11.11.1999

Chapters already renumbered, 2b=3 etc

CONFORMAL FRACTALS, DIMENSIONS AND ERGODIC THEORY

Feliks Przytycki & Mariusz Urbański

This book is an introduction to the theory of iteration of non-uniformly expanding holomorphic maps and topics in geometric measure theory of the underlying invariant fractal sets. Probability measures on these sets yield informations on Hausdorff and other fractal dimensions and properties. The book starts with a comprehensive chapter on abstract ergodic theory followed by chapters on uniform distance expanding maps and *thermodynamical formalism.* This material is applicable in many branches of dynamical systems and related fields, far beyond the applications in this book.

Popular examples of the fractal sets to be investigated are Julia sets for rational functions on the Riemann sphere. The theory which was initiated by Gaston Julia [J] and Pierre Fatou [F] became very popular since the time when Benoit Mandelbrot's book [M] with beautiful computer made pictures appeared. Then it became a field of spectacular achievements by top mathematicians during the last 20 years.

Consider for example the map $f(z) = z^2$ for complex numbers z. Then the unit circle $S^1 = \{|z| = 1\}$ is f-invariant, $f(S^1) = S^1 = f^{-1}(S^1)$. For $c \approx 0, c \neq 0$ and $f_c(z) = z^2 + c$, there still exists an f_c-invariant set $J(f_c)$ called the *Julia set* of f_c, close to S^1, homeomorphic to S^1 via a homeomorphism h satisfying equality $f \circ h = h \circ f_c$. However $J(f_c)$ has a fractal shape. For large c the curve $J(f_c)$ pinches at infinitely many points; it may pinch everywhere to become a dendrite, or even crumble to become a Cantor set.

These sets satisfy two main properties, standard attributes of "conformal fractal sets": 1. Their fractal dimensions are strictly larger than the topological dimension. 2. They are conformally "self-similar", namely arbitrarily small pieces have shapes similar to large pieces *via* conformal mappings, here *via* iteration of f.

To measure fractal sets invariant under holomorphic mappings one applies probability measures corresponding to equilibria in the thermodynamical formalism. This is a beautiful example of interlacing of ideas from mathematics and physics.

A *prototype lemma* [B, Lemma 1.1] at the roots of the thermodynamical formalism says that for given real numbers $a_1, ..., a_n$ the quantity

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

has maximum value $P = \log \sum_{i=1}^n e^{\phi_i}$ as $(p_1, ..., p_n)$ ranges over the simplex $\{(p_1, ..., p_n) : p_i \geq 0, \sum_{i=1}^n p_i = 1\}$ and the maximum is assumed only at

$$\hat{p}_j = e^{\phi_j} \big(\sum_{i=1}^{n} e^{\phi_i}\big)^{-1}$$

We can read $\phi_i, p_i, i = 1, ..., n$ as a function (*potential*), resp. probability distribution, on the finite space $\{1, ..., n\}$. Let us further follow Bowen [B]: The quantity

$$S = \sum_{i=1}^{n} -p_i \log p_i$$

is called *entropy* of the distribution $(p_1, ..., p_n)$. The maximizing distribution $(\hat{p}_1, .., \hat{p}_n)$ is called *Gibbs* or *equilibrium state*. In statistical mechanics $\phi_i = -\beta E_i$, where $\beta = 1/kT$, T is a temperature of an external "heat source" and k a physical (Boltzmann) constant. The quantity $E = \sum_{i=1}^{n} p_i E_i$ is the average energy. The Gibbs distribution maximizes then the expression

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

or equivalently minimizes the so-called *free energy* $E - kTS$. The nature prefers states with low energy and high entropy. It minimizes free energy.

The idea of Gibbs distribution as limit of distributions on finite spaces of configurations of states (spins for example) of interacting particles over increasing growing to ∞ bounded parts of the lattice $\boldsymbol{Z}^d$ introduced in statistical mechanics first by Bogolubov and Hacet [BH] and playing there a fundamental role was applied in dynamical systems to study Anosov flows and hyperbolic diffeomorphisms at the end of sixties by Ja. Sinai, D. Ruelle and R. Bowen. For more historical remarks see [Ru] or [Si]. This theory met the notion of entropy S borrowed from information theory and introduced by Kolmogorov as an invariant of a measure-theoretic dynamical system.

Later the usefulness of these notions to the geometric dimensions has become apparent. It was present already in [Billingsley] but crucial were papers by Bowen [Bo1] and McCluskey & Manning [McM].

In order to illustrate the idea consider the following example: Let $T_i : I \to I$, $i = 1, ..., n > 1$, where $I = [0, 1]$ is the unit interval, $T_i(x) = \lambda_i x + a_i$, where λ_i, a_i are real numbers chosen in such a way that all the sets $T_i(I)$ are pairwise disjoint and contained in I. Define the limit set Λ as follows

$$\Lambda = \bigcap_{k=0}^{\infty} \bigcup_{(i_0,...,i_k)} T_{i_0} \circ ... \circ T_{i_k}(I) = \bigcup_{(i_0,i_1...)} \lim_{k\to\infty} T_{i_0} \circ ... \circ T_{i_k},$$

the latter union taken over all infinite sequences $(i_0, i_1, ...)$, the previous over sequences of length $k + 1$.

