Gibbs Measures and General Thermodynamic Formalism

John Ludovic Pirl

August 10, 2016

Abstract

The present document has been written as part of the Diploma Work of the Master Degree "Probabilités et Modèles Aléatoires" at LPMA[1], Université Pierre et Marie Curie, Paris. It is based on the first two chapters of Rufus Bowen's lecture notes "Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms" [3].

I hope to give the reader a few complements and perspectives to better understand the article as well as explain some fundamental concepts in more detail. A summary is also given.

The goal of the Diploma Work at LPMA is to read and understand an article in full detail, hand in a written report and give a presentation exposing the topic.

My warm thanks go to Yvan Velenik and Sacha Friedli at the University of Geneva, who suggested to me this particularly beautiful topic, as well as for being my formal advisers for this work.

[1] Laboratoire de Probabilités et Modèles Aléatoires

Contents

1 Measure-theoretical preliminaries and complements

1.1 Structural comments on Σ_n and Σ_A

We work on the configuration space $\Sigma_n = \{1, \ldots, n\}^{\mathbb{Z}}$. $\{1, \ldots, n\}$ is given the (metrizable) discrete topology, and the configuration space, endowed with the product topology, is thus compact by Tychonoff's theorem. A configuration is a point $\underline{x} = \{x_i\}_{i=-\infty}^{\infty} \in \Sigma_n$. A basis of the product topology is (for example) given by elements $U_m(\underline{x}) = \{\underline{y} : y_i = x_i \; \forall \; |i| \leq m\}$.

Let A be an $n \times n$ matrix of 0's and 1's such that each row and each column admits at least a 1 (so that one can leave any coordinate and arrive at any coordinate from somewhere), define $\Sigma_A = \{\underline{x} \in \Sigma_n : A_{x_i, x_{i+1}} = 1 \; \forall i \in \mathbb{Z}\}$. Notice that Σ_A is closed. Indeed, $\underline{x} \notin \Sigma_A$ is equivalent to there being an $i \in \mathbb{Z}$ such that $A_{x_i, x_{i+1}} = 0$. But then the cylindrical (open) set $\{\underline{y} : y_i = x_i, y_{i+1} = x_{i+1}\}$ is in $(\Sigma_A)^{\mathsf{c}}$. Thus Σ_A is closed and compact too. Note that $\Sigma_n = \Sigma_A$ where $A_{i,j} = 1$ for all i and j.

There is an important interpretation of the (natural) entries of powers of the matrix A. $A^m_{i,j}$ for $m \geq 1$ will give the number of admissible paths $x_0 x_1 \ldots x_m$ with $x_0 = i, x_m = j$ and $A_{x_k, x_{k+1}} = 1 \; \forall k = 0, \ldots, m-1$. This is immediate by induction on m.

Given a countable collection of metric spaces $\{X_n, \rho_n\}_{n \in \mathbb{N}}$ with $\rho_n \leq 1$ for all n, the function $\rho : X \times X \to \mathbb{R}$ on the Cartesian product $\prod_{n \in \mathbb{N}} X_n$ given by

$$\rho(x, y) = \sum_{n \in \mathbb{N}} \frac{\rho_n(x_n, y_n)}{2^n} \tag{1}$$

defines a metric that induces the product topology. Note that more generally, for an arbitrary metric space (Y, d), the function $d' = \frac{d}{1+d}$ induces the same topology on Y. So for arbitrary (X_n, ρ_n), we can generate the product topology with the distance function

$$\rho(x, y) = \sum_{n \in \mathbb{N}} \frac{\rho_n(x_n, y_n)}{2^n(1 + \rho_n(x_n, y_n))}. \tag{2}$$

In our case, the discrete topology on $\{1, \ldots, n\}$ can be generated by $d(x, y) = 1 - \delta_{x,y}$. It will be useful at times to keep our distance function in mind for the product topology in order to isolate sets of the type $U_m(\underline{x})$ given above for m large enough. Note that the order given to the numbering of spaces X_n is irrelevant.

For a function $\phi : \Sigma_A \to \mathbb{R}$ we define for $k \geq 0$ the variation

$$\text{var}_k \phi = \{|\phi(\underline{x}) - \phi(\underline{y})| : x_i = y_i \; \forall |i| \leq k\} \in [0, \infty]. \tag{3}$$

If $\phi \in \mathscr{C}(\Sigma_A)$, $\text{var}_k \phi \in [0, \infty)$. Considering the metric given above and the fact that Σ_A is compact, we immediately deduce that

$$\phi \text{ continuous} \Leftrightarrow \phi \text{ uniformly continuous} \Leftrightarrow lim_{k \to \infty} \text{var}_k \phi = 0.$$

1.2 Some Ergodic Theory

We discuss the few ergodic notions that are necessary for the article, as well as a couple of fundamental results. We refer the reader to [10] for a very precise and comprehensive treatment of the subject.

Consider $(M, \mathcal{B}, f)$ a measurable space M endowed with a sigma-algebra $\mathcal{B}$ and a measurable application $f : M \to M$. Such a system $(M, \mathcal{B}, f)$ is called a *dynamical system*. A probability measure μ for such a dynamical system is called *f-invariant* if $f_*\mu = \mu$. In other words $\mu(f^{-1}(B)) = \mu(B)$ for all $B \in \mathcal{B}$. A set $B \in \mathcal{B}$ is said to be *invariant* if $f^{-1}(B) = B$.

Definition (Ergodic measure)**.** *An f-invariant probability measure μ on a dynamical system $(M, \mathcal{B}, f)$ is called ergodic if every invariant set has measure 0 or 1.*

The set of all invariant probability measures on a dynamical system forms a convex polytope. It can be shown (quite easily) that the set of ergodic measures are exactly the extremal points of this convex polytope. In other words, an invariant probability measure μ is ergodic if and only if it cannot be written as convex combination $\mu = t\mu_1 + (1-t)\mu_2$ for μ_1, μ_2 invariant probabilities and $t \in (0,1)$. For a proof, see 4.3.2 of [10].

Definition (Mixing)**.** *A dynamical system (M, f) with an invariant measure μ is called mixing if*

$$\lim_{n\to\infty} \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B) \quad \forall A, B \in \mathcal{B}. \tag{4}$$

We immediately notice that mixing implies ergodicity since for an invariant set A, one gets $\mu(A) = \mu(A)^2$, whence $\mu(A) \in \{0, 1\}$.

Let's move on to a central result in Ergodic theory, the proof of which can be found in chapter 3 of [10].

Theorem (Birkhoff)**.** *Let $f : M \to M$ be a measurable transformation and μ an f-invariant measure. Given any integrable function $\phi : M \to \mathbb{R}$, the limit*

$$\tilde{\phi}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \phi(f^j(x)) \tag{5}$$

exists μ-a.s. Furthermore, the function $\tilde{\phi}$ thus defined is integrable and satisfies

$$\int \tilde{\phi}(x) d\mu(x) = \int \phi(x) d\mu(x). \tag{6}$$

If we plug $\phi = \chi_E$ (for E measurable) in equality (5), we get almost surely a proportion of time spend by the dynamical system (the orbit of the point x) in the set E and denote it by $\tau(E, x)$. Equality (6) then says that the mean time-average of the dynamical system is equal to its space average.

We're now in a position to give equivalent formulations of ergodicity. The following theorem (and its proof) can be found in [10], proposition 4.1.3.

Theorem (Ergodicity conditions)**.** *Let μ be an invariant probability measure for a dynamical system $(M, \mathcal{B}, f)$. Equivalent are:*

(a) *μ is ergodic.*

(b) *For each $B \in \mathcal{B}$ one has $\tau(B, x) = \mu(B)$, μ-a.s.*

(c) *For each $B \in \mathcal{B}$, the function $\tau(B, .)$ is constant μ-a.s.*

(d) *For each integrable function $\phi : M \to \mathbb{R}$ one has $\tilde{\phi}(x) = \int \phi d\mu$ μ-a.s.*

(e) *For each integrable function $\phi : M \to \mathbb{R}$ the temporal average $\tilde{\phi} : M \to \mathbb{R}$ is constant μ-a.s.*

(f) *For each integrable and μ-invariant function $\Psi : M \to \mathbb{R}$ one has $\Psi(x) = \int \Psi d\mu$, μ-a.s.*

(g) *Every integrable and μ invariant function $\Psi : M \to \mathbb{R}$ is constant μ-a.s.*

1.3 The Riesz-Markov-Kakutani Theorem, Prokhorov's Theorem

In this section we discuss a bridge between probability theory and functional analysis around the notion of weak convergence. We use mainly [10] and [1] as references and shall omit proofs.

