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*

2

General Thermodynamic Formalism

A. Entropy

In Section D of Chapter 1, we defined the number $h_\mu(T, \mathcal{D})$ when T is an endomorphism of a probability space and $\mathcal{D}$ a finite measurable partition. We now define the *entropy* of μ w.r.t. T by

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

where $\mathcal{D}$ ranges over all finite partitions. We will now turn to some computational lemmas.

We define

$$\begin{aligned} H_\mu(\mathcal{C}|\mathcal{D}) &= H_\mu(\mathcal{C} \vee \mathcal{D}) - H_\mu(\mathcal{D}) \\ &= -\sum_i \mu(D_i) \sum_j \frac{\mu(C_j \cap D_i)}{\mu(D_i)} \log\left(\frac{\mu(C_j \cap D_i)}{\mu(D_i)}\right) \\ &\geq 0\,. \end{aligned}$$

Lemma 1.17 says that $H_\mu(\mathcal{C}|\mathcal{D}) \leq H_\mu(\mathcal{C})$. We write $\mathcal{C} \subset \mathcal{D}$ if each set in $\mathcal{C}$ is a union of sets in $\mathcal{D}$.

2.1. Lemma.

(a) $H_\mu(\mathcal{C}|\mathcal{D}) \leq H_\mu(\mathcal{C}|\mathcal{E})$ *if* $\mathcal{D} \supset \mathcal{E}$.
(b) $H_\mu(\mathcal{C}|\mathcal{D}) = 0$ *if* $\mathcal{D} \supset \mathcal{C}$.
(c) $H_\mu(\mathcal{C} \vee \mathcal{D}|\mathcal{E}) \leq H_\mu(\mathcal{C}|\mathcal{E}) + H_\mu(\mathcal{D}|\mathcal{E})$.
(d) $H_\mu(\mathcal{C}) \leq H_\mu(\mathcal{D}) + H_\mu(\mathcal{C}|\mathcal{D})$.

Proof. Letting $\varphi(x) = -x \log x$, $H_\mu(\mathcal{C}|\mathcal{D}) = \sum_j \sum_i \mu(D_i)\ \varphi\left(\frac{\mu(C_j \cap D_i)}{\mu(D_i)}\right)$. Since $\mathcal{E} \subset \mathcal{D}$, one can rewrite this as

$$H_\mu(\mathcal{C}|\mathcal{D}) = \sum_j \sum_{E \in \mathcal{E}} \mu(E) \sum_{D_i \subset E} \frac{\mu(D_i)}{\mu(E)}\ \varphi\left(\frac{\mu(C_j \cap D_i)}{\mu(D_i)}\right).$$

By the concavity of φ (see the proof of Lemma 1.17) one has $\varphi(\sum a_i x_i) \geq \sum a_i \varphi(x_i)$ where

$$a_i = \frac{\mu(D_i)}{\mu(E)}\ ,\ x_i = \frac{\mu(C_j \cap D_i)}{\mu(D_i)}\,.$$

Hence

$$H_\mu(\mathcal{C}|\mathcal{D}) \leq \sum_j \sum_{E\in\mathcal{E}} \mu(E)\ \varphi\left(\frac{\mu(C_j \cap E)}{\mu(E)}\right) = H_\mu(\mathcal{C}|\mathcal{E})\,.$$

To see (b) one notes that $\mathcal{C} \vee \mathcal{D} = \mathcal{D}$ when $\mathcal{D} \supset \mathcal{C}$. For (c) one writes

$$\begin{aligned} H_\mu(\mathcal{C} \vee \mathcal{D}|\mathcal{E}) &= H_\mu(\mathcal{C} \vee \mathcal{D} \vee \mathcal{E}) - H_\mu(\mathcal{D} \vee \mathcal{E}) + H_\mu(\mathcal{D} \vee \mathcal{E}) - H_\mu(\mathcal{E}) \\ &= H_\mu(\mathcal{C}|\mathcal{D} \vee \mathcal{E}) + H_\mu(\mathcal{D}|\mathcal{E}) \\ &\leq H_\mu(\mathcal{C}|\mathcal{E}) + H_\mu(\mathcal{D}|\mathcal{E}) \end{aligned}$$

by (a). Finally

$$\begin{aligned} H_\mu(\mathcal{C}) &= H_\mu(\mathcal{C} \vee \mathcal{D}) - H_\mu(\mathcal{D}|\mathcal{C}) \\ &\leq H_\mu(\mathcal{C} \vee \mathcal{D}) = H_\mu(\mathcal{D}) + H_\mu(\mathcal{C}|\mathcal{D})\,. \quad \square \end{aligned}$$

2.2. Lemma. *Let T be an endomorphism of a probability space $(X, \mathscr{B}, \mu)$, $\mathcal{C}$ and $\mathcal{D}$ finite partitions. Then*

(a) $H_\mu(T^{-k}\mathcal{C}|T^{-k}\mathcal{D}) = H_\mu(\mathcal{C}|\mathcal{D})$ *for* $k \geq 0$,
(b) $h_\mu(T, \mathcal{C}) \leq h_\mu(T, \mathcal{D}) + H_\mu(\mathcal{C}|\mathcal{D})$,
(c) $h_\mu(T, \mathcal{C} \vee \cdots \vee T^{-n}\mathcal{C}) = h_\mu(T, \mathcal{C})$.

Proof. As μ is T-invariant,

$$\begin{aligned} H_\mu(T^{-k}\mathcal{C}|T^{-k}\mathcal{D}) &= H_\mu(T^{-k}\mathcal{C} \vee T^{-k}\mathcal{D}) - H_\mu(T^{-k}\mathcal{D}) \\ &= H_\mu(\mathcal{C} \vee \mathcal{D}) - H_\mu(\mathcal{D}) = H_\mu(\mathcal{C}|\mathcal{D})\,. \end{aligned}$$

Using Lemma 2.1

$$\begin{aligned} H_\mu(\mathcal{C} \vee \cdots \vee T^{-m+1}\mathcal{C}) &\leq H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) \\ &\quad + H_\mu(\mathcal{C} \vee \cdots \vee T^{-m+1}\mathcal{C}|\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) \\ &\leq H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) \\ &\quad + \sum_{k=0}^{m-1} H_\mu(T^{-k}\mathcal{C}|\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) \\ &\leq H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) + \sum_{k=0}^{m-1} H_\mu(T^{-k}\mathcal{C}|T^{-k}\mathcal{D}) \\ &= H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) + mH_\mu(\mathcal{C}|\mathcal{D})\,. \end{aligned}$$

Dividing by m and letting $m \to \infty$,

$$h_\mu(T, \mathcal{C}) \le h_\mu(T, \mathcal{D}) + H_\mu(\mathcal{C}|\mathcal{D}) .$$

Set $\mathcal{D} = \mathcal{C} \vee \cdots \vee T^{-n}\mathcal{C}$. Then

$$\frac{1}{m} H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) = \frac{1}{m} H_\mu(\mathcal{C} \vee \cdots \vee T^{-m-n+1}\mathcal{C}) .$$

Letting $m \to \infty$, (as $\frac{m}{m+n} \to 1$) we get

$$h_\mu(T, \mathcal{D}) = h_\mu(T, \mathcal{C}) . \quad \square$$

2.3. Lemma. *Let X be a compact metric space, $\mu \in \mathscr{M}(X)$, $\varepsilon > 0$ and $\mathcal{C}$ a finite Borel partition. There is a $\delta > 0$ so that $H_\mu(\mathcal{C}|\mathcal{D}) < \varepsilon$ whenever $\mathcal{D}$ is a partition with* $\operatorname{diam}(\mathcal{D}) < \delta$.