It occurs that its Hausdorff dimension is equal to the only number α for which

$$|\lambda_1|^\alpha + ... + |\lambda_n|^\alpha = 1.$$

Λ is a Cantor set. It is self-similar with small pieces similar to large pieces with the use of linear (more precisely, affine) maps $(T_{i_0} \circ ... \circ T_{i_k})^{-1}$. We call such a Cantor set *linear*. We can distribute measure μ by setting $\mu(T_{i_0} \circ ... \circ T_{i_k}(I)) = \left(\lambda_{i_0} ... \lambda_{i_k}\right)^\alpha$. Then for each interval $J \subset I$ centered at a point of Λ its diameter raised to the power α is comparable to its measure μ (this is immediate for the intervals $T_{i_0} \circ ... \circ T_{i_k}(I)$). (A measure with this property for all small balls centered at a compact set, in a euclidean space of any dimension, is called a *geometric measure.*) Hence $\sum(\text{diam} J)^\alpha$ is bounded away from 0 and ∞ for all economical (of multiplicity not exceeding 2) covers of Λ by intervals J.

Note that for each k μ restricted to the space of unions of $T_{i_0} \circ ... \circ T_{i_k}(I)$, each such interval viewed as one point, is the Gibbs distribution, where we set $\phi((i_0, ..., i_k)) = \phi_\alpha((i_0, ..., i_k)) = \sum_{l=0,...,k} \alpha \log \lambda_{i_l}$. The number α is the unique 0 of the *pressure function* $\mathrm{P}(\alpha) = \frac{1}{k+1} \log \sum_{(i_0,...,i_k)} e^{\phi_a((i_0,...,i_k))}$. In this special affine example this is independent of k. In general non-linear case to define pressure one passes with k to ∞.

The family T_i and compositions is an example of very popular in recent years *Iterated Function System* [Barnsley]. Note that on a neighbourhood of each $T_i(I)$ we can consider $\hat{T} := T_i^{-1}$. Then Λ is an invariant repeller for the distance expanding map $\hat{T}$.)

The relations between dynamics, dimension and geometric measure theory start in our book with the theorem that the Hausdorff dimension of an expanding repeller is the unique 0 of the adequate pressure function for sets built with the help of $C^{1+\varepsilon}$ usually non-linear maps in $I\!\!R$ or conformal maps in $I\!\!R^d$.

This theory was developed for non-uniformly hyperbolic maps or flows in the setting of smooth ergodic theory, see [HK], by Mañé [M], Lai-Sang-Young and Ledrappier [LY]; see [Pesin] for recent developments. The advanced chapters of our book are devoted to this theory, but we restrict ourselves to complex dimension 1. So the maps are non-uniformly expanding and the main technical difficulties are caused by critical points, where we have strong contraction since the derivative by definition is equal to 0 at critical points.

A direction not developed in this book are Conformal Iterated Function Systems with infinitely many generators T_i. They occur naturally as return maps in many important constructions, for example for rational maps with parabolic periodic points or in the *Induced Expansion* construction for polynomials [GS]. Beautiful examples are provided by infinitely generated Kleinian groups [.]. The systematic treatment of Iterated Function Systems with infinitely many generators can be found in [MU1], [MU2], [MU3], [MPU] and [U1] for example.

Below is a short description of the content of the book.

Chapter 1 is an introduction to abstract ergodic theory, here T is a probability measure preserving transformation. The reader will find proofs of the fundamental theorems: Birkhoff Ergodic Theorem and Shannon-McMillan-Breiman Theorem. We introduce entropy, measurable partitions and discuss canonical systems of conditional measures in Rohlin's Lebesgue space the notion of *natural extension* (inverse limit in the appropriate category). We follow here Rohlin's Theory [Ro], see also [FKS]. Next to prepare to applications for finite-to-one rational maps we sketch Rohlin's theory on countable-to-one endomorphisms and introduce the notion of Jacobian, see also [Parry]. Finally we discuss mixing properties (K-propery, exactness, Bernoulli) and probability laws (Central Limit Theorem, abbr. CLT, Law of Iterated Logarithm, LIL, Almost Sure Invariance Principle, ASIP) for the sequence of functions (random variables on our probability space) $\phi \circ T^n, n = 0, 1, ...$

Chapter 2 is devoted to ergodic theory and termodynamical formalism for general continuous maps on compact metric spaces. The main point here is the so called Variational Principle for pressure, compare the *prototype lemma* above. We apply also functional analysis in order to explain Legendre transform duality between entropy and pressure. We follow here [Israel] and [Ruelle]. This material is applicable in *large deviations* and *multifractal analysis*, and is directly related to the uniqueness of Gibbs states question.

In Chapters 1, 2 we often follow the beautiful book by Peter Walters [Wa].

In Ch 3. *distance expanding maps* are introduced. Analogously to Axiom A diffeomorphisms [Smale, Bowen] or endomorphisms [Przy] we outline a topological theory: spectral decomposition, specification, Markov partition, and start a "bounded distortion" play with Hölder continuous functions.

