Lecture Notes in Mathematics

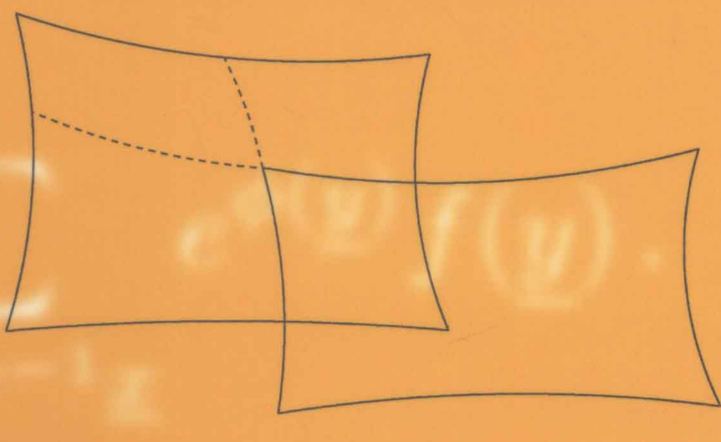
470

Rufus Bowen

Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms

2nd Revised Edition

Edited by Jean-René Chazottes



Springer

Rufus Bowen

Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms

Second revised edition

Jean-René Chazottes
Editor

Preface by David Ruelle

Springer

Author
Robert Edward (Rufus) Bowen
(1947-1978)
University of California at Berkeley
USA

Editor
Jean-René Chazottes
Centre de Physique Théorique
CNRS-École Polytechnique
91128 Palaiseau Cedex
France
jeanrene@cpht.polytechnique.fr

Preface by
David Ruelle
IHÉS
35, Route de Chartres
91440 Bures-sur-Yvette
France
ruelle@ihes.fr

ISBN: 978-3-540-77605-5 e-ISBN: 978-3-540-77695-6
DOI: 10.1007/978-3-540-77695-6

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2008922889

Mathematics Subject Classification (2000): Primary: 37D20, Secondary: 37D35

© 2008, 1975 Springer-Verlag Berlin Heidelberg
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: WMXDesign GmbH

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The Greek and Roman gods, supposedly, resented those mortals endowed with superlative gifts and happiness, and punished them. The life and achievements of Rufus Bowen (1947–1978) remind us of this belief of the ancients. When Rufus died unexpectedly, at age thirty-one, from brain hemorrhage, he was a very happy and successful man. He had great charm, that he did not misuse, and superlative mathematical talent. His mathematical legacy is important, and will not be forgotten, but one wonders what he would have achieved if he had lived longer. Bowen chose to be simple rather than brilliant. This was the hard choice, especially in a messy subject like smooth dynamics in which he worked. Simplicity had also been the style of Steve Smale, from whom Bowen learned dynamical systems theory.

Rufus Bowen has left us a masterpiece of mathematical exposition: the slim volume *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms* (Springer Lecture Notes in Mathematics **470** (1975)). Here a number of results which were new at the time are presented in such a clear and lucid style that Bowen's monograph immediately became a classic. More than thirty years later, many new results have been proved in this area, but the volume is as useful as ever because it remains the best introduction to the basics of the ergodic theory of hyperbolic systems.

The area discussed by Bowen came into existence through the merging of two apparently unrelated theories. One theory was equilibrium statistical mechanics, and specifically the theory of states of infinite systems (Gibbs states, equilibrium states, and their relations as discussed by R.L. Dobrushin, O.E. Lanford, and D. Ruelle). The other theory was that of hyperbolic smooth dynamical systems, with the major contributions of D.V. Anosov and S. Smale. The two theories came into contact when Ya.G. Sinai introduced Markov partitions and symbolic dynamics for Anosov diffeomorphisms. This allowed the poweful techniques and results of statistical mechanics to be applied to smooth dynamics, an extraordinary development in which Rufus Bowen played a major role. Some of Bowen's ideas were as follows. First, only one-dimensional statistical mechanics is discussed: this is a richer theory, which yields what is needed for applications to dynamical systems, and makes use of the powerful analytic tool of transfer operators. Second, Smale's Axiom A dynamical systems are studied rather than the less general Anosov systems. Third, Sinai's Markov partitions are reworked to apply to Axiom A systems and their construction is simplified by the use of *shadowing*. The combination of simplifications and generalizations just outlined led to Bowen's concise and lucid monograph. This text has not aged since it was written and its beauty is as striking as when it was first published in 1975.

Jean-René Chazottes has had the idea to make Bowen's monograph more easily available by retyping it. He has scrupulously respected the original text and notation, but corrected a number of typos and made a few other minor corrections, in particular in the bibliography, to improve usefulness and readability. In his enterprise he has been helped by Jerôme Buzzi, Pierre Collet, and Gerhard Keller. For this work of love all of them deserve our warmest thanks.

