2.15), we can also factorize the Gibbs measure in the sense that the log of

$$\frac{G_{\beta,\Lambda}(\{|H_{\Lambda'}(\sigma_{\Lambda'}) - |\Lambda'|E'| \le \frac{\delta}{4}|\Lambda'|\} \cap \{|H_{\Lambda''}(\sigma_{\Lambda''}) - |\Lambda''|E''| \le \frac{\delta}{4}|\Lambda''|\})}{G_{\beta,\Lambda'}(\{|H_{\Lambda'}(\sigma_{\Lambda'}) - |\Lambda'|E'| \le \frac{\delta}{4}|\Lambda'|\})G_{\beta,\Lambda''}(\{|H_{\Lambda''}(\sigma_{\Lambda''}) - |\Lambda''|E''| \le \frac{\delta}{4}|\Lambda''|\})}$$

divided by $|\Lambda$ vanishes as the cube $\Lambda \to \mathbb{Z}^d$. We can thus replace in (2.3.2.17) the $G_{\beta,\Lambda}$ expectation by the product of expectations with $G_{\beta,\Lambda'}$ and $G_{\beta,\Lambda''}$ and then reproduce the argument used when $E \in \mathcal{E}_{\text{erg}}$.

It remains for us to characterize the set $\mathcal{E}_{\text{allwd}}$ defined as the set of all E which are contained in intervals $[E', E'']$ with E' and E'' expectations of u with respect to two ergodic measures which are DLR at the same value of β. By the analysis of the previous section we know that for each $\beta > 0$ there exist ergodic DLR measures; moreover, any translational invariant DLR measure at β (their collection being denoted by $\mathcal{G}^0_\beta$) can be written as an integral over the subset of ergodic DLR measures (ergodic decomposition theorem). It then follows that

$$\mathcal{E}_{\text{allwd}} = \bigcup_{\beta>0} \{E \in \mathbb{R} : E = \mu(u), \mu \in \mathcal{G}^0_\beta\}. \tag{2.3.2.18}$$

We will next see how the sets $\{E \in \mathbb{R} : E = \mu(u), \mu \in \mathcal{G}^0_\beta\}$ are related to the pressure P_β. Together with the proof of existence of the limit (2.3.2.6) defining P_β we also have that $\pi_\beta := \beta P_\beta$ is a continuous convex function of β. Then by general theorems on convex functions, see Sect. 2.3.6, its right and left derivatives $D^\pm \pi_\beta$ exist, and $D^-\pi_\beta \le D^+\pi_\beta$. We will then prove that for any $\beta > 0$ the values $E'' = -D^-\pi_\beta$ and $E' = -D^+\pi_\beta$, $E' \le E''$ are both in $\mathcal{E}_{\text{erg}}$; thus

$$\mathcal{E} =: -\mathcal{E}_{\text{allwd}} = \bigcup_{\beta>0} [D^-\pi_\beta, D^+\pi_\beta]. \tag{2.3.2.19}$$

It will also follow from general theorems on convex functions that both $[D^-\pi_\beta, D^+\pi_\beta]$ and $\mathcal{E}_{\text{allwd}}$ are bounded intervals. The restriction to a bounded interval of energies is a consequence of the system being a lattice model in which the energy density is bounded both from above and below (in a continuum system it is generally only bounded from below). We will see in the sequel that bounded spin systems have strange properties; in particular negative temperatures can also be defined if we enlarge the set $\mathcal{E}_{\text{allwd}}$ to all E for which the limit (2.3.2.5) exists, a phenomenon which disappears if the energy density is unbounded from above.

2.3.3 Thermodynamic limit of the pressure

In this subsection we will prove

Theorem 2.3.3.1 *For any van Hove sequence $\Lambda \to \mathbb{Z}^d$ and any sequence of b.c. $\sigma_{\Lambda^c} \in \mathcal{X}_{\Lambda^c}$,*

$$\lim_{\Lambda\to\mathbb{Z}^d} P_{\beta,\Lambda}(\sigma_{\Lambda^c}) =: P_\beta, \tag{2.3.3.1}$$

where, recalling that $\mathcal{X}_\Lambda = \{-1, 1\}^\Lambda$,

$$\begin{aligned} P_{\beta,\Lambda}(\sigma_{\Lambda^c}) &= \frac{1}{\beta|\Lambda|} \log Z_{\beta,\Lambda}(\sigma_{\Lambda^c}), \\ Z_{\beta,\Lambda}(\sigma_{\Lambda^c}) &= \sum_{\sigma_\Lambda\in\mathcal{X}_\Lambda} e^{-\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}. \end{aligned} \tag{2.3.3.2}$$

Before starting the proof of the theorem we state and prove the following lemma.

Lemma 2.3.3.2 *There is a constant c so that for any bounded Λ, any σ_Λ and any σ_{Λ^c}*

$$|H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})| \le c|\Lambda|. \tag{2.3.3.3}$$

Moreover, given any $\epsilon > 0$ for all cubes Δ large enough

$$\sum_{x\in\Delta, y\in\Delta^c} |J(x,y)| \le \epsilon|\Delta|. \tag{2.3.3.4}$$

Proof We have

$$|H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})| \le \sum_{x\in\Lambda}\sum_{y\ne x} |J(x,y)| + \sum_{x\in\Lambda}|h|,$$

which proves (2.3.3.3) with $c = |h| + \sum_{x\ne 0}|J(0,x)|$.

Given $R > 0$ we split the sum in the l.h.s. of (2.3.3.4) as

$$\sum_{x\in\Delta, \mathrm{dist}(x,\Delta^c)\le R}\ \sum_{y\in\Delta^c} |J(x,y)| + \sum_{x\in\Delta, \mathrm{dist}(x,\Delta^c)> R}\ \sum_{y\in\Delta^c} |J(x,y)|.$$

The first term is bounded by $cL^{d-1}R$, L being the side of Δ; the second one by

$$|\Delta| \sum_{|x|>R} |J(0,x)| < \frac{\epsilon}{2}|\Delta|,$$

if R is large enough. Given such a R we then choose L so large that $cL^{d-1}R \le (\epsilon/2)L^d$, hence (2.3.3.4). □

Proof of Theorem 2.3.3.1 In this proof we do not make explicit β in the notation. By (2.3.3.3) there is c so that

$$|P_\Lambda(\sigma_{\Lambda^c})| = \frac{1}{\beta|\Lambda|}\,|\log Z_\Lambda(\sigma_{\Lambda^c})| \le c. \tag{2.3.3.5}$$

Then by compactness there exists an increasing sequence of cubes $\Delta \to \mathbb{Z}^d$ such that

$$\lim_{\Delta\to\mathbb{Z}^d} P_\Delta =: P. \tag{2.3.3.6}$$

Now, $P_\Delta = \frac{\log Z_\Delta}{\beta|\Delta|}$, $Z_\Delta = \sum_{\sigma_\Delta} e^{-\beta H_\Delta(\sigma_\Delta)}$, $H_\Delta(\sigma_\Delta)$ as in (2.3.2.1). Let Λ be a van Hove sequence and σ_{Λ^c} a sequence of boundary conditions. We will prove that $\lim_{\Lambda\to\mathbb{Z}^d} P_\Lambda(\sigma_{\Lambda^c}) = P$, P as in (2.3.3.6). Let $\epsilon > 0$ and let Δ be a cube of the sequence in (2.3.3.6) as large as required for (2.3.3.4) to hold. Consider a partition into translates of Δ, call $\Delta(i)$ those in Λ, and Λ' their union. With $c = |h| + \sum_{x\neq 0} |J(0,x)|$ we then have, using Lemma 2.3.3.2,

$$\left| H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) - \sum_{\Delta(i)\subset\Lambda'} H_{\Delta(i)}(\sigma_{\Delta(i)}) \right| \le \epsilon|\Lambda'| + c|\Lambda\setminus\Lambda'|. \tag{2.3.3.7}$$

From (2.3.3.7) we get

$$\left| P_\Lambda(\sigma_{\Lambda^c}) - \frac{|\Lambda'|}{|\Lambda|} P_\Delta \right| \le \epsilon \frac{|\Lambda'|}{|\Lambda|} + c\frac{|\Lambda\setminus\Lambda'|}{|\Lambda|}.$$

Letting $\Lambda \to \mathbb{Z}^d$ and calling $P' \le P''$ the lim inf and lim sup of $P_\Lambda(\sigma_{\Lambda^c})$

$$P_\Delta - \epsilon \le P' \le P'' \le P_\Delta + \epsilon.$$

By letting $\Delta \to \mathbb{Z}^d$ along the sequence in (2.3.3.6) we get $P - \epsilon \le P' \le P'' \le P + \epsilon$ and by the arbitrariness of ϵ, (2.3.3.1). □

2.3.4 *Thermodynamic limit of the entropy*

The precise statement of the results outlined in Sect. 2.3.2 is

Theorem 2.3.4.1 (Main result) *There is a non-empty interval $\mathcal{E}$ (defined in (2.3.2.19)) such that for any $E \in \mathcal{E}_{\text{allw}}$, $\mathcal{E}_{\text{allw}} := -\mathcal{E}$, and for any van Hove sequence $\Lambda \to \mathbb{Z}^d$*

$$\lim_{\delta\to 0} \limsup_{\Lambda\to\mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} = \lim_{\delta\to 0} \liminf_{\Lambda\to\mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} =: S(E), \tag{2.3.4.1}$$

with $S(E)$ a strictly increasing, concave function of E. Moreover, if P_β denotes the pressure defined in (2.3.3.1),

$$\beta P_\beta = \sup_{E\in\mathcal{E}_{\text{allw}}} \{\beta(-E) + S(E)\}, \qquad -S(E) = \sup_{\beta>0}\{\beta(-E) - \beta P_\beta\}, \tag{2.3.4.2}$$

so that the functions $s(E) = -S(-E)$, $E \in \mathcal{E}$, and $\pi_\beta = \beta P_\beta$, $\beta \in \mathbb{R}_+$, are Legendre transforms of each other. $S(E)$ and P_β are both differentiable except at countably many points, while left and right derivatives $D^\pm$ exist everywhere. We say that E and β are conjugate if $e = -E \in [D^-\pi_\beta, D^+\pi_\beta]$; in such a case $\beta \in [D^-s(e), D^+s(e)]$ and, if E and β are conjugate, then

$$P_\beta = -E + \beta^{-1} S(E). \tag{2.3.4.3}$$

Finally, $s(e)$ is linear in $[D^-\pi_\beta, D^+\pi_\beta]$ and π_β in $[D^-s(e), D^+s(e)]$.