Proof. Let $\mathcal{C} = \{C_1, \ldots, C_n\}$. In Lemma 1.23 we showed that, for any $\alpha > 0$, one could find $\delta > 0$ such that whenever $\mathcal{D}$ satisfies $\operatorname{diam}(\mathcal{D}) < \delta$ there is a $\mathcal{E} = \{E_1, \ldots, E_n\} \subset \mathcal{D}$ with

$$\mu(E_i \Delta C_i) < \alpha .$$

The expression

$$H_\mu(\mathcal{C}|\mathcal{E}) = \sum_{i,j} \mu(E_j)\ \varphi\left(\frac{\mu(C_j \cap E_i)}{\mu(E_i)}\right)$$

depends continuously upon the numbers

$$\mu(C_j \cap E_i) \quad \text{and} \quad \mu(E_i) = \sum_j \mu(C_j \cap E_i)$$

and vanishes when $\mu(C_j \cap E_i) = \delta_{ij}\ \mu(E_i)$. Hence, for α small, $H_\mu(\mathcal{C}|\mathcal{E}) < \varepsilon$. Then $H_\mu(\mathcal{C}|\mathcal{D}) \le H_\mu(\mathcal{C}|\mathcal{E}) < \varepsilon$ by 2.1 (a). $\square$

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

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

Proof. Of course $h_\mu(T) \ge \limsup_n h_\mu(T, \mathcal{D}_n)$. Consider any partition $\mathcal{C}$. By Lemmas 2.2 (b) and 2.3

$$h_\mu(T, \mathcal{C}) \le \liminf_n h_\mu(T, \mathcal{D}_n) .$$

Varying $\mathcal{C}$, $h_\mu(T) \le \liminf_n h_\mu(T, \mathcal{D}_n)$. $\square$

A homeomorphism $T : X \to X$ is called *expansive* if there exists $\varepsilon > 0$ so that

$$d(T^k x, T^k y) \le \varepsilon \quad \text{for all}\, k \in \mathbb{Z} \;\; \Rightarrow \;\; x = y .$$

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

Proof. Let $\mathcal{D}_n = T^n\mathcal{D} \vee \cdots \vee \mathcal{D} \vee \cdots \vee T^{-n}\mathcal{D}$. Then $\operatorname{diam}(\mathcal{D}_n) \to 0$ using expansiveness. Hence $h_\mu(T) = \lim_n h_\mu(T, \mathcal{D}_n)$. But $h_\mu(T, \mathcal{D}_n) = h_\mu(T, \mathcal{D})$ by Lemma 2.2 (c). □

Consider the case of $\sigma : \Sigma_A \to \Sigma_A$ and standard partition $\mathcal{U} = \{U_1, \ldots, U_n\}$ where $U_i = \{\underline{x} \in \Sigma_A : x_0 = i\}$. Then σ is expansive and 2.5 gives that $h_\mu(\sigma) = h_\mu(\sigma, \mathcal{U})$ for $\mu \in \mathscr{M}_\sigma(\Sigma_A)$. Now $h_\mu(\sigma, \mathcal{U})$ is what we denoted by $s(\mu)$ in Chapter 1. That $s(\mu) = h_\mu(\sigma)$ implies that the number $s(\mu)$ does not depend on the homeomorphism σ and partition $\mathcal{U}$, but only on σ as an automorphism of the probability space $(\Sigma_A, \mathscr{B}, \mu)$ (because of the definition of $h_\mu(\sigma)$).

2.6. Lemma. $h_\mu(T^n) = nh_\mu(T)$ *for* $n > 0$.

Proof. Let $\mathcal{C}$ be a partition and $\mathcal{E} = \mathcal{C} \vee \cdots \vee T^{-n+1}\mathcal{C}$. Then

$$\begin{aligned} nh_\mu(T, \mathcal{C}) &= \lim_{m\to\infty} \frac{n}{nm} H_\mu(\mathcal{C} \vee \cdots \vee T^{-nm+1}\mathcal{C}) \\ &= \lim_{m\to\infty} \frac{1}{m} H_\mu(\mathcal{E} \vee T^{-n}\mathcal{E} \vee \cdots \vee T^{(-m+1)n}\mathcal{E}) \\ &= h_\mu(T^n, \mathcal{E}) \le h_\mu(T^n) = nh_\mu(T)\,. \end{aligned}$$

Varying $\mathcal{C}$, $nh_\mu(T) \le h_\mu(T^n)$. On the other hand

$$h_\mu(T^n, \mathcal{C}) \le h_\mu(T^n, \mathcal{E})$$

by 2.2 (b) and 2.1 (b). Hence

$$h_\mu(T^n) = \sup_{\mathcal{C}} h_\mu(T^n, \mathcal{C}) \le n \sup_{\mathcal{C}} h_\mu(T, \mathcal{C}) = nh_\mu(T)\,. \quad \square$$

B. Pressure

Throughout this section $T : X \to X$ will be a fixed continuous map on the compact metric space X. We will define the pressure $P(\phi)$ of $\phi \in \mathscr{C}(X)$ in a way which generalizes Section D in Chapter 1.

Let $\mathcal{U}$ be a finite open cover of X, $W_m(\mathcal{U})$ the set of all m-strings

$$\underline{\mathrm{U}} = U_{i_0} U_{i_1} \cdots U_{i_{m-1}}$$

of members of $\mathcal{U}$. One writes $m = m(\underline{\mathrm{U}})$,

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

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

In case $X(\underline{\mathrm{U}}) = \emptyset$, we let $S_m\phi(\underline{\mathrm{U}}) = -\infty$. We say that $\Gamma \subset W_m(\mathcal{U})$ *covers* X if $X = \bigcup_{\underline{\mathrm{U}}\in\Gamma} X(\underline{\mathrm{U}})$. Finally one defines

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

where Γ runs over all subsets of $W_m(\mathcal{U})$ covering X.