Bures sur Yvette, mai 2007 *David Ruelle*

此为试读,需要完整PDF请访问: www.ertongbook.com

Contents

These notes came out of a course given at the University of Minnesota and were revised while the author was on a Sloan Fellowship.

0

Introduction

The main purpose of these notes is to present the ergodic theory of Anosov and Axiom A diffeomorphisms. These diffeomorphisms have a complicated orbit structure that is perhaps best understood by relating them topologically and measure theoretically to shift spaces. This idea of studying the same example from different viewpoints is of course how the subjects of topological dynamics and ergodic theory arose from mechanics. Here these subjects return to help us understand differentiable systems.

These notes are divided into four chapters. First we study the statistical properties of Gibbs measures. These measures on shift spaces arise in modern statistical mechanics; they interest us because they solve the problem of determining an invariant measure when you know it approximately in a certain sense. The Gibbs measures also satisfy a variational principle. This principle is important because it makes no reference to the shift character of the underlying space. Through this one is led to develop a "thermodynamic formalism" on compact spaces; this is carried out in chapter two. In the third chapter Axiom A diffeomorphisms are introduced and their symbolic dynamics constructed: this states how they are related to shift spaces. In the final chapter this symbolic dynamics is applied to the ergodic theory of Axiom A diffeomorphisms.

The material of these notes is taken from the work of many people. I have attempted to give the main references at the end of each chapter, but no doubt some are missing. On the whole these notes owe most to D. Ruelle and Ya. Sinai.

To start, recall that $(X, \mathscr{B}, \mu)$ is a *probability space* if $\mathscr{B}$ is a σ-field of subsets of X and μ is a nonnegative measure on $\mathscr{B}$ with $\mu(X) = 1$. By an *automorphism* we mean a bijection $T : X \to X$ for which

$$\text{(i)} \quad E \in \mathscr{B} \ \text{ iff } \ T^{-1}E \in \mathscr{B},$$
$$\text{(ii)} \quad \mu(T^{-1}E) = \mu(E) \ \text{ for } E \in \mathscr{B}.$$

If $T : X \to X$ is a homeomorphism of a compact metric space, a natural σ-field $\mathscr{B}$ is the family of Borel sets. A probability measure on this σ-field is called a *Borel* probability measure. Let $\mathscr{M}(X)$ be the set of Borel probability measures on X and $\mathscr{M}_T(X)$ the subset of invariant ones, i.e. $\mu \in \mathscr{M}_T(X)$ if $\mu(T^{-1}E) = \mu(E)$ for all Borel sets E. For any $\mu \in \mathscr{M}(X)$ one can define $T^*\mu \in \mathscr{M}(X)$ by $T^*\mu(E) = \mu(T^{-1}E)$.

Remember that the real-valued continuous functions $\mathscr{C}(X)$ on the compact metric space X form a Banach space under $\|f\| = \max_{x\in X} |f(x)|$. The weak $*$-topology on the space $\mathscr{C}(X)^*$ of continuous linear functionals $\alpha : \mathscr{C}(X) \to \mathbb{R}$ is generated by sets of the form

$$U(f, \varepsilon, \alpha_0) = \{\alpha \in \mathscr{C}(X)^* : |\alpha(f) - \alpha_0(f)| < \varepsilon\}$$

with $f \in \mathscr{C}(X)$, $\varepsilon > 0$, $\alpha_0 \in \mathscr{C}(X)^*$.

Riesz Representation. For each $\mu \in \mathscr{M}(X)$ define $\alpha_\mu \in \mathscr{C}(X)^*$ by $\alpha_\mu(f) = \int f d\mu$. Then $\mu \leftrightarrow \alpha_\mu$ is a bijection between $\mathscr{M}(X)$ and

$$\{\alpha \in \mathscr{C}(X)^* : \alpha(1) = 1 \text{ and } \alpha(f) \geq 0 \text{ whenever } f \geq 0\}.$$

We identify α_μ with μ, often writing μ when we mean $\alpha(\mu)$. The weak $*$-topology on $\mathscr{C}(X)^*$ carries over by this identification to a topology on $\mathscr{M}(X)$ (called the weak topology).