Thermodynamic interpretation

As discussed in Sect. 2.3.2, in agreement with the Boltzmann hypothesis we interpret $S(E)$ as the thermodynamic entropy, and then P_β as given in (2.3.4.3) in terms of $S(E)$ is the thermodynamic pressure at the inverse temperature β, as thermodynamics says that entropy and pressure are related as in (2.3.4.3). In conclusion, Theorem 2.3.4.1 and the Boltzmann hypothesis completely determine the thermodynamics of the Ising model and justify the identification of the limit (2.3.3.1) as the thermodynamic pressure.

Proof of the statements of Theorem 2.3.4.1

The proof of the theorem takes most of this section. Convexity of the pressure is proved in Theorem 2.3.5.1; (2.3.4.1) is proved in Theorem 2.3.7.5; (2.3.4.2) in Theorem 2.3.8.1; the properties of $S(E)$ stated in Theorem 2.3.4.1 are proved in Theorem 2.3.8.2, while in Theorem 2.3.8.3 there is a characterization of the set $\mathcal{E}$.

Scheme of the proof of Theorem 2.3.4.1

We will first prove (2.3.4.1) only for $E \in \mathcal{E}_{\text{erg}}$, $\mathcal{E}_{\text{erg}} := \{E : E = \mu(u),\ \mu$ ergodic DLR at inverse temperature $\beta, \beta > 0\}$. For such special values of E we can in fact follow the heuristic argument presented in Sect. 2.3.2 to prove (2.3.4.1). The argument also proves (2.3.4.3), which is thus established for $E \in \mathcal{E}_{\text{erg}}$. The extension to $E \notin \mathcal{E}_{\text{erg}}$ is based on the general properties of convex functions, which are recalled together with the definition and properties of Legendre transforms in Sect. 2.3.6.

Proposition 2.3.4.2 *For any $E \in \mathcal{E}_{\text{erg}}$ and for any van Hove sequence Λ_n, the limit $S(E)$ in* (2.3.4.1) *exists and $S(E)$ satisfies* (2.3.4.3).

We fix in the sequel $E \in \mathcal{E}_{\text{erg}}$ so that there is an ergodic DLR measure μ at some $\beta > 0$ such that $\mu(u) = E$. We postpone the proof of Proposition 2.3.4.2 to first deal with the following three lemmas.

Lemma 2.3.4.3 *For any $\delta > 0$ and any van Hove sequence $\Lambda \to \mathbb{Z}^d$*

$$\lim_{\Lambda \to \mathbb{Z}^d} \mu\left(\frac{1}{|\Lambda|} \left| \sum_{x \in \Lambda} \{\tau_x u(\sigma) - E\} \right| \le \delta \right) = 1. \tag{2.3.4.4}$$

Proof Since μ is ergodic, for any $\epsilon > 0$ there is a cube Δ_ϵ so that for any cube Δ, $|\Delta| \ge |\Delta_\epsilon|$,

$$\mu\left(\left| \sum_{x \in \Delta} \{\tau_x u(\sigma) - E\} \right| \right) < \epsilon |\Delta|. \tag{2.3.4.5}$$

Define a partition of $\mathbb{Z}^d$ into cubes which are translates of Δ; call $\Delta(i)$ those contained in Λ, and Λ' their union. Then

$$\begin{aligned}
&\mu\left(\frac{1}{|\Lambda|}\left|\sum_{x\in\Lambda}\{\tau_x u(\sigma)-E\}\right|>\delta\right)\\
&\quad\le\frac{1}{\delta|\Lambda|}\mu\left(\left|\sum_{x\in\Lambda}\{\tau_x u(\sigma)-E\}\right|\right)\\
&\quad\le\frac{1}{\delta|\Lambda|}\left\{\sum_{\Delta(i)}\mu\left(\left|\sum_{x\in\Delta(i)}\{\tau_x u(\sigma)-E\}\right|\right)+c|\Lambda\setminus\Lambda'|\right\},
\end{aligned}\tag{2.3.4.6}$$

because u is bounded. By (2.3.4.5),

$$\mu\left(\frac{1}{|\Lambda|}\left|\sum_{x\in\Lambda}\{\tau_x u(\sigma)-E\}\right|>\delta\right)\le\frac{c|\Lambda\setminus\Lambda'|}{\delta|\Lambda|}+\frac{\epsilon|\Lambda'|}{\delta|\Lambda|}.$$

In the limit $\Lambda\to\mathbb{Z}^d$, $\frac{c|\Lambda\setminus\Lambda'|}{\delta|\Lambda|}\to 0$, $\frac{\epsilon|\Lambda'|}{\delta|\Lambda|}\to\frac{\epsilon}{\delta}$ and by the arbitrariness of ϵ, (2.3.4.4) is proved. □

Lemma 2.3.4.4 *Given any $\epsilon>0$ for all cubes Δ large enough*

$$\sup_{\sigma_\Delta,\sigma_{\Delta^c}}\left|\sum_{x\in\Delta}[\tau_x u((\sigma_\Delta,\sigma_{\Delta^c}))-\tau_x u((\sigma_\Delta,0_{\Delta^c}))]\right|\le\epsilon|\Delta|,\tag{2.3.4.7}$$

where $u((\sigma_\Delta,0_{\Delta^c}))$ is defined by (2.2.7.3) *putting $\sigma(\cdot)=0$ on Δ^c*;

$$\sup_{\sigma_\Delta,\sigma_{\Delta^c}}\left|H_\Delta(\sigma_\Delta)-H_\Delta((\sigma_\Delta|\sigma_{\Delta^c}))\right|\le\epsilon|\Delta|.\tag{2.3.4.8}$$

Proof Equations (2.3.4.7)–(2.3.4.8) follow straightforwardly from (2.3.3.4). □

Lemma 2.3.4.5 *Let $\Lambda\to\mathbb{Z}^d$ be a van Hove sequence*; *then*, *given any $\epsilon>0$ for all Λ large enough*, *we have*

$$\sup_{\sigma_\Lambda,\sigma_{\Lambda^c}}\left|H_\Lambda(\sigma_\Lambda)-\sum_{x\in\Lambda}\tau_x u((\sigma_\Lambda,\sigma_{\Lambda^c}))\right|\le\epsilon|\Lambda|,\tag{2.3.4.9}$$

$$\sup_{\sigma_\Lambda,\sigma_{\Lambda^c}}\left|H_\Lambda(\sigma_\Lambda)-H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})\right|\le\epsilon|\Lambda|.\tag{2.3.4.10}$$

Proof Let $\Delta(i)$ and Λ' be defined as in the definition of the van Hove sequences; then by (2.3.2.12), the l.h.s. of (2.3.4.9) is bounded by

$$\sum_{\Delta(i)\subset\Lambda'}\left|\sum_{x\in\Delta(i)}[\tau_x u((\sigma_\Lambda,\sigma_{\Lambda^c}))-\tau_x u((\sigma_\Lambda,0_{\Lambda^c}))]\right|+c|\Lambda\setminus\Lambda'|,$$

so that (2.3.4.9) follows from (2.3.4.7), because $|\Lambda \setminus \Lambda'|/|\Lambda| \to 0$. Analogously,

$$H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) = H_{\Lambda\setminus\Delta(i)}(\sigma_{\Lambda\setminus\Delta(i)}|\sigma_{\Lambda^c}) + H_{\Delta(i)}(\sigma_{\Delta(i)}|\sigma_{\Lambda\setminus\Delta(i)}, \sigma_{\Lambda^c}).$$

Hence by (2.3.4.8) for $\Delta(i)$ large enough

$$|H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) - H_{\Delta(i)}(\sigma_{\Delta(i)}) - H_{\Lambda\setminus\Delta(i)}(\sigma_{\Lambda\setminus\Delta(i)}|\sigma_{\Lambda^c})| \le \frac{\epsilon}{4}|\Delta(i)|.$$

By iteration

$$\left| H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) - \sum_{\Delta(i)\subset\Lambda'} H_{\Delta(i)}(\sigma_{\Delta(i)}) \right| \le c|\Lambda \setminus \Lambda'| + \frac{\epsilon}{4}|\Lambda|.$$

By the same argument

$$\left| H_\Lambda(\sigma_\Lambda) - \sum_{\Delta(i)\subset\Lambda'} H_{\Delta(i)}(\sigma_{\Delta(i)}) \right| \le c|\Lambda \setminus \Lambda'| + \frac{\epsilon}{4}|\Lambda|, \tag{2.3.4.11}$$

hence (2.3.4.10). □

Proof of Proposition 2.3.4.2 Let $\Lambda \to \mathbb{Z}^d$ be a van Hove sequence. Then by Lemmas 2.3.4.3 and 2.3.4.5 for Λ large enough,

$$\mu\big(\big|H_\Lambda(\sigma_\Lambda) - |\Lambda|E\big| \le \delta\big) \ge \frac{1}{2}, \tag{2.3.4.12}$$

and by Theorem 2.3.3.1, given any ϵ for Λ large enough and any σ_{Λ^c},

$$e^{-\beta\epsilon|\Lambda|} \le \frac{Z_\Lambda(\sigma_{\Lambda^c})}{e^{\beta P_\beta|\Lambda|}} \le e^{\beta\epsilon|\Lambda|}. \tag{2.3.4.13}$$

We are now ready for the proof of Proposition 2.3.4.2. We first write

$$\begin{aligned} N_{E,\Lambda,\delta} &= \sum_{\sigma_\Lambda} \mathbf{1}_{H_\Lambda(\sigma_\Lambda)-|\Lambda|E|\le\delta|\Lambda|}, \\ &= \int d\mu(\sigma_{\Lambda^c}) \sum_{\sigma_\Lambda} \mathbf{1}_{|H_\Lambda(\sigma_\Lambda)-|\Lambda|E|\le\delta|\Lambda|} \frac{e^{-\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}}{Z_\Lambda(\sigma_{\Lambda^c})} \\ &\quad \times \{Z_\Lambda(\sigma_{\Lambda^c}) e^{\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}\}. \end{aligned}$$

By (2.3.4.13) and (2.3.4.10) for any $\sigma_\Lambda : |H_\Lambda(\sigma_\Lambda) - |\Lambda|E| \le \delta|\Lambda|$ and all Λ large enough,

$$\{Z_\Lambda(\sigma_{\Lambda^c}) e^{\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}\} \le e^{\beta(P_\beta+\epsilon)|\Lambda|} e^{\beta(H_\Lambda(\sigma_\Lambda)+\epsilon)|\Lambda|} \le e^{\beta(P_\beta+E)|\Lambda|} e^{\beta(2\epsilon+\delta)|\Lambda|},$$

so that

$$N_{E,\Lambda,\delta} \le e^{\beta(P_\beta+E)|\Lambda|} e^{\beta(2\epsilon+\delta)|\Lambda|}. \tag{2.3.4.14}$$

