# Introduction

## 0.1 Generalities

The formalism of equilibrium statistical mechanics – which we shall call *thermodynamic formalism* – has been developed since G. W. Gibbs to describe the properties of certain physical systems. These are systems consisting of a large number of subunits (typically $10^{27}$) like the molecules of one liter of air or water. While the physical justification of the thermodynamic formalism remains quite insufficient, this formalism has proved remarkably successful at explaining facts.

In recent years it has become clear that, underlying the thermodynamic formalism, there are mathematical structures of great interest: the formalism hints at the good theorems, and to some extent at their proofs. Outside of statistical mechanics proper, the thermodynamic formalism and its mathematical methods have now been used extensively in *constructive quantum field theory** and in the study of certain *differentiable dynamical systems* (notably Anosov diffeomorphisms and flows). In both cases the relation is at an abstract mathematical level, and fairly inobvious at first sight. It is evident that the study of the physical world is a powerful source of inspiration for mathematics. That this inspiration can act in such a detailed manner is a more remarkable fact, which the reader will interpret according to his own philosophy.

The main physical problem which equilibrium statistical mechanics tries to clarify is that of phase transitions. When the temperature of water is lowered, why do its properties change first smoothly, then suddenly as the freezing point is reached? While we have some general ideas about this, and many special results,

* See for instance Velo and Wightman [1].

a conceptual understanding is still missing.† The mathematical investigation of the thermodynamic formalism is in fact not completed; the theory is a young one, with emphasis still more on imagination than on technical difficulties. This situation is reminiscent of preclassic art forms, where inspiration has not been castrated by the necessity to conform to standard technical patterns. We hope that some of this juvenile freshness of the subject will remain in the present monograph!

The physical systems to which the thermodynamic formalism applies are idealized to be actually infinite, i.e. to fill $\mathbf{R}^\nu$ (where $\nu = 3$ in the usual world). This idealization is necessary because only infinite systems exhibit sharp phase transitions. Much of the thermodynamic formalism is concerned with the study of *states* of infinite systems.

For *classical systems* the states are probability measures on an appropriate space of infinite configurations; such states can also be viewed as linear functionals on an abelian algebra (an algebra of continuous functions in the case of Radon measures). For *quantum systems* the states are "expectation value" linear functionals on non-abelian algebras. Due to their greater simplicity, classical systems have been studied more than quantum systems. In fact attention has concentrated on the simplest systems, the *classical lattice systems* where $\mathbf{R}^\nu$ is replaced by $\mathbf{Z}^\nu$ (a $\nu$-dimensional crystal lattice). For such systems the configuration space is a subset $\Omega$ of $\prod_{x\in\mathbf{Z}^\nu}\Omega_x$ (where $\Omega_x$ is for instance the set of possible "spin values" or "occupation numbers" at the lattice site $x$). We shall assume that $\Omega_x$ is finite. Due to the group invariance (under $\mathbf{Z}^\nu$ or $\mathbf{R}^\nu$) the study of states of infinite systems is closely related to ergodic theory. There are however other parts of the thermodynamic formalism concerned with quite different questions (like analyticity problems).

The present monograph addresses itself to mathematicians. Its aim is to give an account of part of the thermodynamic formalism, and of the corresponding structures and methods. We have restricted ourselves to classical lattice systems. The thermodynamic formalism extends to many other classes of systems, but the theory as it exists now for those systems is less complete, more singular, and filled with technical difficulties. The formalism which we shall describe would not apply directly to the problems of constructive quantum field theory, but it is appropriate to the discussion of Anosov diffeomorphisms and related dynamical systems.

† At a more phenomenological level, a good deal is known about phase transitions and much attention has been devoted to critical points and "critical phenomena"; the latter remain however for the moment inaccessible to rigorous investigations.

The mathematics underlying the thermodynamic formalism consists of general methods and special techniques. We have restricted ourselves in this monograph to the general methods; we hope that a complement on special techniques will be published later. As a rough rule, we have decided that a result was not "general" if it required that the configuration space of the system factorize completely in the form $\Omega = \prod \Omega_x$, where $\Omega_x$ is the finite set of "spin values" at the lattice site $x$. The body of general methods thus defined has considerable unity. As for the special techniques, let us mention the correlation inequalities, the method of integral equations, the Lee-Yang circle theorem, and the Peierls argument. These techniques look somewhat specialized from the general point of view taken in this monograph, but are often extremely elegant. They provide, in special situations, a variety of detailed results of great interest for physics.