Consider a metric space (M, d). It is possible to endow the space of probability measures $\mathcal{M}_1(M)$ on M with a topology, called weak topology.

Given a measure μ in $\mathcal{M}_1(M)$ and a finite set $\Phi = \{\phi_1, \ldots, \phi_N\}$ of continuous bounded functions $\phi_i : M \to \mathbb{R}$ and an $\epsilon > 0$, define

$$V(\mu, \Phi, \epsilon) = \{\nu \in \mathcal{M}_1(M) : |\int \phi_i d\nu - \int \phi_i d\mu| < \epsilon \text{ for all } i\}. \tag{7}$$

Any intersection of two such elements contains another element of this form. Thus, $\{V(\mu, \Phi, \epsilon) : \mathcal{M}_1(M), \Phi = \{\phi_1, \ldots, \phi_N\}, \epsilon > 0\}$ can be taken as a basis of neighbourhoods for any $\mu \in \mathcal{M}_1(M)$. The induced toplogy is called the *weak topology.* Note that this topology is Hausdorff. Indeed, if $\int \phi d\mu = \int \phi d\mu$ for all ϕ bounded continuous then $\mu = \nu$. So there is a ϕ and an $\epsilon > 0$ such that $V(\mu, \{\phi\}, \epsilon) \cap V(\nu, \{\phi\}, \epsilon) = \emptyset$.

The usual notion of weak convergence of probability measures is equivalent to convergence in the weak topology (whence its name):

Lemma (weak convergence)**.** *A sequence of probability measures $(\mu)_{n\in\mathbb{N}}$ converges to a probability measure $\mu \in \mathcal{M}_1(M)$ if and only if $\int \phi d\mu_n \to \int \phi d\mu$ for any bounded continuous function $\phi : M \to \mathbb{R}$.*

In the case of (M, d) being a separable metric space, the weak topology on $\mathcal{M}_1(M)$ is metrizable by the Lévy-Prokhorov metric D. For μ, ν in $\mathcal{M}_1(M)$, D is defined by

$$D(\mu, \nu) = \inf\{\delta : \mu(B) \leq \nu(B^\delta) + \delta \text{ and } \nu(B) \leq \mu(B^\delta) + \delta \text{ for every borelian } B\}, \tag{8}$$

where $A^\delta = \{x \in M : d(x, A) < \delta\}$.

We now move on to a central theorem in probability theory, Prokhorov's theorem.

Definition (Tightness)**.** *Suppose we are given a subset $\mathcal{A}$ of $\mathcal{M}_1(M)$ for a separable metric space (M, d). $\mathcal{A}$ is called tight if for all $\epsilon > 0$ one can find a compact set $K_\epsilon \subset M$ such that $\mu(M \setminus K_\epsilon) < \epsilon$ for all μ in $\mathcal{A}$.*

The intuitive idea of tightness is that a given set of measures doesn't "escape to infinity", meaning their mass is concentrated locally.

Theorem (Prokhorov)**.** *A collection $\mathcal{A} \subset \mathcal{M}_1(M)$ is tight if and only if the closure of $\mathcal{A}$ is (sequentially) compact in $(\mathcal{M}_1(M), D)$.*

If (M, d) is a compact metric space, such as $M = \Sigma_A$, then $\mathcal{M}_1(M)$ is compact in its own right and any subset $\mathcal{A}$ is relatively compact and thus tight. Thus the full power of this theorem is only deployed for non-compact spaces M, but we need this information hereafter.

We now proceed to link measure theory with functional analysis. This is done by means of the Riesz-Markov-Kakutani theorem. It should be noted that there is a whole fauna of such theorems, called alternatively Riesz, or Riesz-Markov. They all treat the duality "measure - linear functional", whether the functional be positive, continous, complex valued, and the measure positive, signed or complex, with varying conditions on the underlying topological space. The greatest care is advised in order not to make too hasty conclusions. We refer the reader to [5] for a detailed exposition of these problems.

Theorem (Riesz-Markov-Kakutani)**.** *Let $(X, \mathcal{T})$ be a locally-compact Hausdorff topological space and $C_0(X)$ denote the space of continuous functions on X that vanish at infinity. $\mathcal{B} = \sigma(\mathcal{T})$ denotes the Borelian sigma-algebra. Then there is a bijective isometry $\Phi : \mathcal{M}_{reg}(\mathcal{B}) \to C_0^*(X)$ between the Banach space of regular signed measures on X and the Banach space of continuous functionals on $C_0(X)$. Φ is given by*

$$\Phi(\nu)(f) := \int_X f d\nu \tag{9}$$

for $f \in C_0(X)$ and $\mathcal{M}_{reg}(\mathcal{B})$.

Recall that for a Banach space $(E, \|\cdot\|)$, the canonical injection i of E into its bi-dual E^{**} is given by $i(v)(f) = f(v)$ for any f in E^* and v in E. It is an isometry. In the case that i is surjective, the space E is called reflexive. One can consider on E^* the coarsest topology for which all functions $i(v)$, $v \in E$, are continuous. This topology is denoted by $\sigma(E^*, E)$ and is called the *weak*-topology.* The weak*-topology is always Hausdorff.

Theorem (Banach-Alaoglu-Bourbaki)**.** *Let E be a Banach space. The closed unit ball $B_{E^*} = \{f \in E^* : \|f\| \leq 1\}$ is compact in the weak* topology.*

Theorem (Metrizability of the unit ball)**.** *Let E be a Banach space. Then $(B_{E^*}, \sigma(E^*, E))$ is metrizable if and only if E is separable in the weak* topology.*

Since B_{E^*} is compact and weak*-topology Hausdorff, *the closed unit ball is closed* in the weak*-topology.

In the case of our metric space (M, d) being compact, we have that $E = (C(M), \|\cdot\|_\infty)$ is separable, and thus the unit ball is metrizable in $\sigma(E^*, E)$. Furthermore, M being compact, we have $C_0(M) = C_b(M) = C(M)$. $\mathcal{M}_1(M)$ is thus endowed with two metrics, the one induced by weak*-convergence and the Prokhorov-Lévy metric. But then convergence in either metric is equivalent on $\mathcal{M}_1(M)$, so the weak topology and the weak* topology coincide (and $\mathcal{M}_1(M)$ is a closed subset of B_{E^*} for the weak*-topology).

The situation is completely different when (M, d) is not compact. In this case the set of probability measures is not closed in the weak*-topology. We refer the reader to [6] for a much more detailed treatment of the subject (the correct measure/linear functional duality has to be specified).

1.4 The Schauder-Tychonoff Fixed Point Theorem

We give a concise panorama of fixed point theorems and steer the discussion to the needs of our article. First we start with Euclidean space:

Theorem (Brouwer)**.** *Let F be a finite dimensional space and let $Q \subset F$ be a non-empty compact convex set. Let $f : Q \to Q$ be a continuous map. Then f has a fixed point.*

This result admits the following generalization to Banach spaces:

Theorem (Schauder)**.** *Let E be a Banach space, and let C be a non-empty, closed convex set in E. Let $f : C \to C$ be a continuous map such that $F(C) \subset K$, where K is a compact subset of C. Then F has a fixed point in K.*

A proof of Brouwer's fixed point theorem may be found in [8]. This version of Schauder's fixed point theorem is given as an exercise with hints in [1]. In our present situation we need an even more general fixed-point theorem, since we are dealing with a general topological vector spaces (obtained with the weak*-topology).

Theorem (Schauder-Tychonoff)**.** *Let C be a non-empty compact convex subset of a locally convex topological vector space. Let $f : C \to C$ be a continuous function, then f has a fixed point.*

A proof of this theorem can be found in chapter V.10 of [9]. A very systematic book that covers all the intricacies of the subject is Frank Bonsall's *Lectures on some fixed point theorems of functional analysis* [2]. Another proof of Schauder-Tychnoff is found in its Appendix.