Notice that the upper bound (2.3.4.14) is valid for any E and not only for the special value $E = \mu(u)$. For the lower bound, instead after proceeding similarly, we use (2.3.4.12) to get

$$N_{E,\Lambda,\delta} \geq e^{\beta(P_\beta+E)|\Lambda|} e^{-\beta(2\epsilon+\delta)|\Lambda|} \frac{1}{2}, \tag{2.3.4.15}$$

and Proposition 2.3.4.2 is proved. □

2.3.5 Equivalence of ensembles

Proposition 2.3.4.2 is a first indication of it, but it is not yet a full justification for interpreting the limit of $\log Z_\Lambda(\sigma_{\Lambda^c})/(\beta|\Lambda|)$ as the thermodynamic pressure. One reason is that the entropy $S(E)$ has only been defined for $E \in \mathcal{E}_{\mathrm{erg}}$ and therefore the identity (2.3.4.3) is only established for such values of E. Notice however that (2.3.4.3) covers all $\beta > 0$ as for any β there is an ergodic DLR measure, hence a value $E \in \mathcal{E}_{\mathrm{erg}}$ for which (2.3.4.3) holds. As we are going to explain this creates a serious consistency problem, which could make the interpretation of P_β as the pressure dubious. In fact, according to thermodynamic principles, if P_β is the thermodynamic pressure at the inverse temperature β, then the function

$$S^{\mathrm{td}}(E) := -\sup_{\beta>0}\{\beta(-E) - \beta P_\beta\} \tag{2.3.5.1}$$

is the thermodynamic entropy. We thus have two entropies, $S(E)$ defined directly from the Boltzmann hypothesis, and $S^{\mathrm{td}}(E)$ defined starting from P_β. Consistency requires that they coincide which indeed will be proved. There is also another consistency problem to check: by thermodynamic principles the pressure is related to the entropy by

$$\beta P_\beta = \sup_E\{\beta(-E) - [-S^{\mathrm{td}}(E)]\}, \tag{2.3.5.2}$$

namely $\pi_\beta = \beta P_\beta$ is the Legendre transform of $s(E) = -S^{\mathrm{td}}(-E)$; hence π_β must be a convex function of β—see the paragraph "Legendre transforms" at the beginning of Sect. 2.3.8.

We thus need to prove (1) that the limit in (2.3.4.1) is well posed also for $E \notin \mathcal{E}_{\mathrm{erg}}$; (2) that $S(E) = S^{\mathrm{td}}(E)$ for all E; (3) that βP_β is a convex function of β, a property which implies (2.3.5.2) because by (2.3.5.1) $s(E) = -S^{\mathrm{td}}(-E)$ is the Legendre transform of βP_β (see again the paragraph "Legendre transforms" in Sect. 2.3.8).

Property (3) is proved as a corollary of Theorem 2.3.5.1 below, so that the requirement (2.3.5.2) from thermodynamics is fulfilled. More serious is the consistency problem (2), i.e. that $S(E) = S^{\mathrm{td}}(E)$, which will be proved in the next subsections together with (1), leading in the end to the proof of Theorem 2.3.4.1. Problem (2) is usually referred to as "equivalence of ensembles." The equivalence is between the "grand-canonical ensemble" used in the definition of $Z_\Lambda(\sigma_{\Lambda^c})$, i.e. all $\sigma_\Lambda \in \mathcal{X}_\Lambda$, and the "micro-canonical ensemble" used in the definition of $N_{E,\Lambda,\delta}$, where σ_Λ is restricted by the energy constraint, and, as we shall see, there are other ensembles related to other variables than the energy.

We will next prove that the pressure is convex. Consider the convex family of hamiltonians $H_\Lambda(\sigma_\Lambda) + V_\Lambda(\sigma_\Lambda)$ where $H_\Lambda(\sigma_\Lambda)$ is defined in (2.3.2.13) and $V_\Lambda(\sigma_\Lambda) := \sum_{\tau_x\Delta\subset\Lambda} \tau_x v(\sigma_\Lambda)$, where v is a cylindrical function on Δ and the sum is over all translates of Δ which are in Λ. Call $Z_\Lambda(v)$ the corresponding partition function and, given any van Hove sequence, let

$$\pi(v) = \lim_{\Lambda\to\mathbb{Z}^d} \pi_\Lambda(v), \qquad \pi_\Lambda(v) := \frac{\log Z_\Lambda(v)}{|\Lambda|}. \tag{2.3.5.3}$$

The existence of the above limit is proved as in Sect. 2.3.3 and it is omitted.

Theorem 2.3.5.1 *The function $\pi(v)$ defined in (2.3.5.3) is convex, i.e. for any $a \in [0,1]$*

$$\pi\big(av^{(1)} + (1-a)v^{(2)}\big) = a\pi(v^{(1)}) + (1-a)\pi(v^{(2)}). \tag{2.3.5.4}$$

Proof Let Λ be a bounded region and

$$v = av^{(1)} + (1-a)v^{(2)}, \qquad a \in [0,1].$$

Then calling $H_\Lambda^{(1)}(\sigma_\Lambda)$ and $H_\Lambda^{(2)}(\sigma_\Lambda)$ the energies with $v^{(1)}$ and $v^{(2)}$, respectively,

$$\begin{aligned} Z_\Lambda(v) &= \sum_{\sigma_\Lambda} e^{-\beta a H_\Lambda^{(1)}(\sigma_\Lambda)} e^{-\beta(1-a)H^{(2)}(\sigma_\Lambda)} \\ &\le \Big[\sum_{\sigma_\Lambda\in\mathcal{X}_\Lambda} \big(e^{-\beta a H^{(1)}(\sigma_\Lambda)}\big)^p\Big]^{1/p} \Big[\sum_{\sigma_\Lambda\in\mathcal{X}_\Lambda} \big(e^{-\beta(1-a)H^{(2)}(\sigma_\Lambda)}\big)^q\Big]^{1/q}, \end{aligned}$$

with $1/p + 1/q = 1$. By choosing $p = a^{-1}$ and $q = (1-a)^{-1}$ we get

$$Z_\Lambda(v) \le Z_\Lambda(v^{(1)})^a \, Z_\Lambda(v^{(2)})^{1-a},$$

which yields $\pi_\Lambda(v) \le a\pi_\Lambda(v^{(1)}) + (1-a)\pi_\Lambda(v^{(2)})$. Then π_Λ is convex and $\pi(U)$ as well because of the limit of convex functions; see the paragraph "Properties of convex functions" in the next subsection. □

By Theorem 2.3.5.1, βP_β is a convex function of β and (2.3.5.2) is proved.

2.3.6 Properties of convex functions

We recall here some properties of convex functions on $\mathbb{R}$ which will often be used in the sequel, referring for their proofs to Chap. I.3 in Simon's book on Statistical Mechanics [200].

- *Definition.* $f(x)$, $x \in \mathbb{R}$, is convex if $f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$ for all x and y and $\alpha \in [0, 1]$.
- *Differentiability.* If f is convex, f is differentiable at all but countably many points, the right and left derivatives, $D^+ f$ and $D^- f$, exist everywhere and

$$D^- f(x) \le D^+ f(x) \le D^- f(y) \le D^+ f(y), \quad x < y. \tag{2.3.6.1}$$

- *Limits of convex functions.* If f_n is a sequence of convex functions which converges point-wise to f, then f is convex and for any x

$$D^- f(x) \le \liminf D^- f_n(x) \le \limsup D^+ f_n(x) \le D^+ f(x). \tag{2.3.6.2}$$

- *Legendre transforms.* Let $f(x)$ be a convex function, $x \in \mathbb{R}$. Its Legendre transform $g(p)$, $p \in \mathbb{R}$, is

$$g(p) = \sup_x \{xy - f(x)\}. \tag{2.3.6.3}$$

- *Properties of the Legendre transform.* The Legendre transform $g(p)$ of a function $f(x)$ is convex and, if $f(x)$ is convex, then

$$f(x) = \sup_p \{xy - g(p)\}. \tag{2.3.6.4}$$

  In general, the Legendre transform h of the Legendre transform g of f is $h = CEf$ the convex envelope of f, namely the largest convex function $\le f$.
- *Conjugate pairs.* Let f be a convex function and g its Legendre transform. Then x and p are conjugate if

$$p \in [D^- f(x), D^+ f(x)] \text{ if and only if } g(p) = px - f(x), \tag{2.3.6.5}$$

  and

$$\text{if } p \in [D^- f(x), D^+ f(x)] \text{ then } x \in [D^- g(p), D^+ g(p)]. \tag{2.3.6.6}$$

Notice that if $D^- f(x) < D^+ f(x)$ then $g(p)$ is linear in $[D^- f(x), D^+ f(x)]$. The Legendre transform has the following geometric interpretation (see Fig. 2.1): for each p consider all the straight lines $px + c$, all with the same slope p, which are not above the graph of f, namely the set of all c so that $f(x) - [px + c] \ge 0$ for all x. Hence $c \le \inf_x \{f(x) - px\}$. Call c^* the sup of all such c, then $g(p) = -c^*$ and hence $g(p)$ is minus the intersection with the vertical axis of the highest line with slope p which is $\le f(x)$ at all x.