Proposition. *$\mathscr{M}(X)$ is a compact convex metrizable space.*

Proof. Let $\{f_n\}_{n=1}^\infty$ be a dense subset of $\mathscr{C}(X)$. The reader may check that the weak topology on $\mathscr{M}(X)$ is equivalent to the one defined by the metric

$$d(\mu, \mu') = \sum_{n=1}^{\infty} 2^{-n} \|f_n\|^{-1} \left| \int f_n d\mu - \int f_n d\mu' \right| . \qquad \square$$

Proposition. *$\mathscr{M}_T(X)$ is a nonempty closed subset of $\mathscr{M}(X)$.*

Proof. Check that $T^* : \mathscr{M}(X) \to \mathscr{M}(X)$ is a homeomorphism and note that $\mathscr{M}_T(X) = \{\mu \in \mathscr{M}(X) : T^*\mu = \mu\}$. Pick $\mu \in \mathscr{M}(X)$ and let $\mu_n = \frac{1}{n}(\mu + T^*\mu + \cdots + (T^*)^{n-1}\mu)$. Choose a subsequence μ_{n_k} converging to $\mu' \in \mathscr{M}(X)$. Then $\mu' \in \mathscr{M}_T(X)$. $\square$

Proposition. *$\mu \in \mathscr{M}_T(X)$ if and only if*

$$\int (f \circ T)\, d\mu = \int f d\mu \quad \text{for all } f \in \mathscr{C}(X).$$

Proof. This is just what the Riesz Representation Theorem says about the statement $T^*\mu = \mu$. $\square$

1

Gibbs Measures

A. Gibbs Distribution

Suppose a physical system has possible states $1, \ldots, n$ and the energies of these states are $E_1, \ldots, E_n$. Suppose that this system is put in contact with a much larger "heat source" which is at temperature T. Energy is thereby allowed to pass between the original system and the heat source, and the temperature T of the heat source remains constant as it is so much larger than our system. As the energy of our system is not fixed any of the states could occur. It is a physical fact derived in statistical mechanics that the probability p_j that state j occurs is given by the *Gibbs distribution*

$$p_j = \frac{e^{-\beta E_j}}{\sum_{i=1}^{n} e^{-\beta E_i}},$$

where $\beta = \frac{1}{kT}$ and k is a physical constant.

We shall not attempt the physical justification for the Gibbs distribution, but we will state a mathematical fact closely connected to the physical reasoning.

1.1. Lemma. *Let real numbers $a_1, \ldots, a_n$ be given. Then the quantity*

$$F(p_1, \ldots, p_n) = \sum_{i=1}^{n} -p_i \log p_i + \sum_{i=1}^{n} p_i a_i$$

has maximum value $\log \sum_{i=1}^{n} e^{a_i}$ as $(p_1, \ldots, p_n)$ ranges over the simplex $\{(p_1, \ldots, p_n) : p_i \geq 0,\ p_1 + \cdots + p_n = 1\}$ and that maximum is assumed only by

$$p_j = e^{a_j} \left(\sum_i e^{a_i} \right)^{-1}.$$

This is proved by calculus. The quantity $H(p_1, \ldots, p_n) = \sum_{i=1}^n -p_i \log p_i$ is called the *entropy* of the distribution $(p_1, \ldots, p_n)$ (*note*: $\varphi(x) = -x \log x$ is continuous on $[0,1]$ if we set $\varphi(0) = 0$.) The term $\sum_{i=1}^n p_i a_i$ is of course the average value of the function $a(i) = a_i$. In the statistical mechanics case $a_i = -\beta E_i$, entropy is denoted S and average energy E. The Gibbs distribution then maximizes

$$S - \beta E = S - \frac{1}{kT} E,$$

or equivalently minimizes $E - kTS$. This is called the free energy. The principle that "nature minimizes entropy" applies when energy is fixed, but when energy is not fixed "nature minimizes free energy." We will now look at a simple *infinite* system, the one-dimensional lattice. Here one has for each integer a physical system with possible states $1, 2, \ldots, n$. A configuration of the system consists of assigning an $x_i \in \{1, \ldots, n\}$ for each i:

$$\cdot \quad \cdot \quad \underset{x_{-2}}{\cdot} \quad \underset{x_{-1}}{\cdot} \quad \underset{x_0}{\cdot} \quad \underset{x_1}{\cdot} \quad \underset{x_2}{\cdot} \quad \underset{x_3}{\cdot} \quad \cdot \quad \cdot$$

Thus a configuration is a point

$$\underline{x} = \{x_i\}_{i=-\infty}^{+\infty} \in \prod_{\mathbb{Z}} \{1, \ldots, n\} = \Sigma_n .$$

We now make assumptions about energy:

(1) associated with the occurrence of a state k is a contribution $\Phi_0(k)$ to the total energy of the system independent of which position it occurs at;
(2) if state k_1 occurs in place i_1, and k_2 in i_2, then the potential energy due to their interaction $\Phi_2^*(i_1, i_2, k_1, k_2)$ depends only on their relative position, *i.e.*, there is a function $\Phi_2 : \mathbb{Z} \times \{1, \ldots, n\} \times \{1, \ldots, n\} \to \mathbb{R}$ so that

$$\Phi_2^*(i_1, i_2, k_1, k_2) = \Phi_2(i_1 - i_2; k_1, k_2)$$

(also: $\Phi_2(j; k_1, k_2) = \Phi_2(-j; k_2, k_1)$).
(3) all energy is due to contributions of the form (1) and (2).