In Chapter 4 termodynamical formalism and mixing properties of Gibbs measures for open distance expanding maps T and Hölder continuous potentials ϕ are studied. To large extend we follow [Bo] and [Ru]. We prove the existence of Gibbs measures (states): m with Jacobian being $\exp -\phi$ up to a constant factor, and T-invariant $\mu = \mu_\phi$ equivalent to m. The idea is to use the *transfer operator* $\mathcal{L}_\phi(u)(x) = \sum_{y\in T^{-1}(x)} u(y)\exp\phi(y)$ on the Banach space of Hölder continuous functions u. We prove the exponential convergence $\xi^{-n}\mathcal{L}_\phi^n(u) \to (\int u\, dm)u_\phi$, where ξ is the eigenvalue of the largest absolute value and u_ϕ the corresponding eigenfunction. One obtains $u_\phi = dm/d\mu$. We deduce CLT, LIL and ASIP, and the Bernoulli property for the natural extension.

We provide three different proofs of the uniqueness of the invariant Gibbs measure. The first, simplest, follows [Keller???], the second relies on the *prototype lemma*, the third one on the differentiability of the pressure function in adequate function directions.

Finally we prove Ruelle's formula

$$d^2P(\phi+tu+sv)/dt\, ds|_{t=s=0} = \lim_{n\to\infty}\frac{1}{n}\int(\sum_{i=0}^{n-1}(u\circ T^i-\int u\, d\mu_\phi)\cdot(\sum_{i=0}^{n-1}(v\circ T^i-\int v\, d\mu_\phi)\, d\mu_\phi.$$

This expression for $u=v$ is equal to σ^2 in CLT for the sequence $u\circ T^n$ and measure μ_ϕ.

(In the book we use the letter T to denote a measure preserving transformation. Maps preserving an additional structure, continuous smooth or holomorphic for example, are usually denoted f or g.)

In Chapter 5 the metric space with the action of an open distance expanding map is embedded in a smooth manifold and it is assumed that the map smoothly extends to a neighbourhood. We call the space with the extended dynamics: *Smooth Expanding Repeller*, abbr. SER. We study smoothness of the density u_ϕ. Finally we provide in detail D. Sullivan's theory classifying line Cantor sets via *scaling function*, sketched in [Su] and discuss the realization problem [PT]. We also discuss applications for *solenoids* for *Feigenbaum maps*.

In Chapter 6 we provide definitions of various "fractal dimensions": Hausdorff, box and packing. We consider also Hausdorff measures with gauge functions difefrent from t^α. We prove "Volume Lemma" linking, roughly speaking, (global) dimension with local dimensions.

In Chapter 7 we finally introduce Conformal Expanding Repellers, abbr. CER, and relate pressure with Hausdorff dimension. We prove $C^{?-1}$ dependence of the dimension on the parameter if the dependence on the parameter of the expanding map is $C^?$. We deal with smooth repellers in $I\!R$ and conformal repellers in $\mathbb{C}$. Here $2 <? \leq \omega$, the real analytic case.

Next we follow the easy (uniform) part of [PUZ]. We prove that for CER (X, f) and Hölder continuous $\phi : X \to R$, for $\kappa = \mathrm{HD}(\mu_\phi)$, Hausdorff dimension of the Gibbs measure μ_ϕ (infimum of Hausdorff dimensions of sets of full measure), either $\mathrm{HD}(X) = \kappa$ the measure μ_ϕ is equivalent to Λ_κ, the Hausdorff measure in dimension κ, and is a *geometric measure*, or μ_ϕ is singular with respect to Λ_κ and the right gauge function for the Hausdorff measure to be compared to μ_ϕ is $\Phi(\kappa) = t^\kappa \exp(c\sqrt{\log 1/t \log\log\log 1/t})$. In the proof we use LIL. This theorem is used to prove a dichotomy for the harmonic measure on a Jordan curve ∂, bounding a domain Ω, which is a repeller for a conformal expanding map. Either ∂ is real analytic or harmonic measure is comparable to the Hausdorff measure with gauge function $\Phi(1)$. This yields an information about the lower and upper growth rates of $|R'(r\zeta)|$, for $r \nearrow 1$, for almost every ζ with $|\zeta| = 1$ and univalent function R from the unit disc $|z| < 1$ to Ω. This is a dynamical counterpart of Makarov's theory of boundary behaviour for general simply connected domains, [Makarov].

We prove in particular that for $f_c(z) = z^2 + c, c \neq 0, c \approx 0$ $1 < \mathrm{HD}(J(f_c)) < 2$.

We show how to express in the language of pressure another interesting function: $\int_{|\zeta|=1} |R'(r\zeta)|^t \, |d\zeta|$ for $r \nearrow 1$.

We also look closer at the Gibbs measures, discuss so called *multifractal analysis*, and study *large deviations*.

Finally we apply our theory to the boundary of von Koch "snowflake" and more general Carleson fractals.

Chapter 8 is devoted to Sullivan's rigidity theorem, saying that two non-linear expanding repellers $(X, f), (Y, g)$ that are Lipschitz conjugate (or more generally there exists a measurable conjugacy that transforms a geometric measure on X to a geometric measure on Y, then the conjugacy extends to a conformal one. This means that measures classify non-linear conformal repellers. This fact, annouced in [Su] only with a sketch of the proof, is proved here rigorously for the first time. We sketch also a generalization by E. Prado.