2.7. Lemma. *The limit*

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

exists and is finite.

Proof. If $\Gamma_m \subset W_m(\mathcal{U})$ and $\Gamma_n \subset W_n(\mathcal{U})$ each cover X, then

$$\Gamma_m\Gamma_n = \{\underline{\mathrm{UV}} : \underline{\mathrm{U}} \in \Gamma_m,\ \underline{\mathrm{V}} \in \Gamma_n\} \subset W_{m+n}(\mathcal{U})$$

covers X. One sees that

$$S_{m+n}\phi(\underline{\mathrm{UV}}) \le S_m\phi(\underline{\mathrm{U}}) + S_n\phi(\underline{\mathrm{V}})$$

and so

$$\sum_{\underline{\mathrm{UV}}\in\Gamma_m\Gamma_n}\exp(S_{m+n}\phi(\underline{\mathrm{UV}})) \le \sum_{\underline{\mathrm{U}}\in\Gamma_m}\exp(S_m\phi(\underline{\mathrm{U}}))\sum_{\underline{\mathrm{V}}\in\Gamma_n}\exp(S_n\phi(\underline{\mathrm{V}})).$$

Thus

$$Z_{m+n}(\phi,\mathcal{U}) \le Z_m(\phi,\mathcal{U})\ Z_n(\phi,\mathcal{U})$$

and $Z_m(\phi,\mathcal{U}) \ge e^{-m\|\phi\|}$. Hence $a_m = \log Z_m(\phi,\mathcal{U})$ satisfies the hypotheses of Lemma 1.18. □

2.8. Proposition. *The limit*

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

exists (but may be $+\infty$*).*

Proof. Suppose $\mathcal{V}$ is an open cover refining $\mathcal{U}$, *i.e.*, every $V \in \mathcal{V}$ lies in some $U(V) \in \mathcal{U}$. For $\underline{\mathrm{V}} \in W_m(\mathcal{V})$ let $U(\underline{\mathrm{V}}) = U(V_{i_0})\cdots U(V_{i_{m-1}})$. If $\Gamma_m \subset W_m(\mathcal{V})$ covers X, then $U(\Gamma_m) = \{U(\underline{\mathrm{V}}) : \underline{\mathrm{V}} \in \Gamma_m\} \subset W_m(\mathcal{U})$ covers X.
Let $\gamma = \gamma(\phi,\mathcal{U}) = \sup\{|\phi(x) - \phi(y)| : x, y \in U \text{ for some } U \in \mathcal{U}\}$.
Then $S_m\phi(U(\underline{\mathrm{V}})) \le S_m\phi(\underline{\mathrm{V}}) + m\gamma$ and so $Z_m(\phi,\mathcal{U}) \le e^{m\gamma}Z_m(\phi,\mathcal{V})$, which gives

$$P(\phi, \mathcal{U}) \leq P(\phi, \mathcal{V}) + \gamma \,.$$

Now for any $\mathcal{U}$, all $\mathcal{V}$ with small diameter refine $\mathcal{U}$ and so

$$P(\phi, \mathcal{U}) - \gamma(\phi, \mathcal{U}) \leq \liminf_{\mathrm{diam}(\mathcal{V}) \to 0} P(\phi, \mathcal{V}) \,.$$

Letting $\mathrm{diam}(\mathcal{U}) \to 0$, $\gamma(\phi, \mathcal{U}) \to 0$ and

$$\limsup_{\mathrm{diam}(\mathcal{U}) \to 0} P(\phi, \mathcal{U}) \leq \liminf_{\mathrm{diam}(\mathcal{V}) \to 0} P(\phi, \mathcal{V}) \,.$$

We are done. □

In cases where confusion may arise we write the topological pressure $P(\phi)$ as $P_T(\phi)$.

2.9. Lemma. *Let $S_n\phi(x) = \sum_{k=0}^{n-1} \phi(T^k x)$. Then*

$$P_{T^n}(S_n \phi) = n P_T(\phi) \ \text{for} \ n > 0 \,.$$

Proof. Let $\mathcal{V} = \mathcal{U} \vee \cdots \vee T^{-n+1}\mathcal{U}$. Then $W_m(\mathcal{V})$ and $W_{mn}(\mathcal{U})$ are in one-to-one correspondence; for $\underline{\mathrm{U}} = U_{i_0} U_{i_1} \cdots U_{i_{mn-1}}$ let $\underline{\mathrm{V}} = V_{i_0} \cdots V_{i_{m-1}}$ where $V_{i_k} = U_{i_{kn}} \cap T^{-1} U_{i_{kn+1}} \cap \cdots \cap T^{-n+1} U_{i_{kn+n-1}}$. One sees that $X(\underline{\mathrm{U}}) = X(\underline{\mathrm{V}})$ and $S^T_{mn}\phi(\underline{\mathrm{U}}) = S^{T^n}_m (S_n\phi)(\underline{\mathrm{V}})$. Thus one gets

$$Z^T_{mn}(\phi, \mathcal{U}) = Z^{T^n}_m (S_n \phi, \mathcal{V}) \quad \text{and} \quad n P_T(\phi, \mathcal{U}) = P_{T^n}(S_n \phi, \mathcal{V}) \,.$$

As $\mathrm{diam}(\mathcal{U}) \to 0$, $\mathrm{diam}(\mathcal{V}) \to 0$ and so $nP_T(\phi) = P_{T^n}(S_n\phi)$. □

We now come to our first interesting result about the pressure $P(\phi)$.

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

$$h_\mu(T) + \int \phi \, d\mu \leq P_T(\phi),$$

for any $\mu \in \mathscr{M}_T(X)$.

We will first need a couple of lemmas.

2.11. Lemma. *Suppose $\mathcal{D}$ is a Borel partition of X such that each $x \in X$ is in the closures of at most M members of $\mathcal{D}$. Then*

$$h_\mu(T, \mathcal{D}) + \int \phi \, d\mu \leq P_T(\phi) + \log M \,.$$

Proof. Let $\mathcal{U}$ be a finite open cover of X each member of which intersects at most M members of $\mathcal{D}$. Let $\Gamma_m \subset W_m(\mathcal{U})$ cover X. For each $B \in \mathcal{D}_m = \mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}$ pick $x_B \in B$ with $\int_B S_m\phi \, d\mu \le \mu(B)\ S_m\phi(x_B)$. Now

$$\begin{aligned} h_\mu(T,\mathcal{D}) + \int \phi \, d\mu &\le \frac{1}{m}\left(H_\mu(\mathcal{D}_m) + \int S_m\phi \, d\mu\right) \\ &\le \frac{1}{m}\sum_B \mu(B)(-\log\mu(B) + S_m\phi(x_B)) \\ &\le \frac{1}{m}\log\sum_B \exp(S_m\phi(x_B)) \end{aligned}$$

by Lemma 1.1. For each x_B pick $\underline{U}_B \in \Gamma_m$ with $x_B \in X(\underline{U}_B)$. This map $B \to \underline{U}_B$ is at most M^m to one. As $S_m\phi(x_B) \le S_m\phi(\underline{U}_B)$, one has

$$\begin{aligned} h_\mu(T,\mathcal{D}) + \int \phi \, d\mu &\le \frac{1}{m}\log\sum_{\underline{U}\in\Gamma_m} M^m\ \exp(S_m\phi(\underline{U})) \\ &\le \log M + \frac{1}{m}\log Z_m(\phi,\mathcal{U})\,. \end{aligned}$$