2.3.7 Concavity of the Boltzmann entropy

With D below denoting derivative with respect to β, define

$$\pi_\beta = \beta P_\beta, \quad \mathcal{A} = \{\beta > 0 : D^- \pi_\beta = D^+ \pi_\beta = D\pi_\beta\}, \tag{2.3.7.1}$$

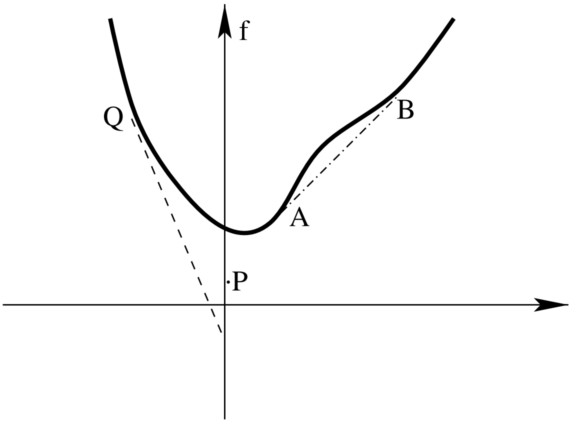

Fig. 2.1 The *dashed line* is tangent at the graph of f at Q, and the value of the Legendre transform is P. The convex envelope is obtained by joining the points A and B as shown

so that $\mathcal{A}$ covers the positive axis but for countably many points; let

$$\mathcal{G}^0_\beta = \text{ the set of translational invariant DLR measures at } \beta. \tag{2.3.7.2}$$

Proposition 2.3.7.1 *For any* $\beta > 0$ *and for any* $\mu \in \mathcal{G}^0_\beta$, $\mathcal{G}^0_\beta$ *as in* (2.3.7.2),

$$D^-\pi_\beta \le -\mu(u) \le D^+\pi_\beta. \tag{2.3.7.3}$$

If $\beta \in \mathcal{A}$, *see* (2.3.7.1), *then*

$$E := -D\pi_\beta \in \mathcal{E}_{\text{erg}} \quad \textit{and} \quad \textit{for any } \mu \in \mathcal{G}^0_\beta, \quad \mu(u) = E = -D\pi_\beta. \tag{2.3.7.4}$$

Proof Calling $\pi_{\beta,\Lambda}(\sigma_{\Lambda^c}) := Z_{\beta,\Lambda}(\sigma_{\Lambda^c})/|\Lambda|$, we claim that for any probability ν on $\mathcal{X}$ and any van Hove sequence $\Lambda \to \mathbb{Z}^d$,

$$\lim_{\Lambda\to\mathbb{Z}^d} \int d\nu(\sigma)\, D\pi_{\beta,\Lambda}(\sigma_{\Lambda^c}) \in [D^-\pi_\beta, D^+\pi_\beta]. \tag{2.3.7.5}$$

The proof of this claim is as follows. Since the pressure is independent of the b.c., see Sect. 2.3.3, and since $\pi_{\beta,\Lambda}(\sigma_{\Lambda^c})$ is uniformly bounded, see (2.3.3.5), by the Lebesgue dominated convergence theorem, $\lim_{\Lambda\to\mathbb{Z}^d} \int \nu(d\sigma)\, \pi_{\beta,\Lambda}(\sigma_{\Lambda^c}) = \pi_\beta$. Thus $\int \nu(d\sigma)\, \pi_{\beta,\Lambda}(\sigma_{\Lambda^c})$ is a sequence of convex, differentiable functions of β which converges to π_β and (2.3.7.5) follows from (2.3.6.2). Thus, the claim is proved.

By explicit computation,

$$D\pi_{\beta,\Lambda}(\sigma_{\Lambda^c}) = -\frac{1}{|\Lambda|} \sum_{x\in\Lambda} \frac{e^{-\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}}{Z_{\beta,\Lambda}(\sigma_{\Lambda^c})} H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}). \tag{2.3.7.6}$$

Rewrite (2.3.7.5) with (2.3.7.6) and use $\nu = \mu \in \mathcal{G}_\beta^0$. By the DLR property we then have

$$\lim_{\Lambda \to \mathbb{Z}^d} \mu\left(- \frac{H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}{|\Lambda|} \right) \in [D^- \pi_\beta, D^+ \pi_\beta]. \tag{2.3.7.7}$$

By Lemma 2.3.4.5 given any $\epsilon > 0$ for all Λ large enough,

$$\left| \mu\left(- \frac{H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}{|\Lambda|} \right) + \frac{1}{|\Lambda|} \mu\left(\sum_{x\in\Lambda} \tau_x u \right) \right| \le 2\epsilon.$$

Since $\mu \in \mathcal{G}_\beta^0$, μ is translation invariant and $\mu(\tau_x u) = \mu(u)$, by (2.3.7.7)

$$\mu(u) \in [D^- \pi_\beta - 2\epsilon, D^+ \pi_\beta + 2\epsilon],$$

which by the arbitrariness of ϵ proves (2.3.7.3).

The second statement in (2.3.7.4) is a corollary of (2.3.7.3) and we only have to prove that $\mu(u) \in \mathcal{E}_{\text{erg}}$, which follows because $\mathcal{G}_\beta^0$ contains at least one ergodic measure, Theorem 2.2.7.2. □

Proposition 2.3.7.2 *Let $\beta \notin \mathcal{A}$; then there are $\mu_\pm \in \mathcal{G}_\beta^0$ so that*

$$-\mu_\pm(u) = D^\pm \pi_\beta. \tag{2.3.7.8}$$

Proof Since π_β is convex there exists an increasing sequence $\beta_n \in \mathcal{A}$ which converges to β as $n \to \infty$. By (2.3.7.4) there are ergodic measures μ_n which are DLR with respect to β_n such that $\mu_n(u) = E_n = -D\pi_{\beta_n}$. By compactness there is a subsequence n_k so that μ_{n_k} converges weakly to some probability measure μ.

We claim that μ is in $\mathcal{G}_\beta^0$. Let f be any cylindrical function, since $\mu_{n_k}(\tau_x f) = \mu_{n_k}(f)$, then also $\mu(\tau_x f) = \mu(f)$, hence μ is translational invariant. We fix arbitrarily a bounded set Λ and $\epsilon > 0$ and call $G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)$ the conditional Gibbs expectation of f at β with b.c. σ_{Λ^c}. Then, by (2.3.3.3), for all n_k large enough

$$\sup_{\sigma_{\Lambda^c}} |G_{\Lambda,\beta_{n_k},\sigma_{\Lambda^c}}(f) - G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)| \le \epsilon. \tag{2.3.7.9}$$

Moreover, since both $f(\sigma)$ and $G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)$ are continuous functions of σ, for all n_k large enough

$$\begin{aligned} \left|\mu\big(G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)\big) - \mu_{n_k}\big(G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)\big)\right| &< \epsilon, \\ \left|\mu(f) - \mu_{n_k}(f)\right| &< \epsilon. \end{aligned} \tag{2.3.7.10}$$

By (2.3.7.9)–(2.3.7.11)

$$\begin{aligned} &\left|\mu\big(G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)\big) - \mu(f)\right| \\ &\quad \le \left|\mu_{n_k}\big(G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)\big) - \mu_{n_k}(f)\right| + 2\epsilon \\ &\quad \le \left|\mu_{n_k}\big(G_{\Lambda,\beta_{n_k},\sigma_{\Lambda^c}}(f)\big) - \mu_{n_k}(f)\right| + 3\epsilon. \end{aligned} \tag{2.3.7.11}$$

Since μ_{n_k} is DLR at β_{n_k}, $\mu_{n_k}(G_{\Lambda,\beta_{n_k},\sigma_{\Lambda^c}}(f)) = \mu_{n_k}(f)$; hence, by the arbitrariness of ϵ, $\mu(G_{\Lambda,\beta,\sigma_{\Lambda^c}}(f)) = \mu(f)$. As the argument applies to any cylindrical f and any bounded Λ, we may conclude the proof of the claim (namely that $\mu \in \mathcal{G}^0_\beta$). By (2.3.6.2),

$$\lim D\pi_{\beta_{n_k}} \le D^-\pi_\beta. \tag{2.3.7.12}$$

By Proposition 2.3.7.1, $\mu_{n_k}(u) = -D\pi_{\beta_{n_k}}$, so that

$$-\mu(u) \le D^-\pi_\beta. \tag{2.3.7.13}$$

By (2.3.7.3) it then follows that $-\mu(u) = D^-\pi_\beta$. By a completely analogous argument we conclude that there is a translational invariant DLR measure ν at inverse temperature β so that $-\nu(u) = D^+\pi_\beta$. □

Corollary 2.3.7.3 *If $\beta \notin \mathcal{A}$ there are ergodic DLR measures $\mu_\pm$ so that (2.3.7.8) holds. Moreover, for any E such that $-E \in [D^-\pi_\beta, D^+\pi_\beta]$ there is $\mu \in \mathcal{G}^0_\beta$ such that $\mu(u) = E$. The converse statement has been proved in (2.3.7.3).*

Proof By Proposition 2.3.7.2, there are translational invariant DLR measures $\nu_\pm$ for which (2.3.7.8) holds. If ν_+ is not ergodic, then by the ergodic decomposition, (2.2.7.2),

$$-D^+\pi_\beta = \nu_+(u) = \int_{\mathcal{X}^0_{gg}\cap\mathcal{G}^0} \mathcal{A}_\sigma(u)\,\nu_+(d\sigma), \tag{2.3.7.14}$$

where $\mathcal{A}_\sigma$ is an ergodic (as $\sigma \in \mathcal{X}^0_{gg}$) DLR measure (as $\sigma \in \mathcal{G}^0$). By (2.3.7.3) $-\mathcal{A}_\sigma(u) \le D^+\pi_\beta$; then by (2.3.7.14) the set of $\sigma \in \mathcal{X}^0_{gg} \cap \mathcal{G}^0$ such that $-\mathcal{A}_\sigma(u) = D^+\pi_\beta$ has ν_+ measure equal to 1 and it is therefore non-empty. The same argument applies to $D^-\pi_\beta$. Finally, if $-E \in [D^-\pi_\beta, D^+\pi_\beta]$, $-E = aD^-\pi_\beta + (1-a)D^+\pi_\beta = a\nu_+(u) + (1-a)\nu_-(u)$, $a \in [0,1]$, hence $-E = \mu(u)$, $\mu = a\nu_+ + (1-a)\nu_-$ and the corollary is proved. □

In summary, we have proved so far that

$$\{-\mu(u), \mu \in \mathcal{G}^0_\beta\} = [D^-\pi_\beta, D^+\pi_\beta], \quad \text{for any } \beta > 0, \tag{2.3.7.15}$$

which implies

$$\mathcal{E} = \Big\{-\mu(u), \mu \in \bigcup_{\beta>0} \mathcal{G}^0_\beta\Big\} = \bigcup_{\beta>0} [D^-\pi_\beta, D^+\pi_\beta]. \tag{2.3.7.16}$$

By the above corollary we also know that $\mathcal{E}_{\text{erg}}$ contains the endpoints of $[D^-\pi_\beta, D^+\pi_\beta]$ as well as those values $E = -D\pi_\beta$ for which $D\pi_\beta$ exists: recall that for $E \in \mathcal{E}_{\text{erg}}$ we can use Proposition 2.3.4.2.