## 0.2 Description of the thermodynamic formalism

The contents of this section are not logically required for later chapters. We describe here, for purposes of motivation and orientation, some of the ideas and results of the thermodynamic formalism.* The reader may go over this material rapidly, or skip it entirely.

### I. Finite systems

Let $\Omega$ be a non-empty finite set. Given a probability measure $\sigma$ on $\Omega$ we define its *entropy*

$$S(\sigma) = -\sum_{\xi\in\Omega} \sigma\{\xi\} \log \sigma\{\xi\},$$

where it is understood that $t \log t = 0$ if $t = 0$. Given a function $U : \Omega \to \mathbf{R}$, we define a real number $Z$ called the *partition function* and a probability measure $\rho$ on $\Omega$ called the *Gibbs ensemble* by

$$Z = \sum_{\xi\in\Omega} \exp[-U(\xi)],$$

$$\rho\{\xi\} = Z^{-1} \exp[-U(\xi)]. \tag{0.1}$$

**Proposition** (Variational principle). *The maximum of the expression*†

$$S(\sigma) - \sigma(U)$$

* We follow in part the Séminaire Bourbaki, exposé 480.
† We write $\sigma(U) = \sum_\xi \sigma\{\xi\}U(\xi)$ or more generally $\sigma(U) = \int U(\xi)\sigma(d\xi)$.

*over all probability measures* $\sigma$ *on* $\Omega$ *is* $\log Z$, *and is reached precisely for* $\sigma = \rho$.

For physical applications, $\Omega$ is interpreted as the space of configurations of a finite system. One writes $U = \beta E$, where $E(\xi)$ is the energy of the configuration $\xi$, and $\beta = 1/kT$, where $T$ is the absolute temperature and $k$ is a factor known as Boltzmann's constant. The problem of why the Gibbs ensemble describes thermal equilibrium (at least for "large systems") when the above physical identifications have been made is deep and incompletely clarified. Note that the energy $E$ may depend on physical parameters called "magnetic field," "chemical potential," etc. Note also that the traditional definition of the energy produces a minus sign in $\exp[-\beta E]$, which is in practice a nuisance. From now on we absorb $\beta$ in the definition of $U$, and call $U$ the *energy*. We shall retain from the above discussion only the hint that the Gibbs ensemble is an interesting object to consider in the limit of a "large system."

The thermodynamic formalism studies measures analogous to the Gibbs ensemble $\rho$ in a certain limit where $\Omega$ becomes infinite, but some extra structure is present. Imitating the variational principle of the above Proposition, one defines *equilibrium states* (see II below). Imitating the definition (0.1), one defines *Gibbs states* (see III below).

## II. Thermodynamic formalism on a metrizable compact set

Let $\Omega$ be a non-empty metrizable compact set, and $x \to \tau^x$ a homomorphism of the additive group $\mathbf{Z}^\nu (\nu \geqslant 1)$ into the group of homeomorphisms of $\Omega$. We say that $\tau$ is *expansive* if, for some allowed metric $d$, there exists $\delta > 0$ such that

$$(d(\tau^x \xi, \tau^x \eta) \leqslant \delta \quad \text{for all } x) \Rightarrow (\xi = \eta).$$

**Definition of the pressure.** If $\mathfrak{A} = (\mathfrak{A}_i)$, $\mathfrak{B} = (\mathfrak{B}_j)$ are covers of $\Omega$, the cover $\mathfrak{A} \vee \mathfrak{B}$ consists of the sets $\mathfrak{A}_i \cap \mathfrak{B}_j$. This notation extends to any finite family of covers. We write

$$\begin{aligned} \tau^{-x}\mathfrak{A} &= (\tau^{-x}\mathfrak{A}_i), \\ \mathfrak{A}^\Lambda &= \bigvee_{x \in \Lambda} \tau^{-x}\mathfrak{A} \qquad \text{if} \quad \Lambda \subset \mathbf{Z}^\nu, \\ \operatorname{diam} \mathfrak{A} &= \sup_i \operatorname{diam} \mathfrak{A}_i, \end{aligned}$$

where $\operatorname{diam} \mathfrak{A}_i$ is the diameter of $\mathfrak{A}_i$ for an allowed metric $d$ on $\Omega$.