In Chapter 9 we start to deal with non-uniform expanding phenomena. A heart of this chapter is the proof of the formula $\mathrm{HD}(\mu) = \mathrm{h}_\mu(f)/\chi_\mu(f)$ for an arbitrary f-invariant ergodic measure μ of positive Laypunov exponent $\chi_\mu := \int \log |f'| \, d\mu$.

(The word non-uniform expanding is used just to say that we consider (typical points of) an ergodic measure with positive Lyapunov exponent. In higher dimension one uses the name non-uniform hyperbolic for measures with all Lyapunov exponents non-zero.)

It is so roughly because a small disc around z, whose n-th image is large, has diameter of order $|(f^n)'(z)|^{-1} \approx \exp -n\chi_\mu$ and measure $\exp -n\mathrm{h}_\mu(f)$ (Shannon-McMillan-Breiman theorem is involved here)

Chapter 10 is devoted to conformal measures, namely probability measures with Jacobian $\mathrm{Const}\exp -\phi$ or more specifically $|f'|^\alpha$ in a non-uniformly expanding situation, in particular for any rational mapping f on its Julia set J. It is proved that there exists a minimal exponent $\delta(f)$ for which such a measure exists and that $\delta(f)$ is equal to each of the following quantities:

Dynamical Dimension $\mathrm{DD}(J) := \sup\{\mathrm{HD}(\mu)\}$, where μ ranges over all ergodic f-

invariant measures on J of positive Lyapunov exponent.

Hyperbolic Dimension $\mathrm{HyD}(J) := \sup\{\mathrm{HD}(Y)\}$, where Y ranges over all Conformal Expanding Repellers in J, or CER's that are Cantor sets.

It is an open problem whether for every rational mapping $\mathrm{HyD}(J) = \mathrm{HD}(J) =$ box dimension of J, but for many nonuniformly expandig mappings these equalities hold. It is often easier to study the continuity of $\delta(f)$ with respect to a parameter, than directly Hausdorff dimension. So one obtains an information about the continuity of dimensions due to the above equalities.

Most of the book was written in the years 1990-1992 and was lectured to graduate students by each of us in Warsaw, Yale and Denton. We neglected finishing writing, but recently unexpectedly to us the methods in Chapter 10, relating hyperbolic dimension to minimal exponent of conformal measure, were used to study the dependence on ε of the dimension of Julia set for $z^2 + 1/4 + \varepsilon$, for $\varepsilon \to 0$ and other parabolic bifurcations, by A. Douady, P. Sentenac and M. Zinsmeister in [DSZ] and by C. McMullen in [McM]. So we decided to make a final effort. Meanwhile nice books appeared on some topics of our book, let us mention [Falconer], [Zinsmeister], [Gora,Boyarsky], [Viana], but a lot of important material in our book is new or was hardly accessible, or is written in an unconventional way.

[Barnsley] ....

[Falconer] K. Falconer, Technics in Fractal Geometry

[Zinsmeister] M. Zinsmeister, Le Formalisme Thermodynamique: Mode d'emploi

[Boyarsky, Góra] A. Boyarsky, P. Góra, Laws of Chaos, Invariant Measures and Dynamical Systems in One Dimension. Birkhäuser, Boston 1997

[Viana] M. Viana,

November 14, 1999
chapters renumbered, 2b=3 etc

CHAPTER 0.

BASIC EXAMPLES AND DEFINITIONS

Let us start with definitions of dimensions. We shall come back to them in a more systematic way in Chapter 5.

Definition 0.1. Let (X, ρ) be a metric space. We call *upper (lower) box dimension* of X the quantity

$$\limsup(\liminf)_{r\to 0} \frac{\log N(r)}{-\log r}$$

where $N(r)$ is the minimal number of balls of radius r which cover X.

Sometimes the names *capacity* or *Minkowski dimension* or *box-counting dimension* are used. The name box dimension comes from the situation where X is a subset of a euclidean space $I\!R^d$. Then one can consider only $r = 2^{-n}$ and $N(2^{-n})$ can be replaced by the number of dyadic boxes $[\frac{k_1}{2^{-n}}, \frac{k_1+1}{2^{-n}}] \times ... \times [\frac{k_d}{2^{-n}}, \frac{k_d+1}{2^{-n}}], k_j \in \boldsymbol{Z}$ intersecting X.

Definition 0.2. Let (X, ρ) be a metric space. For every $\kappa > 0$ we define $\Lambda_\kappa(X) = \lim_{\delta\to 0} \inf\{\sum_{i=1}^\infty (\mathrm{diam} U_i)^\kappa\}$, where the infimum is taken over all countable covers $(U_i, i = 1, 2, ...)$ of X by sets of diameter not exceeding δ. $\Lambda_\kappa(Y)$ defined as above on all subsets $Y \subset X$ is called κ-th outer *Hausdorff measure.*

It is easy to see that there exists $\kappa_0 : 0 \le \kappa_0 \le \infty$ such that for all $\kappa : 0 \le \kappa < \kappa_0$ $\Lambda_\kappa(X) = \infty$ and for all $\kappa : \kappa_0 < \kappa$ $\Lambda_\kappa(X) = 0$. The number κ_0 is called the *Hausdorff dimension* of X.

Note that if in this definition we replace the assumption: sets of diameter not exceeding δ by equal δ, and $\lim_{\delta\to 0}$ by $\liminf$ or $\limsup$, we obtain box dimension.