Letting $m \to \infty$ and then $\operatorname{diam}(\mathcal{U}) \to 0$, we obtain the desired inequality. □

2.12. Lemma. *Let $\mathcal{A}$ be a finite open cover of X. For each $n > 0$ there is a Borel partition $\mathcal{D}_n$ of X so that*

(a) *D lies inside some member of $T^{-k}\mathcal{A}$ for each $D \in \mathcal{D}_n$ and $k = 0,\ldots,n-1$,*
(b) *at most $n|\mathcal{A}|$ sets in $\mathcal{D}_n$ can have a point in all their closures.*

Proof. Let $\mathcal{A} = \{A_1,\ldots,A_m\}$ and $g_1,\ldots,g_m$ be a partition of unity subordinate to A. Then $G = (g_1,\ldots,g_m) : X \to s_{m-1} \subset \mathbb{R}^m$ where s_{m-1} is an $m-1$ dimensional simplex. Now $\mathcal{U} = \{U_1,\ldots,U_m\}$ is an open cover where $U_i = \{\underline{x} \in s_{m-1} : x_i > 0\}$ and $G^{-1}U_i \subset A_i$. Since $(s_{m-1})^n$ is $nm-n$ dimensional, there is a Borel partition $\mathcal{D}_n^*$ of $(s_{m-1})^n$ so that

(a') each member of $\mathcal{D}_n^*$ lies in some $U_{i_1} \times \cdots \times U_{i_n}$, and
(b') at most nm members of $\mathcal{D}_n^*$ can have a common point in all their closures.

Then $\mathcal{D}_n = L^{-1}\mathcal{D}_n^*$ works where

$$L = (G, G\circ T, \ldots, G\circ T^{n-1}) : X \to (s_{m-1})^n\,. \quad □$$

Proof of 2.10. Let $\mathcal{C}$ be a Borel partition and $\varepsilon > 0$. By Lemma 2.3 find an open cover $\mathcal{A}$ so that $H_\mu(\mathcal{C}|\mathcal{D}) < \varepsilon$ whenever $\mathcal{D}$ is a partition every member of which is contained in some member of $\mathcal{A}$. Fix $n > 0$, let $\mathcal{E} = \mathcal{C} \vee \cdots \vee T^{-n+1}\mathcal{C}$ and $\mathcal{D}_n$ as in Lemma 2.12. Then (see the proof of 2.6)

$$\begin{aligned} h_\mu(T, \mathcal{C}) + \int \phi \, d\mu &\le \frac{1}{n}\left(h_\mu(T^n, \mathcal{E}) + \int S_n\phi \, d\mu\right) \\ &\le \frac{1}{n}\left(h_\mu(T^n, \mathcal{D}_n) + \int S_n\phi \, d\mu\right) + \frac{1}{n} H_\mu(\mathcal{E}|\mathcal{D}_n) \\ &\le \frac{1}{n}\left(P_{T^n}(S_n\phi) + \log(n|\mathcal{A}|)\right) + \frac{1}{n} H_\mu(\mathcal{E}|\mathcal{D}_n) \end{aligned}$$

by Lemma 2.11. Now

$$H_\mu(\mathcal{E}|\mathcal{D}_n) \le \sum_{k=0}^{n-1} H_\mu(T^{-k}\mathcal{C}|\mathcal{D}_n)\,.$$

Since $\mathcal{D}_n$ refines $T^{-k}\mathcal{A}$ for each k, one has $H_\mu(T^{-k}\mathcal{C}|\mathcal{D}_n) < \varepsilon$ (since μ is T-invariant, $T^{-k}\mathcal{A}$ bears the same relation to $T^{-k}\mathcal{C}$ as $\mathcal{A}$ to $\mathcal{C}$). Hence, using 2.9,

$$h_\mu(T, \mathcal{C}) + \int \phi \, d\mu \le P_T(\phi) + \frac{1}{n}\log(n|A|) + \varepsilon\,.$$

Now let $n \to \infty$ and then $\varepsilon \to 0$. □

2.13. Proposition. *Let $T_1 : X_1 \to X_1$, $T_2 : X_2 \to X_2$ be continuous maps on compact metric spaces, $\pi : X_1 \to X_2$ continuous and onto satisfying $\pi \circ T_1 = T_2 \circ \pi$. Then*

$$P_{T_2}(\phi) \le P_{T_1}(\phi \circ \pi)$$

for $\phi \in \mathscr{C}(X_2)$.

Proof. For $\mathcal{U}$ an open cover of X_2 one sees that

$$P_{T_2}(\phi, \mathcal{U}) = P_{T_1}(\phi \circ \pi, \pi^{-1}\mathcal{U})\,.$$

As in the proof of 2.8

$$P_{T_1}(\phi \circ \pi, \pi^{-1}\mathcal{U}) \le P_{T_1}(\phi \circ \pi) + \gamma(\phi \circ \pi, \pi^{-1}\mathcal{U})\,.$$

But $\gamma(\phi \circ \pi, \pi^{-1}\mathcal{U}) = \gamma(\phi, \mathcal{U}) \to 0$ as $\mathrm{diam}(\mathcal{U}) \to 0$. Hence, letting $\mathrm{diam}(\mathcal{U}) \to 0$ we get $P_{T_2}(\phi) \le P_{T_1}(\phi \circ \pi)$. □

C. Variational Principle

Let $\mathcal{U}$ be a finite open cover of X. We say that $\Gamma \subset W^*(\mathcal{U}) = \bigcup_{m>0} W_m(\mathcal{U})$ *covers* $K \subset X$ if $K \subset \bigcup_{\underline{\mathrm{U}} \in \Gamma} X(\underline{\mathrm{U}})$. For $\lambda > 0$ and $\Gamma \subset W^*(\mathcal{U})$ define

$$Z(\Gamma, \lambda) = \sum_{\underline{\mathrm{U}} \in \Gamma} \lambda^{m(\underline{\mathrm{U}})} \exp(S_{m(\underline{\mathrm{U}})}\phi(\underline{\mathrm{U}})).$$