The definition of the pressure which we shall now give will not look simple and natural to someone unfamiliar with the subject. This should not alarm the reader: the definition will give us quick access to a general statement of theorems of statistical mechanics. It will otherwise recur only in Chapter 6, with more preparation.

We denote by $\mathcal{C} = \mathcal{C}(\Omega)$ the space of continuous real functions on $\Omega$. Let $A \in \mathcal{C}$, $\mathfrak{A}$ be a finite open cover of $\Omega$, and $\Lambda$ be a finite subset of $\mathbf{Z}^\nu$; define

$$Z_\Lambda(A, \mathfrak{A}) = \min \left\{ \sum_j \exp \left[ \sup_{\xi \in \mathfrak{B}_j} \sum_{x \in \Lambda} A(\tau^x \xi) \right] \right. \\ \left. : (\mathfrak{B}_j) \text{ is a subcover of } \mathfrak{A}^\Lambda \right\}.$$

If $a^1, \ldots, a^\nu$ are integers $> 0$, let $a = (a^1, \ldots, a^\nu)$ and

$$\Lambda(a) = \{(x^1, \ldots, x^\nu) \in \mathbf{Z}^\nu : 0 \leqslant x^i < a^i \text{ for } i = 1, \ldots, \nu\}.$$

The function $a \to \log Z_{\Lambda(a)}(A, \mathfrak{A})$ is subadditive, and one can write (with $|\Lambda(a)| = \text{card } \Lambda(a) = \prod_i a^i$)

$$P(A, \mathfrak{A}) = \lim_{a^1, \ldots, a^\nu \to \infty} \frac{1}{|\Lambda(a)|} \log Z_{\Lambda(a)}(A, \mathfrak{A}) \\ = \inf_a \frac{1}{|\Lambda(a)|} \log Z_{\Lambda(a)}(A, \mathfrak{A}),$$

and

$$P(A) = \lim_{\text{diam } \mathfrak{A} \to 0} P(A, \mathfrak{A}).$$

The function $P : \mathcal{C} \to \mathbf{R} \cup \{+\infty\}$ is the (topological) *pressure*. $P(A)$ is finite for all $A$ if and only if $P(0)$ is finite; in that case $P$ is convex and continuous (for the topology of uniform convergence in $\mathcal{C}$). $P(0)$ is the *topological entropy*; it gives a measure of the rate of mixing of the action $\tau$.

**Entropy of an invariant measure.** If $\sigma$ is a probability measure on $\Omega$, and $\mathfrak{A} = (\mathfrak{A}_i)$ a finite Borel partition of $\Omega$, we write

$$H(\sigma, \mathfrak{A}) = -\sum_i \sigma(\mathfrak{A}_i) \log \sigma(\mathfrak{A}_i).$$

The real measures on $\Omega$ constitute the dual $\mathcal{C}^*$ of $\mathcal{C}$. The topology of weak dual of $\mathcal{C}$ on $\mathcal{C}^*$ is called the *vague* topology. Let $I \subset \mathcal{C}^*$ be the set of probability measures $\sigma$ invariant under $\tau$, i.e. such that $\sigma(A) = \sigma(A \circ \tau^x)$; $I$ is convex and compact for the vague topology. If $\mathfrak{A}$ is a finite Borel partition and $\sigma \in I$, we write

$$h(\sigma, \mathfrak{A}) = \lim_{a^1,\ldots,a^\nu\to\infty} \frac{1}{|\Lambda(a)|} H(\sigma, \mathfrak{A}^{\Lambda(a)})$$
$$= \inf_a \frac{1}{|\Lambda(a)|} H(\sigma, \mathfrak{A}^{\Lambda(a)});$$
$$h(\sigma) = \lim_{\operatorname{diam}\mathfrak{A}\to 0} h(\sigma, \mathfrak{A}).$$

The function $h : I \to \mathbf{R} \cup \{+\infty\}$ is affine $\geqslant 0$; it is called the (mean) *entropy*. If $\tau$ is expansive, $h$ is finite and upper semi-continuous on $I$ (with the vague topology).