Proposition 2.3.7.4 *If $E \in \mathcal{E}_{\rm erg}$ then*

$$S^{\rm td}(E) = S(E), \quad S^{\rm td}(E) \text{ as in (2.3.5.1)}. \tag{2.3.7.17}$$

Proof If $E \in \mathcal{E}_{\rm erg}$ there is β and an ergodic DLR measure μ at β such that $\mu(u) = E$. By Proposition 2.3.4.2, $S(E) = \beta(P_\beta + E)$. By Proposition 2.3.7.1, $-E \in [D^-\pi_\beta, D^+\pi_\beta]$ and therefore by (2.3.6.5), $-S^{\rm td}(E) = \beta(-E) - \pi_\beta$, hence (2.3.7.17). □

Theorem 2.3.7.5 *For any $-E \in \mathcal{E}$ the limit (2.3.4.1) exists and*

$$S(E) = S^{\rm td}(E), \quad \text{for all } -E \in \mathcal{E} \tag{2.3.7.18}$$

($S^{\rm td}(E)$ as in (2.3.5.1)).

Proof We already know that $S(E) = S^{\rm td}(E)$ for all $E \in \mathcal{E}_{\rm erg}$, and hence for all E conjugate to $\beta \in \mathcal{A}$. By (2.3.6.5) for any E conjugate to $\beta \notin \mathcal{A}$,

$$-S^{\rm td}(E) = \beta(-E) - \pi_\beta, \quad -E \in [-E^-, -E^+], \; -E^\pm = D^\pm \pi_\beta. \tag{2.3.7.19}$$

By Corollary 2.3.7.3, $E^\pm \in \mathcal{E}_{\rm erg}$ and by Proposition 2.3.7.4 $S^{\rm td}(E^\pm) = S(E^\pm)$. Theorem 2.3.7.5 then follows from the following proposition. □

Proposition 2.3.7.6 *Suppose that there are two ergodic DLR measures, μ_1 and μ_2 at the same inverse temperature β so that $\mu_1(u) = E_1 < \mu_2(u) = E_2$. Then for any E in $[E_1, E_2]$ the limit in (2.3.4.1) exists and*

$$S(E) - \beta E = \beta P_\beta. \tag{2.3.7.20}$$

Proof Fix $E \in (E_1, E_2)$ and let $a \in (0, 1)$ be such that $E = aE_1 + (1 - a)E_2$. Let $\Lambda \to \mathbb{Z}^d$ be a van Hove sequence; given any cube Δ and a partition into translates of Δ, call $\Delta(i)$, $i = 1, \dots, N$, the cubes of the partition contained in Λ; Λ' is their union. As $\Lambda \to \mathbb{Z}^d$, $N \to \infty$ and $|1 - \frac{N|\Delta|}{|\Lambda|}| \to 0$. We then choose for any Λ a positive integer $n < N$ so that

$$\lim_{\Lambda \to \mathbb{Z}^d} \frac{n}{N} = a, \qquad \lim_{\Lambda \to \mathbb{Z}^d} \frac{N - n}{N} = 1 - a. \tag{2.3.7.21}$$

Thus for any $\zeta > 0$, as soon as $|\Lambda|$ is large enough

$$\left| \frac{n|\Delta|}{|\Lambda|} - a \right| < \zeta, \qquad \left| \frac{(N - n)|\Delta|}{|\Lambda|} - (1 - a) \right| < \zeta.$$

We can now bound $N_{E,\Lambda,\delta}$ from below as follows. We choose $\delta' < \delta$ (its value will be specified later) and consider all σ_Λ such that $|H_{\Delta(i)}(\sigma_{\Delta(i)}) - |\Delta|E_1| < \delta'|\Delta|$ for

$i = 1, \dots, n$ and $|H_{\Delta(i)}(\sigma_{\Delta(i)}) - |\Delta| E_2| < \delta' |\Delta|$ for $i = n+1, \dots, N$. By (2.3.4.11), given any $\epsilon > 0$ if $|\Delta|$ is large enough,

$$\left| H_\Lambda(\sigma_\Lambda) - \sum_{\Delta(i) \subset \Lambda_0} H_{\Delta(i)}(\sigma_{\Delta(i)}) \right| \le c|\Lambda \setminus \Lambda'| + \frac{\epsilon}{4}|\Lambda|,$$

and therefore if ζ, ϵ and δ' are small enough and $|\Lambda|$ large enough,

$$|H_\Lambda(\sigma_\Lambda) - E|\Lambda|\,| \le c|\Lambda \setminus \Lambda'| + \frac{\epsilon}{4}|\Lambda| + \delta'|\Lambda| + \zeta|\Lambda| < \delta|\Lambda|. \tag{2.3.7.22}$$

We have

$$N_{E,\Lambda,\delta} \ge \left\{ \prod_{i=1}^{n} N_{E_1,\Delta(i),\delta'} \right\} \left\{ \prod_{i=n+1}^{N} N_{E_2,\Delta(i),\delta'} \right\},$$

so that, by (2.3.7.21),

$$\liminf_{\Lambda \to \mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \ge a \frac{\log N_{E_1,\Delta,\delta'}}{|\Delta|} + (1-a) \frac{\log N_{E_2,\Delta,\delta'}}{|\Delta|}.$$

We next let $|\Delta| \to \infty$ and then $\delta' \to 0$ and, since E_1 and E_2 are in $\mathcal{E}_{\rm erg}$,

$$\liminf_{\delta \to 0} \liminf_{\Lambda \to \mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \ge a S(E_1) + (1-a) S(E_2).$$

By (2.3.4.3) the r.h.s. is equal to

$$\beta\{a[P_\beta + E_1] + (1-a)[P_\beta + E_2]\} = \beta\{P_\beta + E\}.$$

By taking $\Lambda \to \mathbb{Z}^d$ in (2.3.4.14) and then δ and ϵ to 0, we get the upper bound $\limsup_{\delta \to 0} \limsup_{\Lambda \to \mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \le \beta\{P_\beta + E\}$. □

2.3.8 Variational principles and equivalence of ensembles

In this subsection we discuss three variational principles. The first two, see Theorem 2.3.8.1, are just a reformulation of those stated in Theorem 2.3.4.1. The third one, see Theorem 2.3.9.1, is instead of a rather different nature.

Theorem 2.3.8.1 *For any $\beta > 0$ and with P_β and $S(E)$ as in Theorem 2.3.4.1 of Sect. 2.3.4,*

$$P_\beta = \sup_{-E \in \mathcal{E}} \{-(E - \beta^{-1} S(E))\}, \quad \mathcal{E} \textit{ as in } (2.3.7.16), \tag{2.3.8.1}$$

and for any $-E \in \mathcal{E}$

$$S(E) = -\sup_{\beta > 0} \{-\beta E - \beta P_\beta\}. \tag{2.3.8.2}$$

Proof As a consequence of the identification $S^{\mathrm{td}}(E) = S(E)$, $E \in \mathcal{E}_{\mathrm{allw}}$, see (2.3.7.18), we get (2.3.8.1) from (2.3.5.2) and (2.3.8.2) from (2.3.5.1). □

Remarks

The Gibbs formula and the theory of DLR states have been heavily used in the proofs but they should be regarded just as auxiliary tools. The only physical assumption in the theory has been the Boltzmann hypothesis. Equation (2.3.8.1) may also be proved directly by arguments similar to those in Sect. 2.3.3 as traditionally is done in texts on statistical mechanics. As we already had available the theory of DLR measures, it was simpler to proceed the way we did.

The variational principle (2.3.8.2) may be viewed as a method for determining the entropy in terms of the grand canonical partition function that more easily is handled because the energy constraint is dropped.

Below are some consequences of the above variational principles which have physical relevance.

Theorem 2.3.8.2 *$S(E)$ is an increasing function of E; its right and left derivatives exist everywhere and $\beta \in [D^-S(E), D^+S(E)]$ if and only if $-E \in [D^-\pi_\beta, D^+\pi_\beta]$. The derivative of $S(E)$ exists at all but countably many points and if $DS(E)$ exists*

$$DS(E) = \beta, \quad \beta : -E \in [D^-(\beta P_\beta), D^+(\beta P_\beta)], \tag{2.3.8.3}$$

and $S(E) = \beta E + \beta P_\beta$.

Proof Let $s(e) = -S(-e)$, $e \in \mathcal{E}$, and $\pi_\beta = \beta P_\beta$. Since $s(e)$ is the Legendre transform of π_β, see (2.3.8.2), it is convex and hence its derivative $Ds(e)$ exists at all but countably many $e \in \mathcal{E}$; we denote such a set $\mathcal{E}'$. Recalling from (2.3.7.3) that e is related to β by the relation $e \in [D^-\pi_\beta, D^+\pi_\beta]$, then if $e \in \mathcal{E}'$, $Ds(e)$ must be equal to β by (2.3.6.6). Hence

$$s(e'') - s(e') = \int_{[e',e'']\cap\mathcal{E}'} Ds(e)de > 0, \quad e'' > e'.$$

The claimed relation between e and β is a particular case of (2.3.6.5)–(2.3.6.6). □

Remarks

Theorem 2.3.8.2 establishes the basic principle of thermodynamics that the derivative of the entropy with respect to the energy is the inverse temperature and since Theorem 2.3.8.2 proves that the entropy is an increasing function of the energy, the temperatures are positive as they should (if $S(E)$ is not differentiable, the values of β are identified with the slope of the tangent lines to the graph of S hence $\beta \in [D^-s(e), D^+s(e)]$ and therefore $\beta > 0$ in general).

There are two main questions though, which need to be clarified: where is the limitation to $-E \in \mathcal{E}$ coming from? And what happens if we take $-E \notin \mathcal{E}$?

Theorem 2.3.8.3 *We have*

$$\sup\{E : -E \in \mathcal{E}\} = 0, \qquad \sup_{-E\in\mathcal{E}} S(E) = \ln 2. \tag{2.3.8.4}$$

Moreover, for any increasing sequence Δ of cubes which invades $\mathbb{Z}^d$, the following limit exists:

$$\lim_{\Delta\to\mathbb{Z}^d} \frac{1}{|\Delta|} \min_{\sigma_\Delta} H_\Delta(\sigma_\Delta) =: e_{\min}, \tag{2.3.8.5}$$

and the set $\mathcal{E}$ in Theorem 2.3.4.1 is

$$\mathcal{E} = (e_{\min}, 0). \tag{2.3.8.6}$$

Proof Recall from (2.3.7.16) that $\mathcal{E} = \bigcup_{\beta>0}[D^-\pi_\beta, D^+\pi_\beta]$; then by (2.3.6.1)

$$\sup\{E, -E \in \mathcal{E}\} = -\lim_{\beta\to 0} D^-\pi_\beta.$$

We can take the limit along a sequence β_n where $D\pi_{\beta_n}$ exists. Then there are measures $\mu_n \in \mathcal{C}^0_{\beta_n}$ so that $-D\pi_{\beta_n} = \mu_n(u)$. Using the DLR property

$$\mu_n(\sigma(0)) = \mu_n\big(\tanh\{\beta_n\kappa(\sigma)\}\big), \quad \kappa(\sigma) = h + \sum_{x\neq 0} J(0,x)\sigma(x),$$

and since $\sup_\sigma |\kappa(\sigma)| \le c$, $\lim_{n\to\infty} \mu_n(\sigma(0)) = 0$. An analogous argument proves that for any $x \neq 0$, $\lim_{n\to\infty} \mu_n(\sigma(0)\sigma(x)) = 0$ and the first half of (2.3.8.4) is proved.