In Bowen's article, the topological vector space is the topological dual $\mathscr{C}(\Sigma_A)^*$ endowed with the weak*-topology (and not with the norm topology!). Σ_A being compact metric, $\mathscr{C}(\Sigma_A)$ is separable. Whence the unit ball of $\mathscr{C}(\Sigma_A)^*$ is metrizable in the weak-* topology. The Riesz-Markov-Kakutani gives a bijective isometry between the space of signed measures on Σ_A and $\mathscr{C}(\Sigma_A)^*$. Σ_A being compact, the Prokhorov metric endows $\mathcal{M}_1(\Sigma_A)$ with a metric whose induced topology is the projection of the weak-* topology on $\mathcal{M}_1(\Sigma_A)$. $\mathcal{M}_1(\Sigma_A)$ is closed in $(\mathscr{C}(\Sigma_A)^*, \sigma(\mathscr{C}(\Sigma_A)^*, \mathscr{C}(\Sigma_A)))$. Hence, the Schauder-Tychonoff theorem can be applies to $\mathcal{M}_1(\Sigma_A)$.

2 Technical aspects of the article revisited

2.1 Gibbs Measures

2.1.1 Another metric on Σ_A

For any $\beta \in (0,1)$ it is possible to endow Σ_n (and thus Σ_A) with another metric d_β that generates the product topology.

For $\underline{x}, \underline{y} \in \Sigma_n$ define $d(\underline{x}, \underline{y}) = \beta^N$ where $N = N(\underline{x}, \underline{y})$ is the largest integer such that $x_i = y_i$ for all $|i| < N$. If there is no such integer, then $x = y$ and one sets the distance to zero. In other words N is the smallest integer such that $\underline{x}$ and $\underline{y}$ differ either at position N or $-N$. From this remark, it is immediate that the triangular inequality holds, since a third element $\underline{z}$ will have to differ with either $\underline{x}$ or $\underline{y}$ (or both...) at position N or $-N$.

The distances are thus "quantized" as powers of β. Now notice that for $N \in \mathbb{N}$ the open ball $\{\underline{x} : d(\underline{y}, \underline{x}) < \beta^N\}$ is nothing but $\{\underline{y} : y_i = x_i \text{ for } |i| \leq N\}$. Since all such former and later elements form a basis for their respective topologies, by proving both bases are actually the same, we have proven both topologies are equal.

Lemma. *Fix $\beta \in (0,1)$. Then $f \in \mathscr{F}_A$ if and only if f is Hölder-continuous with positive constant with respect to d_β.*

Proof. Elementary once one notices that $f \in \mathscr{F}_A$ means there is a $b > 0$ and $\alpha \in (0,1)$ such that $|f(\underline{x}) - f(\underline{y})| \leq b\alpha^{N(\underline{x},\underline{y})}$. □

2.1.2 Identifying $\mathscr{C}(\Sigma_A)$ with $\mathscr{C}(\Sigma_A^+)$ and the transfer operator $\mathcal{L}$

Suppose we are given a function $\phi \in \mathscr{C}(\Sigma_A^+)$. Then we can think of ϕ as an element $\tilde{\phi}$ of $\mathscr{C}(\Sigma_A)$ such that $\tilde{\phi}(\underline{x}) = \tilde{\phi}(\underline{y})$ when $x_i = y_i$ for all $i \geq 0$. To see $\tilde{\phi}$ is indeed continuous, just notice that $\tilde{\phi} = \phi \circ \pi$ where π is the (continuous) projection of Σ_A onto Σ_A^+. Reciprocally, suppose we are given a function $\tilde{\phi} \in \mathscr{C}(\Sigma_A)$ such that $\tilde{\phi}(\underline{x}) = \tilde{\phi}(\underline{y})$ when $x_i = y_i$ for all $i \geq 0$. Then we can think of $\tilde{\phi}$ as an element ϕ of $\mathscr{C}(\Sigma_A^+)$. To see this, write $\phi = \tilde{\phi} \circ i$ where i is an inclusion from Σ_A^+ into Σ_A that should be continuous. For example, fixing n extensions $\ldots x_{-3}x_{-2}x_{-1}$ to the left of x_0 for each value x_0 of $\underline{x}$ we have that i is indeed continuous.

A few comments about the transfer operator $\mathcal{L}$ are now presented. Recall that

$$(\mathcal{L}_\phi f)(\underline{x}) = \sum_{\underline{y} \in \sigma^{-1}\underline{x}} e^{\phi(\underline{y})} f(\underline{y}) \tag{10}$$

Suppose first that for $f \in \mathscr{C}(\Sigma_A^+)$ we have $\mathcal{L}_\phi f$ continuous, then for each $\underline{x}$ our sum is taken on at most n elements. Whence $\left\|\mathcal{L}_\phi f\right\|_\infty \leq n e^{\|\phi\|_\infty} \|f\|_\infty$. So $\mathcal{L}_\phi$ is a bounded operator on $\mathscr{C}(\Sigma_A^+)$. To show continuity, we should notice first that the sum at a value $\underline{x}$ depends on the value of x_0: all possible values y_0 for $\underline{y}$ correspond to the lines of A whose value at column x_0 is 1. We can thus see $\mathcal{L}_\phi$ as a patch of n functions $\mathcal{L}_\phi|_{U_i}$, where $U_i = \{\underline{x} : x_0 = i\}$. These n sets are disjoint, compact (both closed and open), so have a positive distance to each other, and it is sufficient to show continuity on each of them, which is trivial, as a sum of continuous functions.

2.1.3 Note on Lemma 1.12

We give an explicit construction to a technicality in the proof of Lemma 1.12.

It is stated that given $g \in \mathscr{C}(\Sigma_A^+)$ and $\epsilon \geq 0$ one can find an r and two functions f_1 and f_2 in $\mathscr{C}_r$ such that $0 \leq f_2 - f_1 \leq \epsilon$.

For this, given such a g, choose an r such that $\mathrm{var}_r g \leq \epsilon$, and define

$$f_2(\underline{x}) = \max\{g(\underline{y}) : y_i = x_i \text{ for } i \in \{0, \ldots, r\}\}. \tag{11}$$

This maximum exists by compactness. Define f_1 similarly with a minimum. To check continuity, proceed exactly as in section 2.1.4.

2.1.4 Constructing a σ-invariant measure on $\mathscr{C}(\Sigma_A)$

We review in more detail the fundamental step of going from a σ-invariant measure on $\mathscr{C}(\Sigma_A^+)$ to a σ-invariant measure on $\mathscr{C}(\Sigma_A)$. The Schauder-Tychnoff theorem provides us with such a measure μ_ϕ on $\mathscr{C}(\Sigma_A^+)$ for any $\phi \in \mathscr{F}_A \cap \mathscr{C}(\Sigma_A^+)$.

Lemma. *For $f \in \mathscr{C}(\Sigma_A)$ define f^* on Σ_A^+ by*

$$f^*(\{x\}_{i=0}^{\infty}) = min\{f(\underline{y}) : \underline{y} \in \Sigma_A \; and \; y_i = x_i \; \forall i \geq 0\}, \tag{12}$$

then f^* is continuous.

Proof. We recall that Σ_A being compact (and keeping in mind a product metric), the assertions f continuous, f uniformly continuous and $\lim_{k\to\infty} \mathrm{var}_k f = 0$ are equivalent.

Fix a k such that $\mathrm{var}_k f \leq \epsilon$ and choose an $\underline{x} \in \mathscr{C}(\Sigma_A^+)$ and an extension $\underline{\tilde{x}} \in \mathscr{C}(\Sigma_A)$ ($x_i = \tilde{x}_i$ for $i \geq 0$) such that $f^*(\underline{x}) = f(\underline{\tilde{x}})$ (possible by compactness).