**Theorem 1** (Variational principle).

$$P(A) = \sup_{\sigma\in I}[h(\sigma) + \sigma(A)]$$

*for all* $A \in C$.

This corresponds to the variational principle for finite systems if $-A$ is interpreted as the contribution to the energy of one lattice site.

Let us assume that $P$ is finite. The set $I_A$ of *equilibrium states* for $A \in \mathcal{C}$ is defined by

$$I_A = \{\sigma \in I : h(\sigma) + \sigma(A) = P(A)\}.$$

$I_A$ may be empty.

**Theorem 2** *Assume that h is finite and upper semi-continuous on I (with the vague topology).*

(a) $I_A = \{\sigma \in \mathcal{C}^* : P(A + B) \geqslant P(A) + \sigma(B)$ for all $B \in \mathcal{C}\}$. *This set is not empty; it is convex, compact; it is a Choquet simplex and a face of I.*
(b) *The set* $D = \{A \in \mathcal{C} : \text{card } I_A = 1\}$ *is residual in* $\mathcal{C}$.
(c) *For every* $\sigma \in I$,

$$h(\sigma) = \inf_{A\in\mathcal{C}}[P(A) - \sigma(A)].$$

The fact that $I_A$ is a metrizable simplex implies that each $\sigma \in I_A$ has a unique integral representation as the barycenter of a measure carried by the extremal points of $I_A$. It is known that $I$ is also a simplex. The fact that $I_A$ is a face of $I$ implies that the extremal points of $I_A$ are also extremal points of $I$ (i.e. $\tau$-ergodic measures on $\Omega$).

Cambridge University Press
0521546494 - Thermodynamic Formalism: The Mathematical Structures of Equilibrium Statistical Mechanics, Second Edition
David Ruelle
Excerpt
More information



## III. Statistical mechanics on a lattice

The above theorems extend results known for certain systems of statistical mechanics (classical lattice systems). For instance, if $F$ is a non-empty finite set (with the discrete topology), we can take $\Omega = F^{\mathbf{Z}^\nu}$ with the product topology, and $\tau^x$ defined in the obvious manner. More generally we shall take for $\Omega$ a closed $\tau$-invariant non-empty subset of $F^{\mathbf{Z}^\nu}$. For the physical interpretation, note that $\Omega$ is the space of infinite configurations of a system of spins on a crystal lattice $\mathbf{Z}^\nu$. Up to sign and factors of $\beta$, $P$ can be interpreted as the "free energy" or the "pressure," depending on the physical interpretation of $F$ as the set of "spin values" or of "occupation numbers" at a lattice site. For simplicity we have retained the word "pressure."

If $x = (x^i) \in \mathbf{Z}^\nu$, we write $|x| = \max |x^i|$. Let $0 < \lambda < 1$; if $\xi, \eta \in \Omega$, with $\xi = (\xi_x)_{x\in\mathbf{Z}^\nu}$, $\eta = (\eta_x)_{x\in\mathbf{Z}^\nu}$, we define

$$d(\xi, \eta) = \lambda^k \qquad \text{with} \quad k = \inf\{|x| : \xi_x \neq \eta_x\}.$$

$d$ is a distance compatible with the topology of $\Omega$. One checks with this definition that $\tau$ is expansive; hence Theorem 2 applies.

We shall henceforth assume that there exists a finite set $\Delta \subset \mathbf{Z}^\nu$ and $G \subset F^\Delta$ such that

$$\Omega = \{\xi \in F^{\mathbf{Z}^\nu} : \tau^x\xi|\Delta \in G \qquad \text{for all} \quad x\}.$$

If $\Lambda \subset \mathbf{Z}^\nu$ we denote by $\mathrm{pr}_\Lambda$, $\mathrm{pr}'_\Lambda$ the projections of $F^{\mathbf{Z}^\nu}$ on $F^\Lambda$ and $F^{\mathbf{Z}^\nu\setminus\Lambda}$ respectively.