2.14. Lemma. *Let $P = P(\phi, \mathfrak{U})$ and $\lambda > 0$. Suppose that $Z(\Gamma, \lambda) < 1$ for some Γ covering X. Then $\lambda \leq e^{-P}$.*

Proof. As X is compact we may take Γ finite and $\Gamma \subset \bigcup_{m=1}^{M} W_m(\mathfrak{U})$. Then $Z(\Gamma^n, \lambda) \leq Z(\Gamma, \lambda)^n$ where $\Gamma^n = \{\underline{\mathrm{U}}_1\underline{\mathrm{U}}_2 \cdots \underline{\mathrm{U}}_n : \underline{\mathrm{U}}_i \in \Gamma\}$. Letting $\Gamma^* = \bigcup_{n=1}^{\infty} \Gamma^n$, one has

$$Z(\Gamma^*, \lambda) = \sum_{n=1}^{\infty} Z(\Gamma^n, \lambda) < \infty\,.$$

Fix N and consider $x \in X$. Since Γ covers X, one can find $\underline{\mathrm{U}} = \underline{\mathrm{U}}_1\underline{\mathrm{U}}_2 \cdots \underline{\mathrm{U}}_n \in \Gamma^*$ with

(a) $x \in X(\underline{\mathrm{U}})$, and
(b) $N \leq m(\underline{\mathrm{U}}) < N + M$.

Let $\underline{\mathrm{U}}^*$ be the first N symbols of $\underline{\mathrm{U}}$. Then

$$S_N\phi(\underline{\mathrm{U}}^*) \leq S_{m(\underline{\mathrm{U}})}\phi(\underline{\mathrm{U}}) + M\|\phi\|\,.$$

For Γ^N the set of $\underline{\mathrm{U}}^*$ so obtained,

$$\lambda^N \sum_{\Gamma^N} \exp S_N\phi(\underline{\mathrm{U}}^*) \leq \max\left\{1, \lambda^{-M}\right\}\; e^{M\|\phi\|}\; Z(\Gamma^*, \lambda),$$

or $\lambda^N Z_N(\phi, \mathfrak{U}) \leq$ constant. It follows that $\lambda \leq e^{-P}$. □

Let δ_x be the unit-measure concentrated on the point x. Define

$$\begin{aligned} &\delta_{x,n} = n^{-1}(\delta_x + \delta_{Tx} + \cdots + \delta_{T^{n-1}x}) \\ \text{and}\qquad &V(x) = \{\mu \in \mathscr{M}(X) : \delta_{x,n_k} \to \mu \text{ for some } n_k \to \infty\}\,. \end{aligned}$$

$V(x) \neq \emptyset$ as $\mathscr{M}(X)$ is a compact metric space. Now $T^*\delta_{x,n} = \delta_{Tx,n}$ and for $f \in \mathscr{C}(X)$, $|T^*\delta_{x,n}(f) - \delta_{x,n}(f)| = n^{-1}|f(T^n x) - f(x)| \leq 2n^{-1}\|f\|$. This shows $V(x) \subset \mathscr{M}_T(X)$.

Let E be a finite set, $\underline{a} = (a_0, \ldots, a_{k-1}) \in E^k$. One defines the distribution $\mu_{\underline{a}}$ on E by

$$\begin{aligned} &\mu_{\underline{a}}(e) = k^{-1}(\text{number of } j \text{ with } a_j = e) \\ \text{and}\qquad &H(\underline{a}) = -\sum_{e \in E} \mu_{\underline{a}}(e) \log \mu_{\underline{a}}(e)\,. \end{aligned}$$

2.15. Lemma. *Let $x \in X$, $\mu \in V(x)$, $\mathfrak{U}$ a finite open cover of X and $\varepsilon > 0$. There are m and arbitrarily large N for which one can find $\underline{\mathrm{U}} \in W_N(\mathfrak{U})$ satisfying the following*

(a) $x \in X(\underline{\mathrm{U}})$,
(b) $S_N\phi(\underline{\mathrm{U}}) \leq N(\int \phi\, d\mu + \gamma(\mathfrak{U}) + \varepsilon)$,

(c) $\underline{\mathrm{U}}$ *contains a subword of length* $km \geq N - m$ *which, when viewed as* $\underline{a} = a_0, \ldots, a_{k-1} \in (\mathcal{U}^m)^k$ *satisfies*

$$\frac{1}{m} H(\underline{a}) \leq h_\mu(T) + \varepsilon.$$

Proof. Let $\mathcal{U} = \{U_1, \ldots, U_q\}$. Recall that

$$\gamma(\mathcal{U}) = \sup\{|\phi(y) - \phi(z)| : y, z \in U_i \text{ for some } i\}.$$

Pick a Borel partition $\mathcal{C} = \{C_1, \ldots, C_q\}$ with $\overline{C}_i \subset U_i$. There is an m so that

$$\frac{1}{m} H_\mu(\mathcal{C} \vee \cdots \vee T^{-m+1}\mathcal{C}) \leq h_\mu(T, \mathcal{C}) + \frac{\varepsilon}{2} \leq h_\mu(T) + \frac{\varepsilon}{2}\cdot$$

Let $\delta_{x,n_j} \to \mu$. For $n' > n$ one has

$$\delta_{x,n'} = \frac{n}{n'}\, \delta_{x,n} + \frac{n' - n}{n'}\, \delta_{T^n x, n'-n}\,.$$

If we replaced n_k by the nearest multiple of m, this formula shows that μ would still be the limit. Thus we assume $n_j = mk_j$.

Let $D_1, \ldots, D_t$ be the nonempty members of $\mathcal{C} \vee \cdots \vee T^{-m+1}\mathcal{C}$ and for each D_i find a compact $K_i \subset D_i$ with $\mu(D_i \backslash K_i) < \beta$ ($\beta > 0$ small). Each D_i is contained in some member of $\mathcal{U} \vee \cdots \vee T^{-m+1}\mathcal{U}$ and one can find an open set $V_i \supset K_i$ for which this is still true. Furthermore we may assume $V_i \cap V_j = \emptyset$ for $i \neq j$. Now enlarge each V_i to a Borel set V_i^* still contained in a member of $\mathcal{U} \vee \cdots \vee T^{-m+1}\mathcal{U}$ and so that $\{V_1^*, \ldots, V_t^*\}$ is a Borel partition of X.