Then for any $\underline{y} \in \Sigma_A^+$ with $y_i = x_i$ for $i \in \{0, \ldots, m\}$, we get, by choosing for $\underline{y}$ the same extension as for $\underline{x}$ that $f^*(\underline{y}) \leq f^*(\underline{x}) + \epsilon$. Symmetrically $f^*(\underline{x}) \leq f^*(\underline{y}) + \epsilon$. Whence $|f^*(\underline{x}) - f^*(\underline{y})| \leq \epsilon$ as soon as $x_i = y_i$ for $i = 0, \ldots, m$, so f^* is uniformly continuous on Σ_A^+. □

Noticing (by staring patiently at the formula ...) that for $n, m \geq 0$

$$\|(f \circ \sigma^n)^* \circ \sigma^m - (f \circ \sigma^{n+m})^*\| \leq \mathrm{var}_n f, \tag{13}$$

we get, by σ-invariance, that

$$|\mu((f \circ \sigma^n)^*) - \mu((f \circ \sigma^{n+m})^*)| = |\mu((f \circ \sigma^n)^*) \circ \sigma^m - \mu((f \circ \sigma^{n+m})^*)| \leq \mathrm{var}_n f. \tag{14}$$

This expression goes to 0 as $n \to \infty$, f being continuous. We thus have a Cauchy sequence $(\mu((f \circ \sigma^n)^*)_{n\geq 0}$, whose limit we shall denote by $\tilde{\mu}(f)$. We immediately check that $\tilde{\mu}$ is positive and that $\tilde{\mu}(\chi_{\Sigma_A}) = \lim_{n\to\infty} \mu((\chi_{\Sigma_A} \circ \sigma^n)^*) = 1$. So the Riesz representation theorem tells us that $\tilde{\mu}$ defines a probability. σ-invariance is immediate since $\tilde{\mu}(f \circ \sigma) = \tilde{\mu}(f)$ for any continuous f. Furthermore, seeing an $f \in \mathscr{C}(\Sigma_A^+)$ as an element of $\mathscr{C}(\Sigma_A)$, we get that $\tilde{\mu}(f) = \mu(f)$. In other words $\tilde{\mu}$ extends μ.

2.1.5 Complement to Theorem 1.16

We deal with a small technicality that is left unsolved in the proof of the theorem. We are given two Gibbs measures μ and μ', and have proven that for a certain constant c, $\mu(E_m(\underline{x}) \leq c\mu'(E_m(\underline{x})$ for all m and $\underline{x}$. By σ-invariance, the inequality holds for any cylindrical set. We wish to show the inequality is then true for any borelian set E.

The classical "jack in the box" of probability, the pi-system/Dynkin-system theorem , fails because one can't go to the complementary without changing the inequality. However, this theorem has an equivalent version:

Theorem (monotone class theorem)**.** *Let $\mathcal{A}$ be an algebra of sets, and let $\mathcal{M}(\mathcal{A})$ be the smallest monotone class containing $\mathcal{A}$. Then $\mathcal{M}(\mathcal{A}) = \sigma(\mathcal{A})$.*

The algebra of sets $\mathcal{A}$ is then simply the set of all finite unions of cylindrical sets. We define $\mathscr{C} = \{U \in \mathcal{B}(\Sigma_A) : \mu(U) \leq c\mu'(U)\}$. $\mathcal{A} \subset \mathscr{C}$, because each element of $\mathcal{A}$ can be written as a disjoint union of cylindrical sets, and one readily checks that $\mathscr{C}$ is a monotone class. Whence $\sigma(\mathcal{A}) \subset \mathscr{C}$.

2.1.6 Complement to Theorem 1.22

We explicit the calculation showing that $\nu' = \mu$.

Recall that $\nu' = \frac{d\nu'}{d\mu}\mu$. On the one hand, by σ-invariance, one has that

$$\begin{aligned} \nu'(A) &= \int_A \frac{d\nu'}{d\mu} d\mu \\ &= \int (\chi_A \circ \sigma)(\frac{d\nu'}{d\mu} \circ \sigma) d\mu. \end{aligned} \tag{15}$$

On the other, by ν-invariance:

$$\begin{aligned} \nu'(A) &= \int (\chi_A \circ \sigma) d\nu' \\ &= \int (\chi_A \circ \sigma) \frac{d\nu'}{d\mu} d\mu. \end{aligned} \tag{16}$$

Setting $A' = \sigma(A)$, we get for all $A' \in \mathcal{B}(\Sigma_A)$ that $\int_{A'} (\frac{d\nu'}{d\mu} \circ \sigma) d\mu = \int_{A'} \frac{d\nu'}{d\mu} d\mu$, so that $\frac{d\nu'}{d\mu} \circ \sigma = \frac{d\nu'}{d\mu}$, μ almost surely. But since σ is ergodic, $\frac{d\nu'}{d\mu}$ is almost surely equal to a constant c. Whence $c = 1$ and we are done.

2.1.7 Lemma 1.29 Revisited

We recall the following. . .

Definition (Baire Space)**.** *A Baire Space is a topological space such that every intersection of a countable collection of open dense sets in the space is also dense.*

. . . as well as the classical

Theorem (Baire Category theorem)**.** *If X is a complete metric space or a locally compact Hausdorff space, then it is a Baire space.*

For the first claim, see Munkres [7], chapter 48. For the second, see Dugundji [4], chapter 11, section 10. Note that Munkres proves the second claim in the special case of a compact Hausdorff space.

Definition (Topological transitivity)**.** *A continuous function $f : M \to M$ on a topological space M is said to be topologically transitive if there exists an $x \in M$ such that $\left\{f^n(x) : n \in \mathbb{N}^*\right\}$ is dense in M.*

It should be noted that different definitions of topological transitivity exist, for instance replacing $\mathbb{N}^*$ by $\mathbb{N}$, which would be more coherent will the usual notion of orbit. Our definition, as in [10], even if stronger, insures we have some notion of periodicity (since for any open set U in the vicinity of x we will then have a $k \geq 1$ with $f^k(x) \in U$). It is not clear which notion is used in Bowen's lecture notes, however this one is sufficient for our purpose.

We now proceed to proving an equivalent formulation of topological transitivity for second-countable spaces (found in [10]).

Lemma. *Suppose that M is a is second-countable Baire space. A continuous function $f : M \to M$ is topologically transitive if and only if for each pair of open sets U and V there exists a $k \geq 1$ such that $f^{-k}(U)$ intersects V.*

Proof. Suppose first that f is transitive and let $x \in M$ be a point whose orbit $\left\{f^n(x) : n \in \mathbb{N}^*\right\}$ is dense in M. We first notice that $\left\{f^n(x) : n > m\right\}$ is then dense in M for any $m \in \mathbb{N}$. So there is an $m \geq 1$ and an $n > m$ such that $f^m(x) \in V$ and $f^n(x) \in U$. Denoting $k = n - m$, we get that $f^m(x) \in f^{-k}(U) \cap V$.

To prove the converse, let $\{U_j : j \in \mathbb{N}^*\}$ be a countable basis of open sets of M. Our hypothesis guarantees that $\cup_{k=1}^{\infty} f^{-k}(U_j)$ is dense in M for any $j \in \mathbb{N}^*$. Then

$$X = \bigcap_{j=1}^{\infty} \bigcup_{k=1}^{\infty} f^{-k}(U_j) \tag{17}$$

is a dense (in particular non-empty) subset of M (because M is a Baire space). On the other hand, if $x \in X$ then for any $j \in \mathbb{N}^*$ there is a $k \geq 1$ such that $f^k(x) \in U_j$. Since the U_j form a basis of the topology of M, we proved that $\{f^k(x) : k \in \mathbb{N}^*\}$ is dense in M.

□

From this characterisation we deduce the following corollary:

Corollary (Lemma 1.29)**.** *Let $f : M \to M$ be a continuous topologically transitive continuous map of a compact metric space M, then there is a point $x \in M$ such that for any $U \neq \varnothing$ and $N > 0$ there exists an $n \geq N$ with $T^n(x) \in U$.*

Proof. X being compact metric, it is a second-countable Baire space. Denote by $(U_i, i \geq 1)$ a basis for the topology. Fix an $N > 0$. By the previous lemma,

$$V_{i,N} = \bigcup_{n \geq N} T^{-n} U_i \tag{18}$$

is a dense open set in M. Whence, by the Baire Category Theorem,

$$V = \bigcap_{(n,i) \in \mathbb{N}^* \times \mathbb{N}^*} T^{-n} U_i \tag{19}$$

is a dense set. Any $x \in V$ does the job. □

The very next step in the continuation of the article is to notice that since σ is bijective, topological mixing implies topological transitivity. This is immediate one more time by the characterisation above.

2.2 General Thermodynamic Formalism

2.2.1 A few notes on expansivity

Let (X, d) be a compact metric space, recall the following

Definition (expansivity)**.** *A homeomorphism $T : X \to X$ is called expansive if there exists $\epsilon > 0$ such that $d(T^k(x), T^k(y)) \leq \epsilon \; \forall \; k \in \mathbb{Z}$ implies $x = y$.*

In other words, taking the contraposition, if $x \neq y$, there is a $k = k(x, y)$ such that $T^k(x)$ and $T^k(y)$ end up being at a distance superior to ϵ.