Under these hypotheses the energy contribution due to x_0 being in the 0th place is

$$\phi^*(\underline{x}) = \Phi_0(x_0) + \sum_{j \neq 0} \frac{1}{2} \Phi_2(j; x_j, x_0).$$

(We "give" each of x_0 and x_k half the energy due to their interaction). We now assume that $\|\Phi_2\|_j = \sup_{k_1, k_2} |\Phi(j; k_1, k_2)|$ satisfies

$$\sum_{j=1}^{\infty} \|\Phi_2\|_j < \infty .$$

此为试读，需要完整PDF请访问：www.ertongbook.com

Then $\phi^*(\underline{x}) \in \mathbb{R}$ and depends continuously on $\underline{x}$ when $\{1, \ldots, n\}$ is given the discrete topology and $\Sigma_n = \prod_{\mathbb{Z}}\{1, \ldots, n\}$ the product topology.

If we just look at $x_{-m} \ldots x_0 \ldots x_m$ we have a finite system (n^{2m+1} possible configurations) and an energy

$$E_m(x_{-m}, \ldots, x_m) = \sum_{j=-m}^{m} \Phi_0(x_j) \quad + \sum_{-m \leq j < k \leq m} \Phi_2(k-j; x_k, x_j)$$

and the Gibbs distribution μ_m assigns probabilities proportional to $e^{-\beta E_m(x_{-m}, \ldots, x_m)}$. Now just suppose that for each $x_{-m}, \ldots, x_m$ the limit

$$\mu(x_{-m} \ldots x_m) = \lim_{k \to \infty} \sum \{\mu_k(x'_{-k} \ldots x'_k) : x'_i = x_i \ \forall |i| \leq m\}$$

exists. Then $\mu \in \mathscr{M}(\Sigma_n)$ and it would be natural to call μ the Gibbs distribution on Σ_n (for the given energy and β). If we are given $\underline{x} = \{x_i\}_{i=-\infty}^{\infty}$, then instead of $E_m(x_{-m}, \ldots, x_m)$ one might add in the contributions by interactions of x_j $(-m \leq j \leq m)$ with *all* other x_k's, *i.e.*,

$$\sum_{j=-m}^{m} \left(\Phi_0(x_j) + \sum_{k=-\infty}^{\infty} \frac{1}{2} \Phi_2(k-j; x_k, x_j) \right) .$$

If we define the (left) *shift* homeomorphism $\sigma : \Sigma_n \to \Sigma_n$ by $\sigma\{x_i\}_{i=-\infty}^{\infty} = \{x_{i+1}\}_{i=-\infty}^{\infty}$, then this expression is just $\sum_{j=-m}^{m} \phi^*(\sigma^j \underline{x})$. This expression differs from $E_m(x_{-m}, \ldots, x_m)$ by at most

$$\sum_{j=-m}^{m} \left(\sum_{k=j+m+1}^{\infty} \frac{1}{2} \|\Phi_2\|_k \ + \sum_{k=m-j+1}^{\infty} \frac{1}{2} \|\Phi_2\|_k \right) \leq \sum_{k=1}^{\infty} k \|\Phi_2\|_k .$$

Thus, if $C = \sum_{k=1}^{\infty} k\|\Phi_2\|_k < \infty$ then $E_m(x_{-m}, \ldots, x_m)$ differs from $\sum_{j=-m}^{m} \phi^*(\sigma^j \underline{x})$ by at most C. If we used $\sum_{j=-m}^{m} \phi^*(\sigma^j \underline{x})$ instead of $E_m(x_{-m}, \ldots, x_m)$ in the Gibbs distribution μ_m, the probabilities would change by factors in $[e^{-2C}, e^{2C}]$. The point is that taking x_i into consideration for $i \notin [-m, m]$ may change μ_m, but not drastically if one assumes $\sum_{k=1}^{\infty} k\|\Phi_2\|_k < \infty$.

We want now to state a theorem we have been leading up to. For $\phi : \Sigma_n \to \mathbb{R}$ continuous define

$$\operatorname{var}_k \phi = \sup\{|\phi(\underline{x}) - \phi(\underline{y})| : x_i = y_i \ \forall \ |i| \leq k\} .$$

As ϕ is uniformly continuous, $\lim_{k \to \infty} \operatorname{var}_k \phi = 0$.