A standard example to compare both notions is the set $\{1/n, n = 1, 2, ...\}$ in $I\!R$. Its box dimension is equal to $1/2$ and Hausdorff dimension is 0. If one considers $\{2^{-n}\}$ instead one obtains both dimensions 0. Also linear Cantor sets in Introduction have Hausdorff and box dimensions equal. The reason for this is self-similarity.

Example 0.3. Shifts spaces. For every natural number d consider the space Σ^d of all infinite sequences $(i_0, i_1, ...)$ with $i_n \in \{1, 2, ..., d\}$. Consider the metric

$$\rho((i_0, i_1, ...), (i'_0, i'_1, ...)) = \sum_{n=0}^{\infty} \lambda^n |i_n - i'_n|$$

for an arbitrary $0 < \lambda < 1$. Sometimes it is more comfortable to use the metric

$$\rho((i_0, i_1, ...), (i'_0, i'_1, ...)) = \lambda^{-\min\{n: i_n \neq i'_n\}}$$

, equivalent to the previous one. Consider $\sigma : \Sigma^d \to \Sigma^d$ defined by $f((i_0, i_1, ...) = (i_1, ...)$. The metric space (Σ^d, ρ) is called *one-sided shift space* and the map σ the *left shift.* Often, if we do not specify metric but are interested only in the cartesian product topology in $\Sigma^d = \{1, ..., d\}^{\boldsymbol{Z}^+}$, we use the name *topological shift space.*

One can consider the space $\tilde{\Sigma}^d$ of all two sides infinite sequences $(..., i_{-1}, i_0, i_1, ...)$. This is called *two-sided shift space.*

Each point $(i_0, i_1, ...) \in \Sigma^d$ determines its forward trajectory under σ, but is equipped with a Cantor set of backward trajectories. Together with the topology determined by the metric $\sum_{n=-\infty}^{\infty} \lambda^{|n|} |i_n - i'_n|$ the set $\tilde{\Sigma}^d$ can be identified with the *inverse limit* (in the topological category) of the system $... \to \Sigma^d \to \Sigma^d$ where all the maps $\to$ are σ.

Note that the limit Cantor set Λ in Introduction, with all $\lambda_i = \lambda$ is Lipschitz homeomorphic to Σ^d, with the homeomorphism h mapping $(i_0, i_1, ...)$ to $\bigcap_k T_{i_0} \circ .. \circ T_{i_k}(I)$. Note that for each $x \in \Lambda$, $h^{-1}(x)$ is the sequence of integers $(i_0, i_1, ...)$ such that for each k, $\hat{T}^k(x) \in T_{i_k}(I)$. It is called a *coding sequence.* If we allow the end points of $T_i(I)$ to overlap, in particular $\lambda = 1/d$ and $a_i = (i-1)/d$, then $\Lambda = I$ and $h^{-1}(x) = \sum_{k=0}^{\infty} (i_k - 1) d^{-k-1}$.

One generalizes the one (or two) -sided shift space, called sometimes *full shift space* by considering the set Σ_A for an arbitrary $d \times d$ – matrix $A = (a_{ij}$ with $a_{ij} = 0$ or 1 defined by

$$\Sigma_A = \{(i_0, i_1, ...) \in \Sigma^d : a_{i_t i_{t+1}} = 1 \quad \text{for every } t = 0, 1, ...\}.$$

By the definition $\sigma(\Sigma_A \subset \Sigma_A$. Σ_A with the mapping σ is called a *topological Markov chain.* Here the word *topological* is substantial, otherwise it is customary to think of a finite number of states stochastic process, see Example 0.8.

Example 0.4. Iteration of rational maps. Let $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ be a holomorphic mapping of the Riemann sphere $\overline{\mathbb{C}}$. Then it must be rational, i.e. ratio of two polynomials. We assume that the topological degree of f is at least 2. The *Julia set* $J(f)$ is defined as follows:

$J(f) = \{z \in \overline{\mathbb{C}} : \forall U \ni z$, U open, the family of iterates $f^n = f \circ ... \circ f|_U$, n times, for $n = 1, 2, ...$ is not *normal* in the sense of Montel $\}$.

A family of holomorphic functions $f_t : U \to \overline{\mathbb{C}}$ is called *normal* (in the sense of Montel) if it is pre-compact, namely from every sequence of functions belonging to the family one can choose a subsequence uniformly convergent (in the spherical metric on the Riemann sphere $\overline{\mathbb{C}}$) on all compact subsets of U.

$z \in J(f)$ implies for example, that for every $U \ni z$ the family $f^n(U)$ covers all $\overline{\mathbb{C}}$ but at most 2 points. Otherwise by Montel's theorem $\{f^n\}$ would be normal on U.

Another characterization of $J(f)$ is that $J(f)$ is the closure of repelling periodic points, namely those points $z \in \overline{\mathbb{C}}$ for which there exists an integer n such that $f^n(z) = z$ and $|(f^n)'(z)| > 1$.

There is only finite number of attracting periodic points, $|(f^n)'(z)| < 1$; they lie outside $J(f)$, an uncountable "chaotic, repelling" Julia set. The lack of symmetry between atracting and repelling phenomena is caused by the non-invertibility of f.