We prove Proposition 2.5 in more detail and explain an interesting consequence, that connects the first two chapters of Bowen's lecture notes.

Lemma (Proposition 2.5)**.** *Suppose $T : X \to X$ is a homeomorphism with expansive constant ϵ. Then $h_\mu(T) = h_\mu(T, \mathcal{D})$ whenever $\mu \in \mathcal{M}_T(X)$ and $diam(\mathcal{D}) \leq \epsilon$.*

Proof. Set $\mathcal{D}_n = T^n\mathcal{D} \vee \ldots \mathcal{D} \vee \ldots T^{-n}\mathcal{D}$. Suppose that $\text{diam}(\mathcal{D}_n)$ does not go to zero. Considering that $\mathcal{D}_{n+1}$ is finer than $\mathcal{D}_n$ there would then exist two distinct points $x \neq y$ such that both x and y are in the same element of $\mathcal{D}_n$ for each n, but then $d(T^n(x), T^n(y)) \leq \epsilon$ for all $n \in \mathbb{Z}$, a contradiction. Thus $\text{diam}(\mathcal{D}_n)$ goes to 0. By proposition 2.4, $h_\mu(T) = lim_n h_\mu(T, \mathcal{D}_n)$. By Lemma 2.2(c), $h_\mu(T, \mathcal{D}_n) = h_\mu(T, \mathcal{D})$. ☐

We're now in a position of showing that s_μ, defined as $h_\mu(\sigma,\mathcal{U})$ for $\mathcal{U} = \{U_1, \ldots, U_n\}$ with $U_i = \{\underline{x} \in \Sigma_A \; x_0 = i\}$ depends exclusively of σ and not on both σ and $\mathcal{U}$. Indeed, notice that the U_i form a partition of (open and compact) sets of Σ_A. Thus for $i \neq j$ we have that $d(U_i, U_j) > 0$. Taking c strictly smaller than the minimum of all such values, we find that σ is expansive with expansivity constant c (each component being shifted to zero). Considering a usual metric on Σ_A we can take n large enough such that each element of $\mathcal{D}_n = \sigma^n\mathcal{U} \vee \ldots \mathcal{U} \vee \ldots \sigma^{-n}\mathcal{U}$ has diameter less or equal to c (this is coherent with the definition of $h_\mu(\sigma,\mathcal{U})$ because by invariance one can shift this whole expression by σ^{-n}; we already used this trick implicitly in the lemma above). By the lemma, we get $s_\mu = h_\mu(\sigma)$. Hence, s_μ depends on σ only, as a σ-invariant measure.

The reader interested in further properties of expansivity in relation to entropy and equilibrium states can (should!) refer to chapters 9 and 10 of [10].

2.2.2 Connecting the two notions of pressure $P(\phi)$

We outline how to reconcile both notions of pressure. For this, using the notation of general thermodynamic formalism (chapter 2), we must prove that $P(\phi,\mathcal{U}) = P(\phi)$ on Σ_A, where $\mathcal{U} := \{U_1, \ldots, U_n\}$ with $U_i = \{\underline{x} : x_0 = i\}$ and $P(\phi) = \lim_{\text{diam}(\mathcal{V}) \to 0} P(\phi, \mathcal{V})$.

We introduce the following notation for $j \in \mathbb{Z}$

$$\mathcal{U}^j = \{U_1^j, \ldots, U_n^j\} \text{ with } U_i^j = \{\underline{x} : x_j = i\}. \tag{20}$$

In other terms $\mathcal{U}^j$ is a partition and open covering of (infinite) words by value at position j. We further introduce for $j \in \mathbb{Z}$ and $k \geq 1$

$$\mathcal{U}^{j,k} = \mathcal{U}^j \vee \sigma\mathcal{U}^j \ldots \vee \sigma^{k-1}\mathcal{U}^j \tag{21}$$

In other terms $\mathcal{U}^{j,k}$ is a partition and open covering of words whose k letters starting at position j are fixed.

The philosophy is the following: by a simple bijection, one can prove that for a fixed $k \geq 1$

$$Z(\phi, \mathcal{U}^{j_1,k}) = Z(\phi, \mathcal{U}^{j_2,k}) \tag{22}$$

for any j_1 and j_2. This translates into

$$P(\phi, \mathcal{U}^{j_1,k}) = P(\phi, \mathcal{U}^{j_2,k}) \tag{23}$$

for any j_1 and j_2.

Then, by an argument using uniform continuity and easy combinatorics on the partition function, one shows that

$$P(\phi, \mathcal{U}^{j,1}) = P(\phi, \mathcal{U}^{j,k}) \tag{24}$$

for any $k \geq 1$. With our terminology, we have $\mathcal{U}$=$\mathcal{U}^{0,1}$, and we have shown how to successively reduce the diameter of the partition to zero (by increasing the length of the word and shifting it left) without changing the value of the pressure. Whence our stated equality holds.

2.2.3 Limiting measures as invariant measures

We explain in more details the discussion given before Lemma 2.15 in Bouwer's lecture notes and rephrase it in a more general set-up. We refer the reader to section 1.3 and to chapter 2 of [10].

Lemma (continuity of push-forward)**.** *Let (M, d) be a metric space and $f : M \to M$ a continuous function. Then the pushforward application $f_* : \mathcal{M}_1(M) \to \mathcal{M}_1(M)$ is continuous with respect to the weak*-topology.*

Proof. Let $\Phi = \{\phi_1, \ldots, \phi_n\}$ be any family of continuous bounded functions on M. Since f is continuous, $\Psi = \{\phi_1 \circ f, \ldots, \phi_n \circ f\}$ also is a family of continuous bounded functions on M. But for μ, ν in $\mathcal{M}_1(M)$ we also have that

$$|\int \phi_i d(f_*\mu) - \int \phi_i d(f_*\nu)| = |\int (\phi_i \circ f) d(f_*\mu) - \int (\phi_i \circ f) d(f_*\nu)|, \tag{25}$$

which shows that $f_*(V(\mu, \Psi, \epsilon) \subset V(f_*\mu, \phi, \epsilon)$ for all μ, Φ and $\epsilon > 0$. This shows that f_* is continuous for the weak*-topology. □

For ν in $\mathcal{M}_1(M)$ we are lead to study the limiting points of series

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \nu. \tag{26}$$

For (M, d) compact, such limiting points always exist, $\mathcal{M}_1(M)$ being compact metric. We have the general

Theorem. *Any limiting point μ of a sequence μ_n as in (26) is a fixed point of f_*. In other words, μ is an invariant measure.*

Proof. Suppose $\mu_{n_k} \to \mu$ as $k \to \infty$. Then for any $\epsilon > 0$ and any family $\Phi = \{\phi_1, \ldots, \phi_n\}$ of continuous bounded functions we have

$$|\frac{1}{n_k} \sum_{j=0}^{n_k-1} \int (\phi_i \circ f^j) d\nu - \int \phi_i d\mu| < \epsilon/2 \tag{27}$$

for each i and each k large enough. By our lemma (continuity of the push-forward):

$$\begin{aligned} f_*\mu &= f_*(\lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f_*^j \nu) \\ &= \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=1}^{n_k} f_*^j \nu. \end{aligned} \tag{28}$$

Now observe that

$$\begin{aligned} |\frac{1}{n_k} \sum_{j=0}^{n_k-1} \int (\phi_i \circ f^j) d\nu - \frac{1}{n_k} \sum_{j=1}^{n_k} \int (\phi_i \circ f^j) d\nu| &= \frac{1}{n_k} | \int \phi_i d\nu - \int (\phi_i \circ f^{n_k}) d\nu| \\ &\leq \frac{1}{n_k} \sup |\phi_i| \end{aligned} \tag{29}$$

with this expression being smaller than $\epsilon/2$ for all i and any k sufficiently large. We thus get

$$|\frac{1}{n_k} \sum_{j=1}^{n_k} \int (\phi_i \circ f^j) d\nu - \int \phi_i d\mu| < \epsilon \tag{30}$$

for each i and any k sufficiently large. This reads

$$\frac{1}{n_k} \sum_{j=1}^{n_k} f_*^j \nu \to \mu \text{ as } k \to \infty. \tag{31}$$

We showed above that this sequence converges to $f_*\mu$. By unicity of the limit, $f_*\mu = \mu$ follows. □

3 Errata for the Article

The convention "line minus x" is used to indicate one should go to line x counting from the bottom of the page.