1.2. Theorem. *Suppose $\phi : \Sigma_n \to \mathbb{R}$ and there are $c > 0$, $\alpha \in (0, 1)$ so that $\operatorname{var}_k \phi \leq c\alpha^k$ for all k. Then there is a unique $\mu \in \mathscr{M}_\sigma(\Sigma_n)$ for which one can find constants $c_1 > 0$, $c_2 > 0$, and P such that*

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

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

This measure μ is written μ_ϕ and called *Gibbs measure* of ϕ. Up to constants in $[c_1, c_2]$ the relative probabilities of the $x_0 \ldots x_m$'s are given by $\exp \sum_{k=0}^{m-1} \phi(\sigma^k \underline{x})$. For the physical system discussed above one takes $\phi = -\beta\phi^*$. In statistical mechanics Gibbs states are not *defined* by the above theorem. We have ignored many subtleties that come up in more complicated systems (*e.g.*, higher dimensional lattices), where the theorem will not hold. Our discussion was a gross one intended to motivate the theorem; we refer to Ruelle [9] or Lanford [6] for a refined outlook.

For later use we want to make a small generalization of Σ_n before we prove the theorem. If A is an $n \times n$ matrix of 0's and 1's, let

$$\Sigma_A = \{\underline{x} \in \Sigma_n : A_{x_i x_{i+1}} = 1 \;\; \forall i \in \mathbb{Z}\}.$$

That is, we consider all $\underline{x}$ in which A says that $x_i x_{i+1}$ is allowable for every i. One easily sees that Σ_A is closed and $\sigma\Sigma_A = \Sigma_A$. *We will always assume that* A *is such that each* k *between* 1 *and* n *occurs at* x_0 *for some* $\underline{x} \in \Sigma_A$. (Otherwise one could have $\Sigma_A = \Sigma_B$ with B an $m \times m$ matrix and $m < n$.)

1.3. Lemma. $\sigma : \Sigma_A \to \Sigma_A$ *is topologically mixing (*i.e., *when* U, V *are non-empty open subsets of* Σ_A, *there is an* N *so that* $\sigma^m U \cap V \neq \emptyset \;\forall m \geq N$*) if and only if* $A^M > 0$ (i.e., $A^M_{i,j} > 0 \;\forall i, j$) *for some* M.

Proof. One sees inductively that $A^m_{i,j}$ is the number of $(m+1)$-strings $a_0 a_1 \ldots a_m$ of integers between 1 and n with

(a) $A_{a_k a_{k+1}} = 1 \quad \forall k$,
(b) $a_0 = i$, $a_m = j$.

Let $U_i = \{\underline{x} \in \Sigma_A : x_0 = i\} \neq \emptyset$.

Suppose Σ_A is mixing. Then $\exists N_{i,j}$ with $U_i \cap \sigma^n U_j \neq \emptyset \;\forall n \geq N_{i,j}$. If $\underline{a} \in U_i \cap \sigma^n U_j$, then $a_0 a_1 \ldots a_n$ satisfies (a) and (b); so $A^m_{i,j} > 0 \;\forall i, j$ when $m \geq \max_{i,j} N_{i,j}$.

Suppose $A^M > 0$ for some M. As each number between 1 and n occurs as x_0 for some $\underline{x} \in \Sigma_A$, each row of A has at least one positive entry. From this it follows by induction that $A^m > 0$ for all $m \geq M$.

Consider open subsets U, V of Σ_A with $\underline{a} \in U$, $\underline{b} \in V$. There is an r so that

$$U \supset \{\underline{x} \in \Sigma_A : x_k = a_k \;\; \forall |k| \leq r\}$$
$$V \supset \{\underline{x} \in \Sigma_A : x_k = b_k \;\; \forall |k| \leq r\}.$$

For $t \geq 2r + M$, $m = t - 2r \geq M$ and $A^m > 0$. Hence find $c_0, \ldots, c_m$ with $c_0 = b_r$, $c_m = a_{-r}$, $A_{c_k c_{k+1}} = 1$ for all k. Then

$$\underline{x} = \cdots b_{-2} b_{-1} b_0 \cdots b_r c_1 \cdots c_{m-1} a_{-r} \cdots a_0 a_1 \cdots$$

is in Σ_A and $\underline{x} \in \sigma^t U \cap V$. So Σ_A is topologically mixing. □

Let $\mathscr{F}_A$ be the family of all continuous $\phi : \Sigma_A \to \mathbb{R}$ for which $\mathrm{var}_k \phi \le b\alpha^k$ (for all $k \ge 0$) for some positive constants b and $\alpha \in (0,1)$. For any $\beta \in (0,1)$ one can define the metric d_β on Σ_A by $d_\beta(\underline{x}, \underline{y}) = \beta^N$ where N is the largest nonnegative integer with $x_i = y_i$ for every $|i| < N$. Then $\mathscr{F}_A$ is just the set of functions which have a positive Hölder exponent with respect to d_β. The theorem we are interested in then reads

1.4. Existence of Gibbs measures. *Suppose Σ_A is topologically mixing and $\phi \in \mathscr{F}_A$. There is 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*

$$c_1 \le \frac{\mu\{\underline{y} : y_i = x_i \text{ for all } i \in [0,m)\}}{\exp\left(-Pm + \sum_{k=0}^{m-1} \phi(\sigma^k \underline{x})\right)} \le c_2$$

for every $\underline{x} \in \Sigma_A$ and $m \ge 0$.