By Theorem 2.3.8.2, $S(E)$ is an increasing function of E; then by the first half of (2.3.8.4), $\sup_{-E\in\mathcal{E}} S(E) = \lim_{E\nearrow 0} S(E)$. As before we take the limit along a sequence E_n such that $D\pi_{\beta_n} = -E_n$ and $\beta_n \to 0$. By (2.3.8.2), $S(E_n) = \beta_n E_n + \pi_{\beta_n}$ and the second half of (2.3.8.4) is then a consequence of the following inequality: there is $c > 0$ so that for any bounded Λ and any σ_{Λ^c}, $e^{-\beta c|\Lambda|}2^{|\Lambda|} \le Z_{\beta,\Lambda}(\sigma_{\Lambda^c}) \le e^{\beta c|\Lambda|}2^{|\Lambda|}$, which thus proves the second half of (2.3.8.4).

The proof of (2.3.8.5) now follows. Since $|H_\Delta(\sigma_\Delta)| \le c|\Delta|$, c the sup norm of u, there is a sub-sequence $\Delta_n \to \mathbb{Z}^d$ such that the limit below exists,

$$\lim_{\Delta_n\to\mathbb{Z}^d} \frac{1}{|\Delta_n|} \min_{\sigma_{\Delta_n}} H_{\Delta_n}(\sigma_{\Delta_n}) =: e. \tag{2.3.8.7}$$

By Lemma 2.3.4.4, given any $\epsilon > 0$ if n is large enough the following holds. Given any Δ, call $\Delta_n(i)$ the cubes of a partition into translates of Δ_n which are in Δ, and Δ' is their union. Then

$$\Big| H_\Delta(\sigma_\Delta) - \sum_i H_{\Delta_n(i)}(\sigma_{\Delta_n(i)}) \Big| \le \epsilon|\Delta| + c|\Delta \setminus \Delta'|. \tag{2.3.8.8}$$

Thus

$$\liminf_{\Delta\to\mathbb{Z}^d}\left|\frac{1}{|\Delta|}\min_{\sigma_\Delta} H_\Delta(\sigma_\Delta)-e\right|\le\epsilon,$$

with the same inequality holding for the limsup; hence (2.3.8.5). The proof of (2.3.8.6) is omitted. □

Remarks

Notice that for any E and not only $-E\in\mathcal{E}$,

$$\lim_{\delta\to 0}\limsup_{\Lambda\to\mathbb{Z}^d}\frac{\log N_{E,\Lambda,\delta}}{|\Lambda|}\le\log 2, \tag{2.3.8.9}$$

as $2^{|\Lambda|}=\mathrm{card}(\Lambda)$. Thus, by Theorem 2.3.8.3, if $E>0$, i.e. $-E\notin\mathcal{E}$, then the entropy $S(E)$ (which can be proved to be well defined) is smaller than the sup of the entropy when $-E\in\mathcal{E}$. Thus, the thermodynamic law that entropy is an increasing function of E is violated, namely temperatures are negative at $E>0$! The physical relevance of states with negative temperatures cannot be completely ruled out in systems where the energy density is bounded, as in our Ising model. Indeed if we are able to thermally isolate the system and to prepare an initial state with high energy, then it must relax to some limit, and the limit state will have by conservation of energy the same high energy as the initial one. There are indeed practical applications of such considerations. Negative temperatures are related to positive temperatures for the hamiltonian $H'=-H$, which is in general limited to spin systems as it may become pathological when variables are unbounded: the "stability condition" on the hamiltonian required for the partition function to exist is in general not satisfied after the transformation $H\to -H$.

2.3.9 A variational principle for measures

We have seen that the pressure and the thermodynamics of the Ising model can be derived under the assumption only that entropy is the log of the number of states at the given energy. Using such an assumption we have then identified the pressure in terms of the log of the partition function. On the other hand, as the partition function is the normalization factor in the Gibbs formula, it seems natural to try to push the argument further to derive the full Gibbs formula as well. The variational principle in Theorem 2.3.8.1 states that the pressure P_β, at a given inverse temperature β, is obtained by maximizing the entropy; see (2.3.8.1). But Propositions 2.3.7.1 and 2.3.7.2 complement the result by stating that the maximizing energy E is indeed the average value $\mu(u)$, with u as in (2.2.7.3), of a translational invariant measure μ DLR at inverse temperature β. Thus, the maximizer gives information on μ; it specifies the average of u. If we could repeat the argument with other functions, we would then identify other expectations of μ, and this may eventually lead to the identification of μ itself. Such a program will be carried out by a generalization of the notion of entropy which will be extended to the "entropy of probability measures," and then the assumption that equilibrium is reached when the entropy is maximal will identify the translational invariant DLR measures.

We have already found in Sect. 2.3.1 an expression for the entropy $S(\mu)$ of a measure μ, given by the formula (2.3.1.8). $S(\mu)$ in Sect. 2.3.1 counts how many "typical messages" are emitted by an independent source μ (of sequences of symbols); namely $S(\mu)$ gives a measure of the storage capacity of the system (it tells how many messages can be stored by μ). The assumption of independence used in Sect. 2.3.1 can be greatly relaxed. The Shannon–MacMillan–Breiman theorem states that if μ is ergodic (on $\{-1,1\}^{\mathbb{Z}^d}$ in our specific application), then there is a unique number $S(\mu)$ such that the following holds. Given any $\epsilon > 0$, for any cube Δ large enough we can split $\{-1,1\}^{\Delta}$ into two sets, A_1 and A_2, so that

$$\mu(A_1) < \epsilon, \qquad \left|\frac{\log \operatorname{card} A_2}{|\Delta|} - S(\mu)\right| < \epsilon,$$

and moreover for any $\sigma_\Delta \in A_2$,

$$\left|\frac{\log \mu(\sigma_\Delta)}{|\Delta|} + S(\mu)\right| < \epsilon. \tag{2.3.9.1}$$

The number $S(\mu)$, called the [information] entropy of μ, is equal to

$$-\lim_{\Delta \to \mathbb{Z}^d} \frac{1}{|\Delta|} \sum_{\sigma_\Delta} \mu(\sigma_\Delta) \log \mu(\sigma_\Delta) =: S(\mu). \tag{2.3.9.2}$$

If μ is our ergodic DLR measure of Sect. 2.3.2 at inverse temperature β, then μ concentrates on configurations with energy $(E \pm \delta)|\Delta|$, $E = \mu(u)$, the complement having vanishing measure as $\Delta \to \mathbb{Z}^d$. We have seen that the number of configurations with energy $(E \pm \delta)|\Delta|$ is $N_{E,\Delta,\delta} \approx e^{S(E)|\Delta|}$ and moreover for any configuration σ_Δ such that $H_\Delta(\sigma_\Delta) \in (E \pm \delta)|\Delta|$,

$$\frac{1}{|\Delta|} \log \mu\big(\{\sigma : \sigma \restriction \Delta = \sigma_\Delta\}\big) \approx -\beta(E \pm \delta \pm \epsilon) + C,$$

C a constant, with $\epsilon|\Delta|$ bounding the interaction with Δ^c. By taking $|\Delta| \to \infty$ this proves that $S(\mu) = S(E)$. Thus the notion of information entropy $S(\mu)$ reduces to the $S(E)$ entropy when $E \in \mathcal{E}_{\mathrm{erg}}$. We can then accept $S(\mu)$ as the entropy of a measure without running into a conceptual conflict with our previous considerations.

Having $S(\mu)$, we can now invoke thermodynamic principles to find the equilibrium states. We start by considering the simple case of a finite region Λ. We suppose that equilibrium states are maximizers of the entropy among all states with a given energy. In the grand-canonical ensemble such a constraint is only imposed in the average by requiring that a measure μ representative of a state should verify

$$\sum_{\sigma_\Lambda} \mu(\sigma_\Lambda) H_\Lambda(\sigma_\Lambda) = E|\Lambda|. \tag{2.3.9.3}$$

The equilibrium measure is then defined as the maximizer of the entropy $S_\Lambda(\mu)$ given by (2.3.9.2) (with $\Delta = \Lambda$ and without taking the limit) under the constraint (2.3.9.3):

$$\left\{ S_\Lambda(\mu), \mu : \sum_{\sigma_\Lambda} \mu(\sigma_\Lambda) H(\sigma_\Lambda) = E|\Lambda| \right\} \longrightarrow \text{ maximum.} \tag{2.3.9.4}$$

By using Lagrange multipliers we need to find the critical points of

$$S_\Lambda(\mu) - \beta \sum_{\sigma_\Lambda} \mu(\sigma_\Lambda) H(\sigma_\Lambda) - \lambda \sum_{\sigma_\Lambda} \mu(\sigma_\Lambda), \tag{2.3.9.5}$$

with β the Lagrange multiplier for the energy constraint and λ for the normalization of μ as a probability. By an elementary computation, we see that the critical point is unique, and it is given by G_Λ, the Gibbs measure at the inverse temperature β, which is a maximizer of (2.3.9.4) by the concavity of $S_\Lambda(\mu)$. Notice also that the variational problem

$$\left\{ S_\Lambda(\mu) - \frac{\beta}{|\Lambda|} \sum_{\sigma_\Lambda} \mu(\sigma_\Lambda) H(\sigma_\Lambda) \right\} \longrightarrow \text{ maximum} \tag{2.3.9.6}$$

is achieved at G_Λ and that the maximum is equal to $\beta P_{\beta,\Lambda}$, $P_{\beta,\Lambda}$ being the finite volume pressure, $\beta P_{\beta,\Lambda} = \log Z_\Lambda / |\Lambda|$, namely

$$\sup_\mu \left\{ S_\Lambda(\mu) - \frac{\beta}{|\Lambda|} \mu(H_\Lambda) \right\} = \beta P_{\beta,\Lambda} = \left\{ S_\Lambda(G_\Lambda) - \frac{\beta}{|\Lambda|} G_\Lambda(H_\Lambda) \right\}. \tag{2.3.9.7}$$

Equation (2.3.9.7) extends to infinite volumes with the finite volume Gibbs measures replaced by the translational invariant DLR measures. We only state the result in our setup (its validity being more general) and refer to the literature for the proof; see for instance Theorems III.4.3, III.4.5 and III.4.9 in Simon's book on statistical mechanics [200].