Let $0 < \alpha \leqslant 1$, and denote by $\mathcal{C}^\alpha$ the Banach space of real Hölder continuous functions of exponent $\alpha$ on $\Omega$ (with respect to the metric d). Let $\xi = (\xi_x) \in \Omega$, $\eta = (\eta_x) \in \Omega$. If $\xi_x = \eta_x$ except for finitely many $x$, and $A \in \mathcal{C}^\alpha$, we can write

$$g_A(\xi, \eta) = \prod_{x\in\mathbf{Z}^\nu} \exp[A(\tau^x\xi) - A(\tau^x\eta)]$$

because $|A(\tau^x\xi) - A(\tau^x\eta)| \to 0$ exponentially fast when $|x| \to \infty$. For finite $\Lambda$, a continuous function $f_\Lambda : \mathrm{pr}_\Lambda\Omega \times \mathrm{pr}'_\Lambda\Omega \to \mathbf{R}$ is then defined by

$$f_\Lambda(\xi) = \begin{cases} \left[\displaystyle\sum_{\eta\in\Omega:\mathrm{pr}'_\Lambda\eta=\mathrm{pr}'_\Lambda\xi} g_A(\eta, \xi)\right]^{-1} & \text{if} \quad \xi \in \Omega, \\ 0 & \text{if} \quad \xi \notin \Omega. \end{cases}$$

**Definition**. Let $A \in \mathcal{C}^\alpha$; we say that a probability measure $\sigma$ on $\Omega$ is a *Gibbs state* if the following holds.

For every finite $\Lambda \subset \mathbf{Z}^\nu$, let $\epsilon_\Lambda$ be the measure on $\mathrm{pr}_\Lambda \Omega$ which gives to each point of this set the mass 1. Then

$$\sigma = f_\Lambda \cdot (\epsilon_\Lambda \otimes \mathrm{pr}'_\Lambda \sigma);$$

(we have again denoted by $\sigma$ the image of this measure by the canonical map $\Omega \mapsto \mathrm{pr}_\Lambda \Omega \times \mathrm{pr}'_\Lambda \Omega$).

Another formulation of the definition is this: $\sigma$ is a Gibbs state if, for every finite $\Lambda$, the conditional probability that $\xi|\Lambda$ is realized in $\Lambda$, knowing that $\xi|(\mathbf{Z}^\nu \backslash \Lambda)$ realized in $\mathbf{Z}^\nu \backslash \Lambda$, is $f_\Lambda(\xi)$.

**Theorem 3** *Let $A \in \mathcal{C}^\alpha$.*

(a) *Every equilibrium state is a $\tau$-invariant Gibbs state.*
(b) *If $\Omega = F^{\mathbf{Z}^\nu}$, every $\tau$-invariant Gibbs state is an equilibrium state.*

In view of (a), the Gibbs states are the probability measures which have the same conditional probabilities $f_\Lambda$ as the equilibrium states. Part (b) of the theorem holds under conditions much more general than $\Omega = F^{\mathbf{Z}^\nu}$. The assumption $A \in \mathcal{C}^\alpha$ can also be considerably weakened. For simplicity we have in this section made an unusual description of statistical mechanics, using (Hölder continuous) functions on $\Omega$, rather than the "interactions" which are much more appropriate to a detailed study.

**Theorem 4** *The set of Gibbs states for $A \in \mathcal{C}^\alpha$ is a Choquet simplex.*

Every Gibbs state has thus a unique integral decomposition in terms of extremal (or "pure") Gibbs states.

**Physical interpretation.** The extremal equilibrium states are $\tau$-ergodic measures. They are interpreted as *pure thermodynamic phases*. Since the equilibrium states correspond to tangents to the graph of $P$ (Theorem 2(a)), the discontinuities of the derivative of $P$ correspond to *phase transitions*. One would thus like to know if $P$ is piecewise analytic (in a suitable sense) on $\mathcal{C}^\alpha$. An extremal equilibrium state $\sigma$ may have a non-trivial decomposition into extremal Gibbs states (those will not be $\tau$-invariant, because of Theorem 3(b)). This is an example of *symmetry breaking* (the broken symmetry is the invariance under $\tau$).

The main problem of equilibrium statistical mechanics is to understand the nature of phases and phase transitions. Because of this, the main object of the thermodynamic formalism is to study the differentiability and analyticity properties of the function $P$, and the structure of the equilibrium states and Gibbs states. As already mentioned, detailed results are known only in special cases, but we shall restrict ourselves in the present monograph to the general theory, as it is known at the time of writing.