This theorem will not be proved for some time. The first step is to reduce the ϕ's one must consider.

Definition. *Two functions ψ, $\phi \in \mathscr{C}(\Sigma_A)$ are* homologous *with respect to σ (written $\psi \sim \phi$) if there is a $u \in \mathscr{C}(\Sigma_A)$ so that*

$$\psi(\underline{x}) = \phi(\underline{x}) - u(\underline{x}) + u(\sigma\underline{x}).$$

1.5. Lemma. *Suppose $\phi_1 \sim \phi_2$ and Theorem 1.4 holds for ϕ_1. Then it holds for ϕ_2 and $\mu_{\phi_1} = \mu_{\phi_2}$.*

Proof.

$$\begin{aligned}\left|\sum_{k=0}^{m-1} \phi_1(\sigma^k \underline{x}) - \sum_{k=0}^{m-1} \phi_2(\sigma^k \underline{x})\right| &= \left|\sum_{k=0}^{m-1} u(\sigma^{k+1}\underline{x}) - u(\sigma^k \underline{x})\right| \\ &= |u(\sigma^m \underline{x}) - u(\underline{x})| \le 2\|u\|.\end{aligned}$$

The exponential in the required inequality changes by at most a factor of $e^{2\|u\|}$ when ϕ_1 is replaced by ϕ_2. Thus the inequality remains valid with c_1, c_2 changed and P, μ unchanged. □

1.6. 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 $i \ge 0$.*

Proof. For each $1 \le t \le n$ pick $\{a_{k,t}\}_{k=-\infty}^{\infty} \in \Sigma_A$ with $a_{0,t} = t$. Define $r : \Sigma_A \to \Sigma_A$ by $r(\underline{x}) = \underline{x}^*$ where

$$x_k^* = \begin{cases} x_k & \text{for } k \ge 0 \\ a_{k,x_0} & \text{for } k \le 0. \end{cases}$$

Let

$$u(\underline{x}) = \sum_{j=0}^{\infty}(\phi(\sigma^j \underline{x}) - \phi(\sigma^j r(\underline{x}))) .$$

Since $\sigma^j \underline{x}$ and $\sigma^j r(\underline{x})$ agree in places from $-j$ to $+\infty$,

$$|\phi(\sigma^j \underline{x}) - \phi(\sigma^j r(\underline{x}))| \le \mathrm{var}_j \phi \le b\alpha^j .$$

As $\sum_{j=0}^{\infty} b\alpha^j < \infty$, u is defined and continuous. If $x_i = y_i$ for all $|i| \le n$, then, for $j \in [0, n]$,

$$|\phi(\sigma^j \underline{x}) - \phi(\sigma^j \underline{y})| \le \mathrm{var}_{n-j}\phi \le b\alpha^{n-j}$$

and

$$|\phi(\sigma^j r(\underline{x})) - \phi(\sigma^j r(\underline{y}))| \le b\alpha^{n-j} .$$

Hence

$$|u(\underline{x}) - u(\underline{y})| \le \sum_{j=0}^{[\frac{n}{2}]} |\phi(\sigma^j \underline{x}) - \phi(\sigma^j \underline{y}) + \phi(\sigma^j r(\underline{x})) - \phi(\sigma^j r(\underline{y}))| + 2 \sum_{j>[\frac{n}{2}]} \alpha^j$$

$$\le 2b \left(\sum_{j=0}^{[\frac{n}{2}]} \alpha^{n-j} + \sum_{j>[\frac{n}{2}]} \alpha^j \right) \le \frac{4b\, \alpha^{[\frac{n}{2}]}}{1-\alpha} .$$

This shows that $u \in \mathscr{F}_A$. Hence $\psi = \phi - u + u \circ \sigma$ is in $\mathscr{F}_A$ also. Furthermore

$$\psi(\underline{x}) = \phi(\underline{x}) + \sum_{j=-1}^{\infty} (\phi(\sigma^{j+1} r(\underline{x})) - \phi(\sigma^{j+1} \underline{x})) + \sum_{j=0}^{\infty} (\phi(\sigma^{j+1} \underline{x}) - \phi(\sigma^j r(\sigma \underline{x})))$$

$$= \phi(r(\underline{x})) + \sum_{j=0}^{\infty} (\phi(\sigma^{j+1} r(\underline{x})) - \phi(\sigma^j r(\sigma \underline{x}))) .$$

The final expression depends only on $\{x_i\}_{i=0}^{\infty}$, as we wanted. D. Lind cleaned up the above proof for us. □

Lemmas 1.5 and 1.6 tell us that in looking for a Gibbs measure μ_ϕ for $\phi \in \mathscr{F}_A$ (*i.e.*, proving Theorem 1.4) we can restrict our attention to functions ϕ for which $\phi(\underline{x})$ depends only on $\{x_i\}_{i=0}^{\infty}$.