Theorem 2.3.9.1 *Let $\Delta \to \mathbb{Z}^d$ be an increasing sequence of cubes and call $\mathcal{M}_0$ the set of all translational invariant probabilities on $\{-1, 1\}^{\mathbb{Z}^d}$. Then we have*

(i) *Existence of entropy: for any $\mu \in \mathcal{M}_0$ the limit (2.3.9.2) exists.*
(ii) *Gibbs variational principle: For any $\mu \in \mathcal{M}_0$ and with u defined in (2.2.7.3),*

$$P_\beta = \sup_{\mu \in \mathcal{M}_0} \{\beta^{-1} S(\mu) - \mu(u)\}. \tag{2.3.9.8}$$

(iii) *Ruelle's theorem*: $\mu \in \mathcal{M}_0$ *is in* $\mathcal{G}_\beta^0$ *if and only if*

$$P_\beta = \beta^{-1} S(\mu) - \mu(u).$$

2.4 Thermodynamics and DLR measures

In this section we complete the analysis of the thermodynamics of the Ising model by proving that the translational invariant DLR measures can be identified as the functionals tangent to the graph of the pressure regarded as a function of the interaction potential. We thus have an alternative way to derive the Gibbs formula from the Boltzmann hypothesis, besides the one in Theorem 2.3.9.1, based on an extended notion of entropy of measures to which thermodynamic principles are then applied. We will conclude the section by briefly discussing (proofs are omitted) a characterization of the translational invariant DLR measures as "tangent functional to the pressure" and the relation between large deviations for Gibbs measures and thermodynamics potentials.

2.4.1 Canonical ensemble and free energy

The Boltzmann hypothesis actually states that entropy is equal to the log of the number of states *available* to the system. Supposing that the energy is a prime integral, the phase space available to the system is the energy surface relative to the energy of the system. There could, however, be other prime integrals or we can imagine that by some external action on the system the phase space available to the system is reduced. The thermodynamic potentials are then modified and new order parameters come into play.

Definitions *Grand-canonical, canonical and micro-canonical ensembles*. We will suppose that the phase space available to the system is determined by the value of an intensive quantity (also called observable). The subset of the phase space where such a value is attained is the "*canonical ensemble*" (called *micro-canonical* when the observable is the energy). Usually the notion is applied to the case where the intensive quantity is the total number of particles or, in Ising systems, the total magnetization. In the case of particle systems "micro-canonical" usually refers to the ensemble where in which both energy and total number of particles are fixed. The *grand-canonical* ensemble is instead the unrestricted phase space.

The equivalence of ensembles is a property which states the equivalence of thermodynamics computed in the canonical ensemble and the thermodynamics computed in the grand-canonical ensemble after a suitable term has been added to the hamiltonian (which is referred to as the variable conjugate to the observable defining the canonical ensemble). We have proved that the micro-canonical ensemble in which energy is fixed gives the same thermodynamic potentials as the grand-canonical ensemble with temperature the variable conjugate to the energy, (2.3.8.1)–(2.3.8.2).

Suppose that the intensive quantity defining the canonical ensemble is

$$V_\Lambda(\sigma_\Lambda) = \sum_{x\in\mathbb{Z}^d:\tau_x\Delta\subset\Lambda} \tau_x v(\sigma_\Lambda), \tag{2.4.1.1}$$

where v is a cylindrical function in a bounded set $\Delta \subset \mathbb{Z}^d$, i.e. such that $v(\sigma)$ is independent of σ_{Δ^c}. The "order parameter" is then the observable v and the values w of its ergodic averages together with the energy E parameterize the equilibrium states of the system. In such a scenario the analogue of the Boltzmann hypothesis (2.3.2.3) is that the finite volume entropy density at the values E and w of the order parameters is

$$\begin{aligned} &S_{\Lambda,\delta}(E, w; v) \\ &:= \frac{1}{|\Lambda|} \log\left\{\sum_{\sigma_\Lambda} \mathbf{1}_{|V_\Lambda(\sigma_\Lambda)-|\Lambda|w|\le\delta|\Lambda|}\mathbf{1}_{|H_\Lambda(\sigma_\Lambda)-|\Lambda|E|\le\delta|\Lambda|}\right\}. \end{aligned} \tag{2.4.1.2}$$

We will study in the sequel a simplified scenario with only one order parameter; namely, we will suppose that the energy is not conserved, because the system exchanges energy with a "reservoir" at inverse temperature β (thermal walls), which thus fixes the temperature of the system at the value β^{-1}. We suppose however that $\sum_x \tau_x v$ is conserved and that the walls are impermeable to exchanges of this quantity, which is then the only order parameter for the system (as the temperature is fixed). In such a scenario the free energy rather than entropy is the relevant thermodynamic quantity and the analogue of (2.4.1.2) is

$$F_{\beta,\Lambda,\delta,\sigma_{\Lambda^c}}(w; v) := -\frac{1}{\beta|\Lambda|}\log\left\{\sum_{\sigma_\Lambda:|V_\Lambda(\sigma_\Lambda)-|\Lambda|w|\le\delta|\Lambda|} e^{-\beta H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c})}\right\}. \tag{2.4.1.3}$$

The minus sign is there because we consider here the free energy and not the pressure. The available phase space, i.e. the ensemble of configurations appearing in (2.4.1.3), is called the (δ-relaxed) "canonical ensemble" relative to the variable v.

There is some arbitrariness in the definition (2.4.1.3) regarding the boundary conditions, which is fortunately unimportant because any surface correction becomes negligible in the thermodynamic limit.

As in the previous section the crucial point is an equivalence of ensemble property which says that in the infinite volume limit it is equivalent to study the above problem or another one where we drop the constraint in the phase space and add an external field which forces $V_\Lambda(\sigma_\Lambda)$ to have the desired value (in the previous section where v was u, this was achieved by choosing properly β). With the addition of the external field the new energy with "zero boundary conditions" is

$$H_{\Lambda,\lambda,v}(\sigma_\Lambda) = H_\Lambda(\sigma_\Lambda) - \lambda V_\Lambda(\sigma_\Lambda) \tag{2.4.1.4}$$

(the minus sign in front of λ is just conventional), while if the boundary condition is σ_{Λ^c},

$$H_{\Lambda;\lambda,v}(\sigma_\Lambda|\sigma_{\Lambda^c}) = H_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) - \lambda V_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}), \tag{2.4.1.5}$$

where $V_\Lambda(\sigma_\Lambda|\sigma_{\Lambda^c}) := \sum_{x\in\mathbb{Z}^d:\tau_x\Delta\cap\Lambda\neq\emptyset} \tau_x v(\sigma_\Lambda, \sigma_{\Lambda^c})$. The new pressure is

$$P_{\beta,\Lambda,\sigma_{\Lambda^c}}(\lambda; v) := \frac{1}{\beta|\Lambda|}\log\left\{\sum_{\sigma_\Lambda} e^{-\beta[H_{\Lambda;\lambda,v}(\sigma_\Lambda|\sigma_{\Lambda^c})}\right\}. \tag{2.4.1.6}$$

Notice that the additional energy in the hamiltonian has a finite range, as this is equal to the diameter of the set Δ introduced above.

Theorem 2.4.1.1 *For any $\beta > 0$ and any cylindrical function v the following holds*:

- *For any van Hove sequence $\{\Lambda\}$, any sequence σ_{Λ^c} and any $\lambda \in \mathbb{R}$*

$$\lim_{\Lambda\to\mathbb{Z}^d} P_{\beta,\Lambda,\sigma_{\Lambda^c}}(\lambda; v) \text{ exists and we call it } P_\beta(\lambda; v). \tag{2.4.1.7}$$

Moreover, for any $w \in \mathcal{E}_v := \bigcup_{\lambda\in\mathbb{R}} [D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)]$,

$$\begin{aligned}&\lim_{\delta\to 0}\liminf_{\Lambda\to\mathbb{Z}^d} F_{\beta,\Lambda,\delta,\sigma_{\Lambda^c}}(w; v)\\ &\quad = \lim_{\delta\to 0}\limsup_{\Lambda\to\mathbb{Z}^d} F_{\beta,\Lambda,\delta,\sigma_{\Lambda^c}}(w; v) =: F_\beta(w; v).\end{aligned} \tag{2.4.1.8}$$

- *$P_\beta(\lambda; v)$ and $F_\beta(w; v)$ are Legendre transforms of each other,*

$$P_\beta(\lambda; v) = \sup_{w\in\mathcal{E}_v}\{\lambda w - F_\beta(w; v)\}, \quad F_\beta(w; v) = \sup_{\lambda\in\mathbb{R}}\{\lambda w - P_\beta(\lambda; v)\}, \tag{2.4.1.9}$$

so that $w \in [D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)]$ iff $\lambda \in [D^- F_\beta(w; v), D^+ F_\beta(w; v)]$ and for such conjugate pairs,

$$P_\beta(\lambda; v) = \lambda w - F_\beta(w; v). \tag{2.4.1.10}$$

- *(λ, w) is a conjugate pair if and only if there exists a translational invariant measure μ DLR at β with hamiltonian (2.4.1.4)–(2.4.1.5) such that*

$$\mu(v) = w. \tag{2.4.1.11}$$

Moreover, if $w = D^\pm P_\beta(\lambda; v)$, then μ can be chosen to be ergodic and if the derivative $DP_\beta(\lambda; v)|_{\lambda=0} =: w$ exists, then all $\mu \in \mathcal{G}^0$ are such that $\mu(v) = w$.

- *Finally,*

$$P_\beta(\lambda; v) = P_\beta(1; \lambda v) =: P_\beta(\lambda v), \tag{2.4.1.12}$$

and $P_\beta(v)$ is a convex function on the space whose elements are the cylindrical functions v.