- Page 6, line 1. One should read: $i = 0, \dots, m-1$ instead of $i = 0, \dots, m$.
- Page 7, in the statement of Lemma 1.5. One should specificy $\psi, \phi \in \mathscr{F}_A$.
- Page 10, line 8. The right parenthesis should be moved to before $f(\underline{y})$.
- Page 11, lines 5, 6 and 7. The constants u_1 and u_2 should be interchanged.
- Page 14, line 1. Read $f \in \mathscr{C}(\Sigma_A^+)$ instead of $f \in \mathscr{C}(\Sigma_A)^*$.
- Page 15, in the statement of Theorem 1.16. One should read: μ is the unique Gibbs measure for $\phi \in \mathscr{F}_A \cap \mathscr{C}\left(\Sigma_A^+\right)$.
- Page 16, lines 18, 19 and 20. We're actually dealing with pull-backs. In Bowen's article, the usual convention $T_*\mu$ and $T^*\mu$ for the push-forward and pull-back of a measure μ by a function T have been swapped, except here. Indeed (using the standard convention), for any A borelian we have

$$\begin{aligned} \int \chi_A d((\sigma^{-1})_\star \mu') &= \int \chi_A \circ \sigma^{-1} \ d\mu' \\ &= \int (\chi_A \circ \sigma^{-1}) f d\mu \\ &= \int \chi_A (f \circ \sigma) d((\sigma^{-1})_\star \mu). \end{aligned} \tag{32}$$

- Page 19, line -3. One should read $\underline{x}$ instead of $\underline{y}$ in the second term on the right-hand side of the inequality.
- Page 21, line 3. Replace $\log a_i$ by $\log p_i$.
- Page 20, line 4. The inequality can be replaced by an equality.
- Page 21, line -6. The inequality can be replaced by an equality.
- Page 22, lines 3, 4. Replace c_2 by c_1.
- Page 25, line 25. One should read "...topological mixing is stronger than topological transitivity ...".
- Page 25, line -7. Replace $u\left(\underline{w}\right)$ by $u\left(\underline{z}\right)$
- Page 31, line 4. Read $\frac{m+n}{m}$ instead of $\frac{m}{m+n}$.
- Page 32, 4$^{\text{th}}$ line of the proof of Lemma 2.6. Cross out "$\dots = n_\mu(T)$".
- Page 35, line -2 of the proof of Lemma 2.11. The inequality should be crossed out and replaced by: ...whence, this holding for any $\Gamma_m \in W_m(\mathcal{U})$, we get

$$h_\mu(T, \mathcal{D}) + \int \phi \ \mathrm{d}\mu \le logM + \frac{1}{m} logZ_m(\phi, \mathcal{U}). \tag{33}$$

- Page 37, line 6. The equality should be replaced by an inequality because Γ, Γ^2, Γ^3, ... are note necessarily disjoint.
- Page 37, lines 13 and 14. The notation Γ^N for "Γ^N the set of $\underline{U}^*$ so obtained" overlaps with the previous distinct notion of $\Gamma^n = \{\underline{U}_1\underline{U}_2 \ldots \underline{U}_n : \underline{U}_i \in \Gamma\}$. Note Γ_N instead for instance.
- Page 41, line -6. The inequality can be replaced by an equality.
- Page 43, line 3. Replace 2 by 3 (using the triangular inequality).

4 Summary of chapter 1: Gibbs Measures

4.1 Construction of Gibbs Measures and a few further properties

The goal of this first chapter is to prove the following

Theorem (Existence of Gibbs measures, preliminary formulation)**.** *Suppose $\phi : \Sigma_n \to \mathbb{R}$ and there are $c > 0$ and $\alpha \in (0,1)$ so that $var_k\phi \leq c\alpha^k$ for all k. Then there is a unique $\mu \in \mathcal{M}(\Sigma_n)$ for which one can find constants $c_1 > 0$, $c_2 > 0$, and P such that*

$$c_1 \leq \frac{\mu\{\underline{y} : y_i = x_i \; \forall \; i = 0, \dots, m-1\}}{\exp(-Pm + \sum_{k=0}^{m-1} \phi(\sigma^k \underline{x}))} \leq c_2 \tag{34}$$

for every $\underline{x} \in \Sigma_n$ and $m \geq 0$.

Such functions are denoted by $\phi \in \mathscr{F}_A$ and the measure μ or μ_ϕ associated to ϕ is called a *Gibbs measure*. The discussion is generalized by considering Σ_A instead for a topologically mixing system ($\sigma : \Sigma_A \to \Sigma_A$ is topologically mixing if for U, V open subsets of Σ_A there is an N such that $\sigma^m U \cap V \neq \emptyset$ for $m \geq N$). Topological mixing is proven to be equivalent to there being an $M > 0$ such that $A^M > 0$. The preliminary formulation is a particular case of the following theorem, with A a matrix whose entries are 1's.

Theorem (Existence of Gibbs measures, general version)**.** *Suppose Σ_A is topologically mixing and $\phi \in \mathscr{F}_A$. There is a unique σ-invariant Borel probability measure μ on Σ_A for which one can find constants $c_1 > 0$, $c_2 > 0$ and P such that*

$$c_1 \leq \frac{\mu\{\underline{y} : y_i = x_i \; \forall \; i = 0, \dots, m-1\}}{\exp(-Pm + \sum_{k=0}^{m-1} \phi(\sigma^k \underline{x}))} \leq c_2 \tag{35}$$

for every $\underline{x} \in \Sigma_A$ and $m \geq 0$. Furthermore, ϕ is ergodic.

The strategy to tackle the proof, loosely speaking, will be to reduce the problem to Σ_A^+ with $\phi \in \mathscr{F}_A \cap \mathscr{C}(\Sigma_A^+)$, solve it on Σ_A^+ (proceeding to standard identifications, see section 2), and then extend the obtained measure back to Σ_A.

This strategy is motivated by the following few considerations/results.

Definition (homology)**.** *Two functions ψ, ϕ in $\mathscr{C}(\Sigma_A)$ are homologous with respect to σ if there is a $u \in \mathscr{C}(\Sigma_A)$ so that $\psi(\underline{x}) = \phi(\underline{x}) - u(\underline{x}) + u(\sigma\underline{x})$.*

Lemma. *If ϕ_1, ϕ_2 are homologous and Gibbs' theorem holds for ϕ_1, then it holds for ϕ_2 for c_1, c_2, and P unchanged.*

Lemma. *If $\phi \in \mathscr{F}_A$, then ϕ is homologous to some $\psi \in \mathscr{F}_A$ with $\psi(\underline{x}) = \psi(\underline{y})$ whenever $x_i = y_i$ for all $i \geq 0$.*

We recall the defnition of the transfer operator $\mathcal{L} = \mathcal{L}_\phi$ on $\mathscr{C}(\Sigma_A^+)$ for $\phi \in \mathscr{C}(\Sigma_A^+)$:

$$(\mathcal{L}_\phi f)(\underline{x}) = \sum_{\underline{y} \in \sigma^{-1}\underline{x}} e^{\phi(\underline{y})} f(\underline{y}). \tag{36}$$

The "atomic bomb to crack a nut" that will solve our problem is the Schauder-Tychonoff theorem *Let E be a non-empty compact convex subset of a locally convex topological vector space, then any continuous function $G : E \to E$ has a fixed point.* It is implemented twice to prove the central

Theorem (Ruelle's Perron-Frobenius Theorem)**.** *Let Σ_A be topologically mixing, $\phi \in \mathscr{F}_A \cap \mathscr{C}(\Sigma_A^+)$ and $\mathcal{L} = \mathcal{L}_\phi$ as above. There are $\lambda > 0$, $h \in \mathscr{C}(\Sigma_A^+)$ with $h > 0$ and $\nu \in \mathcal{M}(\Sigma_A^+)$ for which $\mathcal{L}h = \lambda h$, $\mathcal{L}^*\nu = \lambda\nu$, $\nu(1)$ and*

$$\lim_{m\to\infty} \left\|\lambda^{-m}\mathcal{L}^m g - \nu(g)h\right\| = 0 \text{ for all } g \in \mathscr{C}(\Sigma_A^+). \tag{37}$$

Since $\mathcal{L}$ is a positive operator with $\mathcal{L}1 > 0$, one gets by the Riesz-Markov-Kakutani theorem that $G(\mu) = (\mathcal{L}^*\mu)^{-1}\mathcal{L}^*\mu \in \mathcal{M}(\Sigma_A^+)$ for $\mu \in \mathcal{M}(\Sigma_A^+)$. G is trivially seen to be continuous. One can apply the Schauder-Tychonoff theorem to the set $E = \mathcal{M}(\Sigma_A^+)$ and find a probability measure ν such that $G(\nu) = \nu$. In other words $\mathcal{L}^*\nu = \lambda\nu$ with $\lambda = (\mathcal{L}^*\nu)(1)$.