Now fix $n_j = mk_j$. Let M_i be the number of $s \in [0, n_j)$ with $T^s x \in V_i^*$ and $M_{i,r}$ the number of such $s \equiv r \pmod m$.
Define

$$p_{i,r} = M_{i,r}/k_j$$

and $p_i = M_i/n_j = \frac{1}{m}(p_{i,0} + \cdots + p_{i,m-1})$. As $\delta_{x,n_j} \to \mu$, one has

$$\liminf_{j\to\infty} p_i \geq \mu(K_i) \geq \mu(D_i) - \beta,$$

and $\limsup_{j\to\infty} p_i \leq \mu(K_i) + t\beta \leq \mu(D_i) + t\beta$. For β small enough and j large enough one has

$$\begin{aligned}\frac{1}{m}\left(-\sum_i p_i \log p_i\right) &\leq \frac{1}{m}\left(-\sum_i \mu(D_i)\log\mu(D_i)\right) + \frac{\varepsilon}{2} \\ &\leq h_\mu(T) + \varepsilon\,.\end{aligned}$$

By the concavity of $\varphi(x) = -x\log x$ (see 1.17)

$$\varphi(p_i) \geq \sum_{r=0}^{m-1} \frac{1}{m}\, \varphi(p_{i,r})$$

and so

$$\sum_i \varphi(p_i) \geq \frac{1}{m} \sum_{r=0}^{m-1} \sum_i \varphi(p_{i,r}) .$$

For some $r \in [0, m)$ one must have $\sum_i \varphi(p_{i,r}) \leq \sum_i \varphi(p_i)$ and so

$$\frac{1}{m} \sum_i \varphi(p_{i,r}) \leq h_\mu(T) + \varepsilon .$$

For $N = n_j + r$ with j large we form $\underline{\mathrm{U}} = U_0 U_1 \cdots U_{N-1} \in \mathcal{U}^N$ as follows. For $s < r$ pick $U_s \in \mathcal{U}$ containing $T^s x$. For each V_i^* we choose $U_{0,i} \cap T^{-1} U_{1,i} \cap \cdots \cap T^{-m+1} U_{m-1,i} \supset V_i^*$. For $s > r$ we write $s = r + mp + q$ with $p \geq 0$, $m > q \geq 0$, pick i with $T^{r+mp} x \in V_i^*$ and let $U_s = U_{q,i}$. Letting

$$a_p = U_{0,i} U_{1,i} \cdots U_{m-1,i}$$

we have

$$\underline{\mathrm{U}} = U_0 \cdots U_{r-1} a_0 a_1 \cdots a_{k_j - 1} .$$

Now $\underline{a} = (a_0 a_1 \cdots a_{k_j-1})$ has its distribution $\mu_{\underline{a}}$ on $\mathcal{U}^m$ given by the probabilities $\{p_{i,r}\}_{i=1}^t$ and some zeros.

So

$$\frac{1}{m} H(\underline{a}) = \frac{1}{m} \sum_i \varphi(p_{i,r}) \leq h_\mu(T) + \varepsilon.$$

We have yet to check (b). Since $\delta_{x,n_j} \to \mu$, for j large we will have $\left|\frac{1}{N}\delta_{x,N}(\phi) - \int \phi \, d\mu\right| < \varepsilon$ or $S_N\phi(x) \leq N(\int \phi \, d\mu + \varepsilon)$. As $x \in X(\underline{\mathrm{U}})$, $S_N\phi(\underline{\mathrm{U}}) \leq S_N\phi(x) + N\gamma(\mathcal{U})$. $\square$

2.16. Lemma. *Fix a finite set E and $h \geq 0$. Let $R(k,h) = \{\underline{a} \in E^k : H(\underline{a}) \leq h\}$. Then*

$$\limsup_{k\to\infty} \frac{1}{k} \log |R(k,h)| \leq h .$$

Proof. For any distribution ν on E and $\alpha \in (0,1)$ consider

$$R_k(\nu) = \{\underline{a} \in E^k : |\mu_{\underline{a}}(e) - \nu(e)| < \alpha \ \ \forall e \in E\} .$$

Let μ be the Bernoulli measure on $\Sigma = \prod_{i=0}^\infty E$ with the distribution

$$\mu(e) = (1-\alpha)\nu(e) + \alpha/|E| .$$

Each $\underline{a} \in R_k(\nu)$ corresponds to a cylinder set $C_{\underline{a}}$ of Σ. Since each $e \in E$ occurs in $\underline{a}$ at most $k(\nu(e) + \alpha)$ times,

$$\mu(C_{\underline{a}}) \geq \prod_e \mu(e)^{k(\nu(e)+\alpha)} .$$

As the $C_{\underline{a}}$ are disjoint and have total measure 1,

$$1 \geq |R_k(\nu)| \prod_e \mu(e)^{k(\nu(e)+\alpha)},$$

$$\text{or} \quad \frac{1}{k} \log |R_k(\nu)| \leq \sum_e -(\nu(e)+\alpha) \log \mu(e)$$
$$\leq H(\mu) + \sum_e 3\alpha |\log \mu(e)| \,.$$

As $\mu(e) \geq \alpha/|E|$, we get

$$\frac{1}{k} \log |R_k(\nu)| \leq H(\mu) + 3\alpha |E| (\log |E| - \log \alpha) \,.$$

When $\alpha \to 0$, the second term on the right approaches 0 and $H(\mu) \to H(\nu)$ uniformly in ν. Hence, for any $\varepsilon > 0$ one can find α small enough that

$$\frac{1}{k} \log |R_k(\nu)| \leq H(\mu) + \varepsilon,$$

for all k and ν.

Once α is so chosen, let $\mathcal{N}$ be a finite set of distributions on E so that

(a) $H(\nu) \leq h$ for $\nu \in \mathcal{N}$, and
(b) if $H(\nu') \leq h$ then for some $\nu \in \mathcal{N}$ one has

$$|\nu'(e) - \nu(e)| < \alpha \quad \text{for all } \ e.$$

Then $R(k,h) \subset \bigcup_{\nu \in \mathcal{N}} R_k(\nu)$,

$$\frac{1}{k} \log |R(k,h)| \leq \frac{1}{k} \log |\mathcal{N}| + h + \varepsilon$$
$$\text{and} \qquad \limsup_{k \to \infty} \frac{1}{k} \log |R(k,h)| \leq h + \varepsilon.$$

Now let $\varepsilon \to 0$. □

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

$$P_T(\phi) = \sup_\mu \left(h_\mu(T) + \int \phi \; d\mu \right)$$

where μ runs over $\mathscr{M}_T(X)$.

Proof. Let $\mathfrak{U}$ be a finite cover of X and $\epsilon > 0$. For each $m > 0$ let X_m be the set of points $x \in X$ for which 2.15 holds with this m and some $\mu \in V(x)$. By 2.15 $X = \bigcup_m X_m$ since $V(x) \neq \emptyset$. For $u \in \mathbb{R}$ let $Y_m(u)$ be the set of $x \in X_m$ for which 2.15 holds for some $\mu \in V(x)$ with $\int \phi d\mu \in [u-\varepsilon, u+\varepsilon]$. Set

$$c = \sup_{\mu} \left(h_\mu(T) + \int \phi \, d\mu \right) .$$

For $x \in Y_m(u)$ the μ satisfies $h_\mu(T) \le c - u + \varepsilon$.