B. Ruelle's Perron-Frobenius Theorem

We introduce now one-sided shift spaces. One writes $\underline{x}$ for $\{x_i\}_{i=0}^{\infty}$ (we will continue to write $\underline{x}$ for $\{x_i\}_{i=-\infty}^{\infty}$ but never for both things at the same time). Let

$$\Sigma_A^+ = \left\{ \underline{x} \in \prod_{i=0}^{\infty} \{1, \ldots, n\} \ : \ A_{x_i, x_{i+1}} = 1 \text{ for all } i \ge 0 \right\}$$

and define $\sigma : \Sigma_A^+ \to \Sigma_A^+$ by $\sigma(\underline{x})_i = x_{i+1}$. σ is a finite-to-one continuous map of Σ_A^+ onto itself. If $\phi \in \mathscr{C}(\Sigma_A^+)$ we get $\phi \in \mathscr{C}(\Sigma_A)$ by $\phi(\{x_i\}_{i=-\infty}^{\infty}) = \phi(\{x_i\}_{i=0}^{\infty})$. Suppose $\phi \in \mathscr{C}(\Sigma_A)$ satisfies $\phi(\underline{x}) = \phi(\underline{y})$ whenever $x_i = y_i$ for all $i \geq 0$. Then one can think of ϕ as belonging to $\mathscr{C}(\Sigma_A^+)$ as follows: $\phi(\{x_i\}_{i=0}^{\infty}) = \phi(\{x_i\}_{i=-\infty}^{\infty})$ where x_i for $i \leq 0$ are chosen in any way subject to $\{x_i\}_{i=-\infty}^{\infty} \in \Sigma_A$. The functions $\mathscr{C}(\Sigma_A^+)$ are thus identified with a certain subclass of $\mathscr{C}(\Sigma_A)$. We saw in Lemmas 1.5 and 1.6 that one only needs to get Gibbs measures for $\phi \in \mathscr{C}(\Sigma_A^+) \cap \mathscr{F}_A$ in order to get them for all $\phi \in \mathscr{F}_A$.

In this section we will prove a theorem of Ruelle that will later be used to construct and study Gibbs measures. For $\phi \in \mathscr{C}(\Sigma_A^+)$ define the operator $\mathcal{L} = \mathcal{L}_\phi$ on $\mathscr{C}(\Sigma_A^+)$ by

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

It is the fact that σ is not one-to-one on Σ_A^+ that will make this operator useful.

1.7. Ruelle's Perron-Frobenius Theorem [10, 11]. *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 $\nu \in \mathscr{M}(\Sigma_A^+)$ for which $\mathcal{L}h = \lambda h$, $\mathcal{L}^*\nu = \lambda\nu$, $\nu(h) = 1$ and*

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

Proof. Because $\mathcal{L}$ is a positive operator and $\mathcal{L}1 > 0$, one has that $G(\mu) = (\mathcal{L}^*\mu(1))^{-1}\mathcal{L}^*\mu \in \mathscr{M}(\Sigma_A^+)$ for $\mu \in \mathscr{M}(\Sigma_A^+)$. There is a $\nu \in \mathscr{M}(\Sigma_A^+)$ with $G(\nu) = \nu$ by the Schauder-Tychonoff Theorem (see Dunford and Schwartz, Linear Operators I, p. 456): Let E be a nonempty compact convex subset of a locally convex topological vector space. Then any continuous $G : E \to E$ has a fixed point. In our case $G(\nu) = \nu$ gives $\mathcal{L}^*\nu = \lambda\nu$ with $\lambda > 0$.

We will prove 1.7 via a sequence of lemmas. Let $b > 0$ and $\alpha \in (0,1)$ be any constants so that $\mathrm{var}_k\phi \leq b\alpha^k$ for all $k \geq 0$. Set $B_m = \exp\left(\sum_{k=m+1}^{\infty} 2b\alpha^k\right)$ and define

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

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

Proof. One checks that $\lambda^{-1}\mathcal{L}f \in \Lambda$ when $f \in \Lambda$. Clearly $\lambda^{-1}\mathcal{L}f \geq 0$ and

$$\nu(\lambda^{-1}\mathcal{L}f) = \lambda^{-1}\mathcal{L}^*\nu(f) = \nu(f) = 1 .$$