Proof The proof is completely analogous to the one for the entropy–energy conjugation and it will only be sketched. Using that the additional term in the hamiltonian due to v has a finite range the proofs are essentially unchanged. We thus start by considering all the ergodic measures μ which are DLR at (β, λ) for any $\lambda \in \mathbb{R}$. Call $w = \mu(v)$, then the analogue of Proposition 2.3.4.2 holds and proves that the limit in (2.4.1.8) exists and is equal to

$$F_\beta(w; v) = -P_\beta(\lambda; v) + \lambda w.$$

Analogously to Proposition 2.3.7.1, $\mu(v) \in [D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)]$ for any translational invariant measure μ DLR at β with hamiltonian (2.4.1.4). The converse is also true; its proof is given by the analogues of Propositions 2.3.7.1 and 2.3.7.2: namely for any $w \in [D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)]$, there is a translational invariant measure μ DLR at β with hamiltonian (2.4.1.4), such that $\mu(v) = w$. Since any such measure μ verifies

$$\mu(v) \in [D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)]$$

there must be ergodic ones, $\mu_\pm$, such that $\mu_\pm(v) = D^\pm P_\beta(\lambda; v)$, which is proved as in (2.3.7.8). Then, calling $w_\pm := D^\pm P_\beta(\lambda; v)$, and repeating the proof of Proposition 2.3.7.6, we conclude that the limit (2.4.1.8) exists also when $w \in (D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v))$ and, moreover,

$$F_\beta(w; v) = aF_\beta(w_-; v) + (1-a)F_\beta(w_+; v), \quad w = aw_- + (1-a)w_+,\ a \in [0, 1].$$

Thus $F_\beta(w; v)$ is the Legendre transform of $P_\beta(\lambda; v)$ when

$$w \in \big(D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)\big)$$

and since for the others we have already proved that (2.4.1.10) holds (as for such w there are ergodic measures with $\mu(v) = w$), we then conclude the proof of (2.4.1.9). Equation (2.4.1.12) is obvious and the other statements in the theorem are either already proved or are a consequence of general convexity properties. □

We conclude this subsection with an interesting corollary of Theorem 2.4.1.1 which gives a characterization of the absence or presence of phase transitions, identifying their absence in the DLR context (i.e. uniqueness of translational invariant DLR measures) with the thermodynamic notion based on differentiability of the pressure.

Theorem 2.4.1.2 (Phase transitions) *There is only one invariant DLR measure at β with hamiltonian defined by* (2.3.2.13) *if and only if for any cylindrical function v*

$$D^+ P_\beta(\lambda; v)|_{\lambda=0} = D^- P_\beta(\lambda; v)|_{\lambda=0}. \tag{2.4.1.13}$$

Proof Suppose there is only one DLR measure μ. In the proof of Theorem 2.4.1.1 we have seen that there are ergodic measures $\mu_{\pm;\lambda,v}$ DLR at β with respect to the hamiltonian (2.4.1.4) and such that $\mu_{\pm;\lambda,v}(v) = D^{\pm}P_\beta(\lambda; v)$. Since $\mu_{\pm;0,v}$ are DLR at β with hamiltonian (2.3.2.13), $\mu_{+;0,v} = \mu_{-;0,v}$; hence (2.4.1.13). Thus $P_\beta(\lambda; v)$ is differentiable at $\lambda = 0$ for any v.

Suppose now that $P_\beta(\lambda; v)$ is differentiable at $\lambda = 0$ for any v and let μ' and μ'' be two translational invariant DLR measures with hamiltonian (2.3.2.13). We need to prove that $\mu' = \mu''$. By Theorem 2.4.1.1

$$\mu'(v) \in [D^- P_\beta(\lambda; v)|_{\lambda=0}, D^+ P_\beta(\lambda; v)|_{\lambda=0}] \Rightarrow \mu'(v) = DP_\beta(\lambda; v)|_{\lambda=0}$$

by the assumption of differentiability. The same argument applied to μ'' shows that $\mu'(v) = \mu''(v)$. Since this holds for all v, by density $\mu' = \mu''$. □

2.4.2 Tangent functionals to the pressure

There is a converse to Theorem 2.4.1.2, which shows that the DLR measures are characterized by the discontinuities in the intervals

$$[D^- P_\beta(\lambda; v)|_{\lambda=0}, D^+ P_\beta(\lambda; v)|_{\lambda=0}].$$

We need to introduce the notion of tangent functional to the graph of the pressure. We fix β and regard $P_\beta(v)$, defined in (2.4.1.12), as a function on the space $\mathcal{V}$ of all v.

Definition 2.4.2.1 A tangent functional to $P_\beta(\cdot)$ at v is a *linear functional* $\alpha(\cdot)$ on $\mathcal{V}$ such that

$$\alpha(v') \le P_\beta(v') - P_\beta(v), \quad \text{for any } v' \in \mathcal{V}. \tag{2.4.2.1}$$

Theorem 2.4.2.2 *Any translational invariant measure μ DLR at β with hamiltonian* (2.3.2.13) *is such that $\mu(v)$ is tangent to $P_\beta(\cdot)$ at $v = 0$. Vice versa, if α is tangent to $P_\beta(\cdot)$ at $v = 0$ then there is a unique translational invariant DLR measure μ such that $\mu(v) = \alpha(v)$ for all v.*

Proof Let μ be DLR at β with hamiltonian (2.3.2.13). Calling $P_\beta = P_\beta(0; v)$, we need to prove that $\mu(v) \le P_\beta(v) - P_\beta$ for any $v \in \mathcal{V}$. Call $\mathcal{E}'_v$ the set of all $\lambda \in \mathbb{R}$ where $DP_\beta(\lambda; v)$ exists, which is all $\mathbb{R}$ except for countably many points. Then

$$P_\beta(v) - P_\beta = \int_{\mathcal{E}'_v \cap [0,1]} DP_\beta(\lambda; v)\, d\lambda. \tag{2.4.2.2}$$

By (2.3.6.1) for $\lambda > 0$, $DP^-_\beta(\lambda; v) \ge D^+ P_\beta(\lambda; v)|_{\lambda=0}$, so that (2.4.2.2) yields

$$P_\beta(v) - P_\beta \ge DP^+_\beta(\lambda; v)|_{\lambda=0}, \tag{2.4.2.3}$$

and $DP^+_\beta(\lambda; v)|_{\lambda=0} \ge \mu(v)$: in fact, by (2.4.1.11), $(0, \mu(v))$ is a conjugate pair and $\mu(v) \in [DP^-_\beta(0; v), D^+ P_\beta(0; v)]$. For the converse statement we refer to the literature; see for instance [200]. □

Remarks

We can therefore recover the set of equilibrium states from the pressure, as the functionals tangent to the pressure are the translational invariant DLR measures. Thus ultimately the only axiom needed for the whole theory is the Boltzmann identification of the entropy in terms of the number of states.

2.4.3 Large deviations

Suppose that our Ising system is in equilibrium at an inverse temperature β for which there is a unique ergodic DLR measure μ. We then know that the ergodic average $\mathcal{A}_{\Lambda,\sigma}(u)$ defined in (2.3.2.12) is with large probability close to $\mu(u)$ so that we can read out of σ what is the temperature of the system. In principle, at least the quantity $\mathcal{A}_{\Lambda,\sigma}(u)$ can be observed experimentally, thus providing an alternative way to measure temperatures. However, $\mathcal{A}_{\Lambda,\sigma}(u)$ is close to $\mu(u)$ "only with large probability" and the question then arises of what the probability is to mistake the correct values $\mu(u)$ and β. As we will see such a probability is exponentially small with the volume and for such a reason it is called "a large deviation." As already discussed $\mathcal{A}_{\Lambda,\sigma}(u)$ is for large $|\Lambda|$ close to $H_\Lambda(\sigma_\Lambda)$. Let E be an energy density corresponding to an ergodic DLR measure with inverse temperature β'. Then

$$\begin{aligned}
&G_{\Lambda,\beta}\big(|H_\Lambda(\sigma_\Lambda) - E|\Lambda|| \le \delta\big)\\
&\quad= \sum_{|H_\Lambda(\sigma_\Lambda)-E|\Lambda||\le\delta} \frac{e^{-\beta H_\Lambda(\sigma_\Lambda)}}{Z_{\beta,\Lambda}} \frac{e^{-(\beta'-\beta)H_\Lambda(\sigma_\Lambda)}}{Z_{\beta',\Lambda}} \{e^{(\beta'-\beta)H_\Lambda(\sigma_\Lambda)} Z_{\beta',\Lambda}\}\\
&\quad\approx \frac{Z_{\beta',\Lambda}}{Z_{\beta,\Lambda}} e^{(\beta'-\beta)E|\Lambda|} G_{\beta',\Lambda}\big(|H_\Lambda(\sigma_\Lambda) - E|\Lambda|| \le \delta\big).
\end{aligned}$$

We thus expect that

$$\frac{\log G_{\Lambda,\beta}\big(|H_\Lambda(\sigma_\Lambda) - E|\Lambda|| \le \delta\big)}{|\Lambda|} \approx (\beta'-\beta)E + \beta' P_{\beta'} - \beta P_\beta,$$

in which case the rate of large deviations, namely the l.h.s., is related to thermodynamics and given by $(\beta'-\beta)E + \beta' P_{\beta'} - \beta P_\beta$. The above argument can be made rigorous and generalized as follows (proofs will be omitted).

Let μ be a DLR measure at inverse temperature β with hamiltonian as in (2.3.2.13). Let v be a cylindrical function, $\Delta \to \mathbb{Z}^d$ an increasing sequence of cubes, $w \in \mathbb{R}$ and

$$A_{w,\delta;\Delta} := \mu\left(\frac{1}{|\Delta|}\left|\sum_{x\in\Delta}(\tau_x v - w)\right| \le \delta\right). \tag{2.4.3.1}$$

We have

Theorem 2.4.3.1 *Let Δ be an increasing sequence of cubes*; *then*

$$\lim_{\delta\to 0}\lim_{\Delta\to\mathbb{Z}^d}\frac{\log A_{w,\delta;\Delta}}{\beta|\Delta|} = -[\lambda w - P_\beta(\lambda; v) + P_\beta], \tag{2.4.3.2}$$

where λ is such that $w \in [D^- P_\beta(\lambda; v), D^+ P_\beta(\lambda; v)]$. If the Hamiltonian is instead defined by $u - \lambda_0 v$ ($\lambda_0 = 0$ previously), then (2.4.3.2) *becomes*

$$\lim_{\delta\to 0}\lim_{\Delta\to\mathbb{Z}^d}\frac{\log A_{w,\delta;\Delta}}{\beta|\Delta|} = -[(\lambda - \lambda_0)w - P_\beta(\lambda; v) + P_\beta(\lambda_0; v)].$$