The remainder of the proof of Ruelle's theorem consists in a series of very technical lemmas, the most important of which we state here. To this end, we analyse then set

$$\Lambda = \{f \in \mathscr{C}(\Sigma_A^+) : \ f \geq 0, \ \nu(f) = 1, \ f(\underline{x}) \leq B_m f(\underline{x}') \text{ when } x_i = x_i' \text{ for all } i = 0, \dots, m\} \tag{38}$$

where $B_m = \exp(\sum_{k=m+1}^{\infty} 2b\alpha^k)$ and $b > 0$, $\alpha \in (0,1)$ are such that $\mathrm{var}_k\phi \leq b\alpha^k$ for all $k \geq 0$. One then shows

Lemma. *There is and $h \in \Lambda$ with $\mathcal{L}h = \lambda h$ and $h > 0$.*

Part of the proof consists of showing that the operator $\lambda^{-1}\mathcal{L}$ is defined on Λ, and that Λ is compact by means of the Ascoli-Arzela theorem. One then applies Schauder-Tychonoff once more to $\lambda^{-1}\mathcal{L}$.

We now have the necessary tools in hand to construct an invariant measure on Σ_A. For $\phi \in \mathscr{F}_A \cap \mathscr{C}(\Sigma_A^+)$ and ν, h, λ given by Ruelle's Perron-Frobenius Theorem, $\mu = h\nu$ is a probability measure on Σ_A^+ with $\mu(f) = \nu(hf) = \int f(\underline{x})h(\underline{x})d\nu(\underline{x})$. Showing that σ is invariant is pretty straight-forward:

Lemma. *μ is invariant under $\sigma : \Sigma_A^+ \to \Sigma_A^+$*

Proof. This is equivalent to showing that $\mu(f) = \mu(f \circ \sigma)$ for any $f \in \mathscr{C}(\Sigma_A^+)$. First note that

$$\begin{aligned} ((\mathcal{L}f)\cdot g)(\underline{x}) &= \sum_{\underline{y}\in\sigma^{-1}\underline{x}} f(\underline{y})g(\underline{x}) \\ &= \sum_{\underline{y}\in\sigma^{-1}\underline{x}} f(\underline{y})g(\sigma\underline{y}) \\ &= \mathcal{L}(f\cdot(g\circ\sigma))(\underline{x}), \end{aligned} \tag{39}$$

This yields

$$\begin{aligned} \mu(f) &= \nu(hf) \\ &= \nu(\lambda^{-1}\mathcal{L}h\cdot f) \\ &= \lambda^{-1}\nu(\mathcal{L}(h\cdot(f\circ\sigma))) \\ &= \lambda^{-1}(\mathcal{L}^*\nu)(h\cdot(f\circ\sigma)) \\ &= \nu(h\cdot(f\circ\sigma)) \\ &= \mu(f\circ\sigma). \end{aligned} \tag{40}$$

□

Extending μ to a σ-invariant measure $\tilde{\mu}$ on Σ_A from Σ_A^+ is an important but simple technicality that is discussed in more detail in section 2. It should be (crucially) noted that $\tilde{\mu}(f) = \mu(f)$ for $f \in \mathscr{C}(\Sigma_A^+)$ (with the usual identification of seeing a function f on Σ_A such that $f(\underline{x}) = f(\underline{y})$ when $x_i = y_i,\ i \geq 0$ as a function on Σ_A^+, and conversely), which is why we shall say that $\tilde{\mu}$ "extends" μ. In the sequel, we shall denote μ indifferently for μ or $\tilde{\mu}$.

The very next step is showing that μ is mixing for $\sigma : \Sigma_A \to \Sigma_A$ and thus ergodic.

One is finally left with the culmination of the exposition, which is showing that μ (or μ_ϕ) is indeed a Gibbs measure (satisfying the inequality with the constants c_1 and c_2), and finally that it is the unique such measure. That μ is indeed a Gibbs measure uses the technical machinery developed for Ruelle's Perron-Frobenius theorem. Unicity uses an "algebra - monotone class" argument followed by a Radon-Nikodym argument and is mostly classical.

A few more results are discussed in the final section of chapter 1 of Bowen's lecture notes. We shall simply mention that conditions under which $\mu_\phi = \mu_\psi$ are provided (for Σ_A topologically mixing and ϕ, ψ in $\mathscr{F}_A$) and present one last theorem.

To this end, we shall introduce the space $\mathscr{H}_\alpha$ of functions $f \in \mathscr{C}(\Sigma_A)$ such that $\mathrm{var}_k f \leq c\alpha^k$ for some c and fixed α. $\mathscr{H}_\alpha$ is then a Banach space under the norm

$$\|f\|_\alpha = \|f\| + \sup_{k\geq 0}(\alpha^{-k}\mathrm{var}_k f). \tag{41}$$

Theorem (Exponential Cluster Property)**.** *For a fixed $\alpha \in (0,1)$ there are constants $D > 0$ and $\gamma \in (0,1)$ such that*

$$|\mu(f \cdot (g \circ \sigma^n))| \leq D\|f\|_\alpha \|g\|_\alpha \, \gamma^n \tag{42}$$

for all f, g in $\mathscr{H}_\alpha$, $n \geq 0$.

4.2 The Variational Principle

The Gibbs measures μ_ϕ for $\phi \in \mathscr{F}_A$ are the only ones satisfying a *variational principle.* A series of definitions is needed at first. It should be kept in mind that these definitions and properties are special cases of the more general theory of thermodynamics presented in the next section. They are fully sufficient, however, for the purpose of this variational principle.

For a (measurable) partition $\mathcal{C} = \{C_1, \ldots ; C_k\}$ of a probability space $(X, \mathcal{B}, \mu)$, the *entropy of* $\mathcal{C}$ is defined as

$$H_\mu(\mathcal{C}) = \sum_{i=1}^{k}(-\mu(C_i)\log\mu(C_i)). \tag{43}$$

If $\mathcal{D}$ is another finite partition of X, $\mathcal{C} \vee \mathcal{D} := \{C_i \cap D_j : C_i \in \mathcal{C}, D_j \in \mathcal{D}\}$. $H_\mu(\mathscr{C}) \geq 0$ and the calculus that results from this definition is based on the concave positive function $\phi(x) = -x\log(x)$ with $\phi(0) = 0,\ x \in [0,1]$ and sub-additivity: one shows that $H_\mu(\mathscr{C} \vee \mathcal{D}) \leq H_\mu(\mathscr{C}) + H_\mu(\mathcal{D})$.

Lemma. *If $\mathcal{D}$ is a finite partition of $(X, \mathcal{B}, \mu)$ and $T : X \to X$ a μ-invariant measurable function, then the limit*

$$h_\mu(T, \mathcal{D}) = \lim_{m\to\infty} \frac{1}{m} H_\mu(\mathcal{D} \vee T^{-1}\mathcal{D} \ldots \vee T^{-m+1}\mathcal{D}) \tag{44}$$

exists in $[0, \infty)$.

Definition (entropy of μ)**.** *For $\mu \in \mathcal{M}_\sigma(\Sigma_A)$ and the partition $\mathcal{U} = \{U_i, \ldots, U_n\}$ where $U_i = \{\underline{x} \in \Sigma_A : \ x_0 = i\}$, the quantity $s(\mu) := h_\mu(\sigma,\mathcal{U})$ is called the entropy of μ.*

We now define pressure directly translating the given definitions leading to it in the language of general thermodynamic formalism (chapter 2 of Bowen's notes).