For "one-dimensional systems," i.e. for $\nu = 1$, there are fairly complete results, which can be summarized by saying that there are no phase transitions. Let us assume that

$$\Omega = \{\xi = (\xi_x)_{x \in \mathbf{Z}} \in F^{\mathbf{Z}} : t_{\xi_x \xi_{x+1}} = 1 \qquad \text{for all} \quad x\},$$

where $t = (t_{uv})$ is a matrix with elements 0 or 1. We assume also that there exists an integer $N > 0$ such that all the matrix elements of $t^N$ are $> 0$.

**Theorem 5** *If the above conditions are satisfied,* $P : \mathcal{C}^\alpha \to \mathbf{R}$ *is real analytic. Furthermore for every* $A \in \mathcal{C}^\alpha$ *there is only one Gibbs state (which is also the only equilibrium state).*

All these properties are false for $\nu > 1$.

## 0.3 Summary of contents

Chapters 1 to 5 of this monograph are devoted to the general theory of equilibrium statistical mechanics of classical lattice systems; complete proofs are generally given. Chapters 6 and 7 extend the thermodynamic formalism outside of the traditional domain of statistical mechanics: here the proofs are largely omitted or only sketched.* We give now some more details.

Chapters 1 and 2 give the theory of Gibbs states, without assuming invariance under lattice translations (the lattice $\mathbf{Z}^\nu$ is thus replaced by a general infinite countable set $L$). Chapter 3 assumes translation invariance and develops the theory of equilibrium states and of the pressure for classical lattice systems; general results on phase transitions are also obtained. Chapter 4 is central, and establishes the connexion between Gibbs states and equilibrium states. Chapter 5 deals with one-dimensional systems and prepares Chapter 7. Chapter 6 extends the theory of equilibrium states to the situation where the configuration space $\Omega$ is replaced by a general compact metrizable space on which $\mathbf{Z}^\nu$ acts by homeomorphisms. Chapter 7 extends the theory of Gibbs states (and related topics) to a certain class of compact metric spaces, which we call *Smale spaces*, on which $\mathbf{Z}$ acts by homeomorphisms. Smale spaces include Axiom A basic sets and in particular manifolds with an Anosov diffeomorphism.

Some extra material is given in the form of exercises at the end of the chapters.

Bibliographic references are given either in the text or in notes at the end of the chapters. For purposes of orientation, it may be good to read these notes

* Of course references to the literature are indicated as needed.

before the corresponding chapter. The reader is particularly advised to consult the following original papers: Ruelle [1], Dobrushin [2], [3], Lanford and Ruelle [1], Israel [1], and Sinai [4].

Some background material has been collected in Appendices A.1 to A.5. These appendices recall some well-known facts to establish terminology, and also provide access to less standard results. In general the reader is assumed to be familiar with basic facts of functional analysis, but no knowledge of physics is presupposed.

A few open problems are collected in Appendix B. Appendix C contains a brief introduction to flows.

Concerning notation and terminology we note the following points. We shall often write $|X|$ for the cardinality of a finite set $X$. We shall use in Chapters 5–7 the notation $\mathbf{Z}_{>}, \mathbf{Z}_{\geqslant}, \mathbf{Z}_{<}, \mathbf{Z}_{\leqslant}$ for the sets of integers which are respectively $>0$, $\geqslant 0$, $<0$, $\leqslant 0$. A measure $\rho$ will (unless otherwise indicated) be a *Radon measure* on a compact set $\Omega$. If $f : \Omega \to \Omega'$ is a continuous map, the image of $\rho$ by $f$ (see Appendix A.4) is denoted by $f\rho$ (*not* $f^*\rho$).

We refer the reader to Ruelle [3] for a wider study of equilibrium statistical mechanics, and to the excellent monograph by Bowen [6] for applications to differentiable dynamical systems.* Let us also mention the monograph by Israel [2] and the notes by Lanford [2], Georgii [1], and Preston [1], [2]. Monographs are planned by various authors on aspects of statistical mechanics not covered here, but at this time, much interesting material is not available in book form.

Before proceeding with Chapter 1, the reader is invited to go rapidly through the Appendices A.1–A.5.

* For modern introductions to ergodic theory and topological dynamics, see Walters [2]: Denker, Grillenberger, and Sigmund [1].