It is easy to prove that $J(f)$ is compact, completely invariant: $f(J(f)) = J(f) = f^{-1}(J(f))$, either nowhere dense or equal to the whole sphere (to prove this use Montel's theorem).

For polynomials, the set of points whose images under iterates $f^n, n = 1, 2, ...,$ tend to ∞, *basin of attraction to* ∞, is connected and completely invariant. Its boundary is the

Julia set.

Check that all these general definitions and statements are compatible with the discussion of $f(z)=z^2+c$ in Introduction. As introduction to this theory we recommend for example the books [Beardon], [Carleson, Gamelin] and [Steinmetz].

Below are the computer pictures of some Julia sets

FIGURES: Rabbit, Sierpinski carpet (rational function of degree 2), Newton's method

A Julia set can have Hausdorff dimension arbitrarily close to 0 (but not 0) and arbitrarily close to 2 and even exactly 2 (being in the same time nowhere dense). It is not known whether it can have positive Lebesgue measure. We shall come back to these topics in Chapters 6, 10.

Example 0.5. Complex linear fractals. The linear Cantor set construction in $\mathbb{R}$ described in Introduction can be generalized to conformal linear Cantor and other fractal sets in $\overline{\mathbb{C}}$:

Let $U \subset \mathbb{C}$ be a bounded connected domain and $T_i(z) = \lambda_i z + a_i$, where λ_i, a_i are complex numbers, $i = 1, ..., n > 1$. Assume that closures $\mathrm{cl}T_i(U)$ are pairwise disjoint and contained in U. The limit Cantor set Λ is defined in the same way as in Introduction.

In Ch.7 we shall prove that it cannot be the Julia set for a holomorphic extension of $\hat{T} = T_i^{-1}$ on $T_i(U)$ for each i, to the whole sphere $\overline{\mathbb{C}}$.

If we allow that the boundaries of $T_i(U)$ intersect or intersect ∂U we obtain other interesting examples

FIGURES: Sierpinski carpet, Sierpinski gasket, boundary of von Koch snowflake

Examples 0.6. Action of Kleinian groups. Beautiful examples of fractal sets arise as limit sets of the action of Kleinian groups on $\overline{\mathbb{C}}$.

Let Ho be the group of all *homographies*, namely the rational mappings of the Riemann sphere of degree 1, i.e. of the form $z \mapsto \frac{az+b}{cz+d}$ where $ad - bc \neq 0$. Every discrete subgroup of Ho is called *Kleinian group*. If all the elements of a Kleinian group preserve the unit disc $\mathbb{D} = \{|z| < 1\}$, the group is called *Fuchsian*.

Consider for example a regular hyperbolic $4n$-gon in $\mathbb{D}$ (equipped with the hyperbolic metric) centered at 0. Denote the consecutive sides by $a_i^j, i = 1, ..., n, j = 1, ..., 4$ in the lexicographical order: $a_1^1, ...a_1^4, a_2^1, ...$ Each side is contained in the corresponding circle C_i^j intersecting $\partial\mathbb{D}$ at the right angles. Denote the disc bounded by C_i^j by D_i^j.

It is not hard to see that the closures of D_i^j and D_i^{j+2} are disjoint for each i and $j = 1, 2$.

FIGURE: regular hyperbolic octagon, $\mathbb{D}/G$.

Let $g_i^j, j = 1, 2$ be the unique homography preserving $\mathbb{D}$ mapping a_i^j to a_i^{j+2} and D_i^j to the complement of $\mathrm{cl}D_i^{j+2}$. It is easy to see that the family $\{g_i^j\}$ generates a Fuchsian group G. For an arbitrary Kleinian group G, the *Poincaré limit set* $\Lambda(G) = \bigcup \lim_{k\to\infty} g_k(z)$, the union taken over all sequences of pairwise different $g_k \in G$ such that $g_k(z)$ converges, where z is an arbitrary point in $\overline{\mathbb{C}}$. It is not hard to prove that $\Lambda(G)$ does not depend on z.

For the above example $\Lambda(G) = \partial\mathbb{D}$. If we change slightly g_i^j (the circles C_i^j change slightly), then either $\Lambda(G)$ is a circle S (all new C_i^j intersect S at the right angle), or it is a fractal Jordan curve. The phenomenon is similar to the case of the maps $z \mapsto z^2 + c$ described in Introduction. For details see [Bowen], [Bowen, Series], [Sullivan]. We provide a sketch of the proof in Chapter 7.

If all the closures of the discs $D_i^j, i = 1, ..., n, j = 1, ..., 4$ become pairwise disjoint, $\Lambda(G)$ becomes a Cantor set (the group is called then a *Schottky group* or a *Kleinian group of Schottky type*).