For this partition (as well as open cover) $\mathcal{U}$ we can define the set $W_m(\mathcal{U})$ of all m-strings $\underline{U} = U_{i_0}U_{i_1} \ldots U_{i_{m-1}}$. One writes $m = m(\underline{U})$ and define

$$X(\underline{U}) = \{x \in X : \sigma^k x \in U_{i_k} \text{ for } k = 0, \ldots, m-1\}. \tag{45}$$

This set will be non-empty if and only if $A_{i_k,i_{k+1}} = 1$ for $k = 0, \ldots, m-1$. Define then

$$S_m(\phi)(\underline{U}) = \sup \left\{ \sum_{k=0}^{m-1} \phi(\sigma^k x) : x \in X(\underline{U}) \right\} \tag{46}$$

and $S_m\phi(\underline{U}) = -\infty$ if $X(\underline{U}) = \emptyset$. Finally define the *partition function*

$$Z_m(\phi) := \sum_{\underline{U}\in W_m(\mathcal{U})} \exp(S_m\phi(\underline{U})). \tag{47}$$

Lemma. *For $\phi \in \mathscr{C}(\Sigma_A)$, the limit $P(\phi) := \lim_{m\to\infty} \frac{1}{m} \log Z_m(\phi)$ exists and is called the pressure of ϕ.*

Theorem. *Suppose $\phi \in \mathscr{C}(\Sigma_A)$ and $\mu \in \mathcal{M}_\sigma(\Sigma_A)$. Then*

$$s(\mu) + \int \phi d\mu \leq P(\phi). \tag{48}$$

Theorem (Variational Principle)**.** *Let $\phi \in \mathscr{F}_A$, Σ_A topologically mixing and μ_ϕ the Gibbs measure of ϕ. Then μ_ϕ is the only $\mu \in \mathcal{M}_\sigma(\Sigma_A)$ for which*

$$s(\mu) + \int \phi d\mu = P(\phi). \tag{49}$$

The proof entails showing that actually $P(\phi) = P$, where P is given with the constants c_1 and c_1 for the Gibbs measure μ_ϕ. The equality is then showed to be verified for μ_ϕ. Supposing another measure ν satisfied the variational principle, one yields a contradiction by first supposing ν singular to μ and that $\nu = \mu$ otherwise. The proof is tricky and very technical.

5 Summary of chapter 2: General Thermodynamic Formalism

The objective of the second chapter is to rephrase the machinery developed in the first chapter in a more general set-up called general thermodynamic formalism. Entropy and pressure notions are given and a variational principle is derived.

Definition (entropy)**.** *Let T be a measure-preserving endomorphism of a probability space (X, μ) and $\mathcal{D}$ denote a finite partition of X. The entropy of μ with respect to T is defined by*

$$h_\mu(T) = \sup_{\mathcal{D}} h_\mu(T, \mathcal{D}) \tag{50}$$

where $\mathcal{D}$ ranges over all finite partitions of X.

Definition (conditional entropy)**.** $H_\mu(\mathscr{C}|\mathcal{D}) := H_\mu(\mathscr{C} \vee \mathcal{D}) - H_\mu(\mathcal{D})$.

Intuitively, conditional entropy measures the additional information given by partition $\mathscr{C}$ once we know the information given by partition $\mathcal{D}$. $H_\mu(\mathscr{C}|\mathcal{D}) \geq 0$. If $\mathscr{C} = \mathcal{D}$ it is worth zero.

A "calculus of entropy" is then developed, that culminates in the following lemma, the proof of which essentially relies on the Lebesgue ϵ for finite open coverings of compact metric spaces.

Lemma. *Suppose $T : X \to X$ is a continuous map on a compact metric space, $\mu \in \mathcal{M}_T(X)$ and that $\mathcal{D}_n$ is a sequence of partitions with diam($\mathcal{D}_n$) $\to 0$, then*

$$h_\mu(T) = \lim_{n\to\infty} h_\mu(T, \mathcal{D}_n) \tag{51}$$

Using the concept of expansivity, the entropy $s(\mu)$ previously defined for Σ_A is shown to be equal to the more general concept of entropy of this chapter: $s(\mu) = h_\mu(\sigma)$. See section 2.

The next step is to generalize the notion of pressure. We work with a continuous map $T : X \to X$ on a compact metric space X and a $\phi \in \mathscr{C}(X)$. For any open cover $\mathcal{U}$ of X,we define the set of m-strings $W_m(\phi)$, $X(\underline{U})$, $S_m(\phi)(\underline{U})$ in exactly the same way as in the first chapter. One then says that $\Gamma \subset W_m(\underline{U})$ covers X if $X = \cup_{\underline{U}\in\Gamma} X(\underline{U})$. One finally defines the partition function

$$Z_m(\phi, \mathcal{U}) = \inf_{\Gamma} \sum_{\underline{U}} \exp(S_m\phi(\underline{U})) \tag{52}$$

where Γ runs over all subsets of $W_m(\underline{U})$ covering X. It's important to notice that in the standard open cover $\mathcal{U} = \{U_1, \ldots, U_n\}$ with $U_i = \{\underline{x} : x_0 = i\}$ of first chapter of the lecture notes, $W_m(\mathcal{U}$ itself forms a partition of X (σ being a homeomorphism), whence the infimum is superfluous. Two lemmas are derived

Lemma. *The limit*

$$P(\phi, \mathcal{U}) = lim_{m\to\infty} \frac{1}{m} \log Z_m(\phi, \mathcal{U}) \tag{53}$$

exists and is finite.

Lemma. *The limit*

$$P(\phi) = lim_{diam(\mathcal{U})\to 0} P(\phi,\mathcal{U}) \tag{54}$$

exists (but may be infinite) and is called pressure of ϕ.

For now, the pressure $P(\phi)$ defined in chapter 1 of the lecture notes with ϕ in $\mathcal{C}(\Sigma_A)$ is what we defined here as $P(\phi,\mathcal{U})$. For a direct explanation showing why both definitions coincide, we refer to section 2. Expos-facto, we can show this is the case after exposing the culmination of our theory:

Theorem (Variational Principle)**.** *Let $T : X \to X$ be a continuous map on a compact metric space and $\phi \in \mathscr{C}(X)$. Then*

$$sup_{\mu\in\mathcal{M}_T(X)} \left(h_\mu(T) + \int \phi d\mu \right) = P_T(\phi). \tag{55}$$

Let's rephrase the variational principle of chapter 1 in our new notation. We showed in section 2 that $s(\mu) = h(\sigma,\mathcal{U}) = h_\mu(\sigma)$ for the partition $\mathcal{U} = \{U_1, \dots, U_n\}$ where $U_i = \{\underline{x} : x = i\}$ of Σ_A. Then for any σ-invariant measure μ on Σ_A with $\phi \in \mathscr{F}_A$ one has

$$h_\mu(\sigma) + \int \phi d\mu \leq P_T(\phi,\mathcal{U}) \tag{56}$$

with equality if and only if $\mu = \mu_\phi$.

By comparing both variational principles, and taking the notations of the second chapter one concludes that $P_T(\phi,\mathcal{U}) = P_T(\phi)$, and so we showed a posteriori that the definition of pressure from the first chapter is compatible with the one from the second chapter.

References

[1] Haïm Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.

[2] Frank Bonsall, *Lectures on some Fixed Point Theorems Of Functional Analysis*, Tata Institute of Fundamental Research, Bombay, 1962.

[3] Rufus Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lectures Notes in Mathematics 470, Springer, April 14^{th}, 2013 (Original lecture notes 1975).

[4] James Dugundji, *Topology*, Allyn and Bacon, 1966.

[5] Jürgen Elstrodt, *Mass und Integrationstheorie*, 6. korrigierte Auflage, Springer-Verlag.

[6] Milan Merkle, *Topics in weak convergence of probability measures*, Zb. radova Mat. Inst. Beograd, 9(17) (2000), 235-274.

[7] James Munkres, *General Topology, Second Edition*, Prentice-Hall International.

[8] Louis Nirenberg, *Topics in nonlinear functional analysis*, Courant lecture notes in mathematics.

[9] Nelson Dunford and Jacob T. Schwartz, *Linear Operators, Part I: General Theory*, Willey Classics Library.

[10] Kerley Oliveira and Marcelo Viana, *Fundamentos da Teoria Ergódica*, Lecture notes at IMPA. Also available in English as *Foundations of Ergodic theory*, Cambridge Studies in Advanced Mathematics, Volume 151.

Index