Assume $x_i = x_i'$ for $i \in [0,m]$. Then

$$\mathcal{L}f(\underline{x}) = \sum_j e^{\phi(j\underline{x})} f(j\underline{x})$$

where the sum ranges over all j with $A_{jx_0} = 1$. For $\underline{x}'$ the expression runs over the same j; as $j\underline{x}$ and $j\underline{x}'$ agree in places 0 to $m+1$

$$e^{\phi(j\underline{x})} f(j\underline{x}) \leq e^{\phi(j\underline{x}')} e^{b\alpha^{m+1}} B_{m+1} f(j\underline{x}') \leq B_m e^{\phi(j\underline{x}')} f(j\underline{x}')$$

and so

$$\mathcal{L}f(\underline{x}) \leq B_m \mathcal{L}f(\underline{x}')\,.$$

Consider any $\underline{x}, \underline{z} \in \Sigma_A^+$. Since $A^M > 0$ there is a $\underline{y}' \in \sigma^{-M}\underline{x}$ with $y_0' = z_0$. For $f \in \Lambda$

$$\begin{aligned}\mathcal{L}^M f(\underline{x}) &= \sum_{\underline{y}\in\sigma^{-M}\underline{x}} \exp\left(\sum_{k=0}^{M-1} \phi(\sigma^k \underline{y}) f(\underline{y})\right) \\ &\geq e^{-M\|\phi\|} f(\underline{y}')\,.\end{aligned}$$

Let $K = \lambda^M e^{M\|\phi\|} B_0$. Then $1 = \nu(\lambda^{-M}\mathcal{L}^M f) \geq K^{-1} f(\underline{z})$ gives $\|f\| \leq K$ as $\underline{z}$ is arbitrary. As $\nu(f) = 1$, $f(\underline{z}) \geq 1$ for some $\underline{z}$ and we get $\inf \lambda^{-M}\mathcal{L}^M f \geq K^{-1}$.

If $x_i = x_i'$ for $i \in [0, m]$ and $f \in \Lambda$, one has

$$|f(\underline{x}) - f(\underline{x}')| \leq (B_m - 1)K \to 0$$

as $m \to \infty$, since $B_m \to 1$. Thus Λ is equicontinuous and compact by the Arzela-Ascoli Theorem. $\Lambda \neq \emptyset$ as $1 \in \Lambda$. Applying Schauder-Tychonoff Theorem to $\lambda^{-1}\mathcal{L} : \Lambda \to \Lambda$ gives us $h \in \Lambda$ with $\mathcal{L}h = \lambda h$. Furthermore $\inf h = \inf \lambda^{-M}\mathcal{L}^M h \geq K^{-1}$. □

1.9. Lemma. *There is an $\eta \in (0,1)$ so that for $f \in \Lambda$ one has $\lambda^{-M}\mathcal{L}^M f = \eta h + (1-\eta) f'$ with $f' \in \Lambda$.*

Proof. Let $g = \lambda^{-M}\mathcal{L}^M f - \eta h$ where η is to be determined. Provided $\eta\|h\| \leq K^{-1}$ we will have $g \geq 0$. Assume $x_i = x_i'$ for all $i \in [0, m]$. We want to pick η so that $g(\underline{x}) \leq B_m g(\underline{x}')$, or equivalently

$$(\star) \quad \eta(B_m h(\underline{x}') - h(\underline{x})) \leq B_m \lambda^{-M}\mathcal{L}^M f(\underline{x}') - \lambda^{-M}\mathcal{L}^M f(\underline{x})\,.$$

We saw above that $\mathcal{L}f_1(\underline{x}) \leq B_{m+1}\, e^{b\alpha^{m+1}} \mathcal{L}f_1(\underline{x}') \leq B_{m+1}\, e^{b\alpha^m} \mathcal{L}f_1(\underline{x}')$ for any $f_1 \in \Lambda$. Applying this to $f_1 = \lambda^{-M+1}\mathcal{L}^{M-1} f$ one has

$$\lambda^{-M}\mathcal{L}^M f(\underline{x}) \leq B_{m+1}\, e^{b\alpha^m} \lambda^{-M}\mathcal{L}^M f(\underline{x}')\,.$$

Now $h(\underline{x}) \geq B_m^{-1} h(\underline{x}')$ because $h \in \Lambda$. To get $(\star)$ it is therefore enough to have

$$\eta(B_m - B_m^{-1}) h(\underline{x}') \leq (B_m - B_{m+1} e^{b\alpha^m})\, \lambda^{-M}\mathcal{L}^M f(\underline{x}')$$

or

$$\eta(B_m - B_m^{-1})\|h\| \leq (B_m - B_{m+1} e^{b\alpha^m}) K^{-1}\,.$$

此为试读,需要完整PDF请访问: www.ertongbook.com