Examples 0.7. Higher dimensions. Though the book is devoted to 1-dimensional real and complex iteration and arising fractals, Chapters 1-3 apply to general situations. A basic example is Smale's horseshoe. Take a square $K = [0,1] \times [0,1]$ in the plane $\mathbb{R}^2$ and map it affinely to a strip by squeezing in the horizontal direction and stretching in the vertical, for example $f(x,y) = (\frac{1}{3}x + \frac{1}{9}, 3y - \frac{1}{3})$ and bend the strip by a new map g so that the rectangle $[\frac{1}{9}, \frac{4}{9}] \times [\frac{4}{3}, \frac{8}{3}]$ is mapped to $[\frac{5}{9}, \frac{8}{9}] \times [-\frac{1}{3}, 1]$. The resulting composition $T = g \circ f$ maps K to a "horseshoe", see [Smale, p.773]

FIGURE: horseshoe, stadium extension

The map can be easily extended to a C^∞-diffeomorphism of $\overline{\mathbb{C}}$ by mapping a "stadium" extending K to a bent "stadium", and its complement to the respective complement. The set Λ^K of points not leaving K under action of $T^n, n = ..., -1, 0, 1, ...$ is the cartesian product of two Cantor sets. This set is T-invariant, "uniformly hyperbolic". In the horizontal direction we have contraction, in the vertical direction uniform expansion. The situation is different from the previous examples of Σ^d or linear Cantor sets, where we had uniform expansion in all directions.

Smale's horseshoe is a universal phenomenon. It is always present for an iterate of a diffeomorphism f having a *transversal homoclinic point* q for a saddle p. The stable and unstable manifolds $W^s(p) := \{y : f^n(y) \to p\}, W^u(p) := \{y : f^{-n}(y) \to p\}$ as $n \to \infty$, intersect transversally at q. For more details on hyperbolic sets see [HK].

FIGURE: homoclinic point and embedded horseshoe.

Note that $T|_{\Lambda^K}$ is *topologically conjugate* to the left shift σ on the two-sided shift space $\tilde{\Sigma}^2$, namely there exists a homeomorphism $h : \Lambda^K \to \tilde{\Sigma}^2$ such that $h \circ T = \sigma \circ h$. Compare h i Example 0.3. T on Λ^K is the inverse limit of the mapping $\hat{T}$ on the Cantor set described in Introduction, similarly to the inverse limit $\tilde{\Sigma}^2$ of σ on Σ^2. The philosophy is that hyperbolic systems appear as inverse limits of expanding systems.

A partition of a hyperbolic set Λ into local stable (unstable) manifolds: $W^s(x) = \{y \in \Lambda : (\forall n \geq 0)\rho(f^n(x), f^n(y)) \leq \varepsilon(x)\}$ for a small positive measurable function ε, is an illustration of an abstract ergodic theory measurable partition ξ such that $f(\xi)$ is finer than ξ, $f^n(\xi), n \to \infty$ converges to the partition into points and the conditional entropy $\mathrm{H}_\mu(f(\xi)|\xi)$ is maximal possible, equal to the entropy $\mathrm{h}_\mu(f)$; all this holds for an ergodic invariant measure μ.

The inverse limit of the system $... \to S^1 \to S^1$ where all the maps are $z \mapsto z^2$, is called a solenoid. It has a group structure: $(..., z_{-1}, z_0) \cdot (..., z'_{-1}, z'_0) = (..., z_{-1} \cdot z'_{-1}, z_0 \cdot z'_0)$, which is a trajectory if both factors are, since the map $z \mapsto z^2$ is a homomorphism of the group S^1. Topologically the solenoid can be represented as the attractor A of the mapping of the solid torus $I\!\!D \times S^1$ into itself $f(z,w) = (\frac{1}{3}z + \frac{1}{2}w, w^2)$. Its Hausdorff dimension is equal in this special example to $1 + \mathrm{HD}(A \cap \{w = w_0\}) = 1 + \frac{\log 2}{\log 3}$ for an arbitrary w_0, as Cantor sets $A \cap \{w = w_0\}$ have Hausdorff dimensions $\frac{\log 2}{\log 3}$. These are linear Cantor sets discussed in Introduction.

Especially interesting is the question of Hausdorff dimension of A if $z \mapsto \frac{1}{3}z$ is replaced by $z \mapsto \phi(z)$ not conformal. But this higher dimensional problem goes beyond the scope of our book. See [Pesin].

If the map $z \mapsto z^2$ in the definition of solenoid is replaced by an arbitrary rational mapping then if f is expanding on the Julia set, the solenoid is locally the cartesian product of an open set in $J(f)$ and the Cantor set of all possible choises of backward trajectories. If however there are critical points in $J(f)$ (or converging under the action of f^n to parabolic points in $J(f)$) the solenoid (inverse limit) is more complicated, see [LM] for an attempt to describe it, together with a neighbourhood composed of trajectories outside $J(f)$. We shall not discuss this in our book.

Examples 0.8. Bernoulli shifts and Markov chains. For every positive numbers $p_1, ..., p_d$ such that $\sum_{i=1}^d p_i = 1$, one introduces on the Borel subsets of Σ^d (or $\tilde{\Sigma}^d$) a probability measure μ by extending to the σ-algebra of all Borel sets the function $\mu(C_{i_0,i_1,...,i_t}) = p_0 p_1 ... p_t$, where $C_{i_0,i_1,...,i_t} = \{(i'_0, i'_1, ...) : i'_s = i_s \text{ for every } s = 0, 1, ..., t\}$. Each such C is called a *finite cylinder*.

The space Σ^d with the left shift σ and the measure μ is called one-sided *Bernoulli shift*.