The $\underline{a} \in (\mathcal{U}^m)^k$ appearing in 2.15 (c) for $x \in Y_m(u)$ lie in $R(k, m(c-u+2\varepsilon), \mathcal{U}^m)$. The number of possibilities for $\underline{\mathrm{U}}$ for any fixed $N = km$ is hence at most

$$b(N) = |E|^m \; |R(k, m(c-u+2\varepsilon), \mathcal{U}^m)| .$$

By 2.16

$$\limsup_{N\to\infty} \frac{1}{N} \log b(N) \le c - u + 2\varepsilon .$$

Let $\Gamma = \Gamma_{m,u}$ be the collection of all $\underline{\mathrm{U}}$ showing up in the present situation for some N greater than a fixed N_0. Then Γ covers $Y_m(u)$ and

$$Z(\Gamma, \lambda) \le \sum_{N=N_0}^{\infty} \lambda^N \; b(N) \; \exp(N(u + 2\varepsilon + \gamma(\mathcal{U}))) .$$

For large enough N_0, $b(N) \le \exp(N(c - u + 3\varepsilon))$ and

$$\begin{aligned} Z(\Gamma, \lambda) &\le \sum_{N=N_0}^{\infty} \lambda^N \; \exp(N(c + 5\varepsilon + \gamma(\mathcal{U}))) . \\ &\le \sum_{N=N_0}^{\infty} \beta^N = \frac{\beta^{N_0}}{1-\beta} , \end{aligned}$$

where $\beta = \lambda \exp(c + 5\varepsilon + \gamma(\mathcal{U})) < 1$.

We have seen that for $\lambda < \exp(-(c + 5\varepsilon + \gamma(\mathcal{U})))$ any $Y_m(u)$ can be covered by $\Gamma \subset W^*(\mathcal{U})$ with $Z(\Gamma, \lambda)$ as small as desired. As $X = \bigcup_{m=1}^{\infty} X_m$ and $X_m = Y_m(u_1) \cup \cdots \cup Y_m(u_r)$ where $u_1, \ldots, u_r$ are ε-dense in $[-\|\phi\|, \|\phi\|]$, taking unions of such $\Gamma's$ we obtain a Γ covering X with $Z(\Gamma, \lambda) < 1$. By Lemma 2.14, $\lambda \le e^{-P(\phi, \mathcal{U})}$ or

$$P(\phi, \mathcal{U}) \le c + 5\varepsilon + \gamma(\mathcal{U}) .$$

As ε was arbitrary, $P(\phi, \mathcal{U}) \le c + \gamma(\mathcal{U})$.
Finally

$$\begin{aligned} P(\phi) &\le \lim_{\mathrm{diam}(\mathcal{U})\to 0} P(\phi, \mathcal{U}) \\ &\le \lim_{\mathrm{diam}(\mathcal{U})\to 0} (c + \gamma(\mathcal{U})) = c . \end{aligned}$$

The inequality $c \le P(\phi)$ follows from Theorem 2.10. □

2.18. Corollary. *Suppose $\{X_\alpha\}_{\alpha \in \Lambda}$ is a family of compact subsets of X and $TX_\alpha \subset X_\alpha$ for each α. Then*

$$P_T(\phi) = \sup_{\alpha} P_{T|_{X_\alpha}}(\phi|_{X_\alpha}) .$$

Proof. If $\mu \in \mathscr{M}_T(X_\alpha)$, then $\mu \in \mathscr{M}_T(X)$ and

$$P_T(\phi) \geq h_\mu(T) + \int \phi \, d\mu \, .$$

Hence

$$P_T(\phi) \geq \sup_{\mu \in \mathscr{M}_T(X_\alpha)} \left(h_\mu(T) + \int \phi \, d\mu \right) = P_{T|X_\alpha}(\phi|_{X_\alpha}) \, .$$

If $x \in X_\alpha$, then $V(x) \subset \mathscr{M}_T(X_\alpha)$ and so

$$\begin{aligned} c' &= \sup \left\{ h_\mu(T) + \int \phi \, d\mu \ : \ \mu \in \bigcup_{x \in X} V(x) \right\} \\ &\leq \sup_\alpha P_{T|X_\alpha}(\phi|_{X_\alpha}) \, . \end{aligned}$$

In the proof of 2.17 what was actually used about the number c was $c \geq h_\mu(T) + \int \phi \, d\mu$ for $\mu \in V(x)$. So c' would work just as well there to yield $P_T(\phi) \leq c'$. □

D. Equilibrium States

If $\mu \in \mathscr{M}_T(X)$ satisfies $h_\mu(T) + \int \phi \, d\mu = P_T(\phi)$, then μ is called an *equilibrium state* for ϕ (w.r.t. T). The Gibbs state μ_ϕ of $\phi \in \mathscr{F}_A$ in Chapter 1 was shown to be the unique equilibrium state for such a ϕ.

2.19. Proposition. *Suppose that for some $\varepsilon > 0$ one has $h_\mu(T, \mathcal{D}) = h_\mu(T)$ whenever $\mu \in \mathscr{M}_T(X)$ and* $\operatorname{diam}(\mathcal{D}) < \varepsilon$. *Then every $\phi \in \mathscr{C}(X)$ has an equilibrium state.*

Proof. We show that $\mu \mapsto h_\mu(T)$ is upper semi-continuous on $\mathscr{M}_T(X)$. Then $\mu \mapsto h_\mu(T) + \int \phi d\mu$ will be also, and the proposition follows from 2.17 and the fact that an u.s.c. function on a compact space assumes its supremum.

Fixing $\mu \in \mathscr{M}_T(X)$, $\alpha > 0$, and $\mathcal{D} = \{D_1, \ldots, D_n\}$ with $\operatorname{diam}(\mathcal{D}) < \varepsilon$, one has $\frac{1}{m} H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) \leq h_\mu(T) + \alpha$ for some m. Let $\beta > 0$ and pick a compact set $K_{i_0,\ldots,i_{m-1}} \subset \bigcap_{k=0}^{m-1} T^{-k} D_{i_k}$ with

$$\mu \left(\bigcap_k T^{-k} D_{i_k} \backslash K_{i_0,\ldots,i_{m-1}} \right) < \beta \, .$$

Then $D_i \supset L_i = \bigcup_{j=0}^{m-1} \bigcup \{T^j K_{i_0,\ldots,i_{m-1}} \, : \, i_j = i\}$. As the L_i are disjoint compact sets, one can find a partition $\mathcal{D}' = \{D'_1, \ldots, D'_n\}$ with $\operatorname{diam}(\mathcal{D}') < \varepsilon$ and $L_i \subset \operatorname{int}(D'_i)$. One then has

$$K_{i_0,\ldots,i_{m-1}} \subset \operatorname{int} \left(\bigcap_k T^{-k} D'_{i_k} \right) .$$

If ν is close to μ in the weak topology, one will have

$$\nu\left(\bigcap_k T^{-k} D'_{i_k}\right) \geq \mu(K_{i_0,\ldots,i_{m-1}}) - \beta$$

and $\left|\nu\left(\bigcap_k T^{-k} D'_{i_k}\right) - \mu\left(\bigcap_k T^{-k} D_{i_k}\right)\right| \leq 2\beta n^m$. For β small enough, this implies

$$\begin{aligned} h_\nu(T) = h_\nu(T, \mathcal{D}') &\leq \frac{1}{m} H_\nu(\mathcal{D}' \vee \cdots \vee T^{-m+1}\mathcal{D}') \\ &\leq \frac{1}{m} H_\mu(\mathcal{D} \vee \cdots \vee T^{-m+1}\mathcal{D}) + \alpha \leq h_\mu(T) + 2\alpha\,. \quad \square \end{aligned}$$

2.20. Corollary. *If T is expansive, every $\phi \in \mathscr{C}(X)$ has an equilibrium state.*

Proof. Recall 2.5. □

One notices that the condition in 2.19 has nothing to do with ϕ. Taking $\phi = 0$, one defines the *topological entropy* of T by

$$h(T) = P_T(0)\,.$$

The motivation for this chapter comes from two places: the theory of Gibbs states from statistical mechanics and topological entropy from topological dynamics (see references). Conditions on ϕ become important for the *uniqueness* of equilibrium state and then only after stringent conditions have been placed on the homeomorphism T. The Axiom A diffeomorphisms will be close enough to the subshifts $\sigma : \Sigma_A \to \Sigma_A$ so that one can prove uniqueness statements.

References

The definition of $h_\mu(T)$ is due to Kolmogorov and Sinai (see [2]). For expansive T Ruelle [15] defined $P_T(\phi)$ and proved Theorems 2.10, 2.17 and 2.20. For general T the definition and results are due to Walters [16].

In the transition from Σ_A to a general compact metric space X, most of the work has to do with the more complicated topology of X. Walters' paper is therefore closely related to earlier work on the topological entropy $h(T)$, *i.e.*, the case $\phi = 0$. The definition of $h(T)$ was made by Adler, Konheim and McAndrew [1]. The theorems for this case are due to Goodwyn [10] (Theorem 2.10), Dinaburg [6] (X finite dimensional, 2.17), Goodman [8] (general X, 2.17), and Goodman [9] (2.20). For subshifts these results were proved earlier by Parry [14]. The proofs we have given in these notes are adaptations of [4].

Gurevič [11] gives a T where $\phi = 0$ has no equilibrium states and Misiurewicz [13] gives such a T which is a diffeomorphism. The condition in 2.19 is satisfied by a class of maps which includes all affine maps on Lie groups [3] and Misiurewicz [13] showed that equilibrium states exist under a somewhat weaker condition.

Ruelle [15] showed that for expansive T a Baire set of ϕ have unique equilibrium states. Goodman [9] gives a minimal subshift where $\phi = 0$ has more than one equilibrium state. I believe mathematical physicists know of specific ϕ on Σ_n which do not have unique equilibrium states; in statistical mechanics one looks at $\mathbb{Z}^m$ actions instead of just homeomorphisms and gets nonuniqueness for $m \geq 2$ even with simple ϕ's. Uniqueness was proved in [5] for certain ϕ when T satisfies expansiveness and a very restrictive condition called specification; this result has been carried over to flows by Franco-Sanchez [7].
Finally we mention a very interesting result in a different direction. Let $T : M \to M$ be a continuous map on a compact manifold and λ an eigenvalue of the map $T_* : H_1(M) \to H_1(M)$ on one-dimensional homology. Then Manning [12] showed $h(T) \geq \log|\lambda|$. It is conceivable that this inequality is true for λ for any $H_k(M)$ (not just $k = 1$) provided T is a diffeomorphism.

1. R. L. Adler, A. G. Konheim and M. H. McAndrew: Topological entropy, *Trans. Amer. Math. Soc.* **114** (1965), 309-319.
2. P. Billingsley: *Ergodic Theory and Information*, Wiley (1965). ([1])
3. R. Bowen: Entropy-expansive maps, *Trans. Amer. Math. Soc.* **164** (1972), 323-331.
4. R. Bowen: Topological entropy for noncompact sets, *Trans. Amer. Math. Soc.* **184** (1973), 125-136.
5. R. Bowen: Some systems with unique equilibrium states, *Math. Systems Theory* **8** (1974/75), no. 3, 193-202.
6. E. I. Dinaburg: On the relations among various entropy characteristics of dynamical systems, *Math. USSR Izvestia* **5** (1971), 337-378.
7. E. Franco-Sanchez: *Flows with unique equilibrium states*, Ph.D. thesis. Univ. of California, Berkeley, 1974.
8. T. N. T. Goodman: Relating topological entropy and measure entropy, *Bull. London Math. Soc.* **3** (1971), 176-180.
9. T. N. T. Goodman: Maximal measures for expansive homeomorphisms, *J. London Math. Soc.* (2) **5** (1972), 439-444.
10. L. W. Goodwyn: Topological entropy bounds measure-theoretic entropy, *Proc. Amer. Math. Soc.* **23** (1969), 679-688.
11. B. M. Gurevič: Topological entropy of denumerable Markov chains, *Soviet Math. Dokl.* **10** (1969), 911-915.
12. A. Manning: *Topological entropy and the first homology group*, in "Dynamical Systems" - Warwick 1974 (Proc. Symp. Appl. Topology and Dynamical Systems, Univ. Warwick, Coventry, 1973/1974), pp. 185-190. Lect. Notes in Math. **468**, Springer, Berlin, 1975.
13. M. Misiurewicz: Diffeomorphism without any measure with maximal entropy, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **21** (1973), 903–910.
14. W. Parry: Intrinsic Markov chains, *Trans. Amer. Math. Soc.* **112** (1964), 55-66.
15. D. Ruelle: Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification, *Trans. Amer. Math. Soc.* **185** (1973), 237-251.
16. P. Walters: A variational principle for the pressure of continuous transformations, *Amer. J. Math.* **97** (1975), no. 4, 937-971.

[1] Reprint: Robert E. Krieger Publishing Co., Huntington, N.Y., 1978 (note of the editor).