On a topological Markov chain $\Sigma_A \subset \Sigma^d$ with $A = (a_{ij})$ and an arbitrary $d \times d$ matrix $M = p_{ij}$ such that $\sum_{j=1}^d p_{ij} = 1$ for every $i = 1, ..., d$, $p_{ij} \geq 0$ and $p_{ij} = 0$ if $a_{ij} = 0$, one can introduce a probability measure μ on all Borel subsets of Σ_A by extending $\mu(C_{i_0,i_1,...,i_t}) = p_{i_0} p_{i_0 i_1} ... p_{i_{t-1} i_t}$. Here $(p_1, ..., p_d)$ is an eigenvector of M^*, namely $\sum_i p_i p_{ij} = p_j$, such that $p_i \geq 0$ for every $i = 1, ..., d$ and $\sum_{i=1}^d = 1$.
The space Σ_A with the left shift σ and the measure μ is called one-sided *Markov chain*.
Note that μ is σ-invariant. Indeed,

$$\mu(\bigcup_i (C_{i,i_0,...,i_t})) = \sum_i p_i p_{i i_0} p_{i_0 i_1} ... p_{i_{t-1} i_t} = p_{i_0} p_{i_0 i_1} ... p_{i_{t-1} i_t} = \mu(C_{i_0,...,i_t}).$$

As in the topological case if we consider $\tilde{\Sigma}^d$ rather than Σ^d, we obtain two-sided Bernoulli shifts and two-sided Markov chains.

Example 0.9. Tchebyshev polynomial Let us consider the mapping $T : [-1,1] \to [-1,1]$ of the real interval $[-1,1]$ defined by $T(x) = 2x^2 - 1$. In the co-ordinates $z \mapsto 2z$ it is just a restriction to an invariant interval of the mapping $z \mapsto z^2 - 2$ discussed already in Introduction. The interval $[-1,1]$ is Julia set of T.

Notice that this map is the factor of the mapping $z \mapsto z^2$ on the unit circle $\{|z| = 1\}$ in $\mathbb{C}$ by the orthogonal projection P to the real axis. Since the length measure l is preserved by $z \mapsto z^2$ its projection is preserved by T. Its density with respect to the Lebesgue measure on $[-2,2]$ is proportional to $(dP/dl)^{-1}$, after normalization is equal to $\frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}$. This measure satisfiesmany properties of Gibbs invariant measures discussed in Chapter 4, though T is not expanding; it has a critical point at 0. This T is the simplest example of non-uniformly expanding maps to which the advanced parts of the book are devoted.

[Smale] S. Smale, Differentiable Dynamical Systems. Bulletin of the American Mathematical Society 73 (1967), 747-817.

[Steinmetz] N. Steinmetz, Rational Iteration, Complex Dynamics, Dynamical Systems, Walter de Gruyter, Berlin 1993

version Dec.7, 1999
last revised Feb. 23, 2002

CHAPTER 1
MEASURE PRESERVING ENDOMORPHISMS

§1.1 MEASURE SPACES AND MARTINGALE THEOREM

We assume that the reader knows basic elements of measure and integral theory. For a complete treatment see for example [Halmos] or [Billingsley, 1979]. We start with some basics to fix notation and terminology.

A family $\mathcal{F}$ of subsets of a set X is said to be a *σ-algebra* if the following conditions are satisfied:

$$X \in \mathcal{F}, \tag{1.1.1}$$

$$A \in \mathcal{F} \;\Rightarrow\; A^c \in \mathcal{F} \tag{1.1.2}$$

and

$$\{A_i\}_{i=1}^{\infty} \subset \mathcal{F} \;\Rightarrow\; \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \tag{1.1.3}$$

It follows from this definition that $\emptyset \in \mathcal{F}$, that the σ-algebra $\mathcal{F}$ is closed under countable intersections and subtractions of sets. If (1.1.3) is assumed only for finite subfamilies of $\mathcal{F}$ then $\mathcal{F}$ is called an *algebra*. Fixed $\mathcal{F}$, elements of the σ–algebra $\mathcal{F}$ will be frequently called *measurable* sets. For any family $\mathcal{F}_0$ of subsets of X, we denote by $\sigma(\mathcal{F}_0)$ the minimal σ–algebra that contains $\mathcal{F}_0$ and call it the σ–algebra *generated* by $\mathcal{F}_0$.

A function on a σ-algebra $\mathcal{F}$, $\mu : \mathcal{F} \to [0, \infty]$, is said to be *$\sigma$-additive* if for any countable subfamily $\{A_i\}_{i=1}^{\infty}$ of $\mathcal{F}$ consisting of mutually disjoint sets, we have

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) \tag{1.1.4}$$

We say then that μ is a *measure*. If we consider in (1.1.4) only finite families of sets, we say μ is *additive*. The two notions: of *additive* and of *σ-additive*, make sense for a σ-algebra as well as for an algebra, provided in the algebra case in (1.1.4) that all A_i and their union belong to $\mathcal{F}$. The simplest consequences of the definition of measure are the following:

$$\mu(\emptyset) = 0; \tag{1.1.5}$$

(1.1.6) $\qquad$ if $A, B \in \mathcal{F}$ and $A \subset B$ then $\mu(A) \le \mu(B)$;

(1.1.7) $\qquad$ if $A_1 \subset A_2 \subset \ldots$ and $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ then $\mu(\bigcup_{i=1}^{\infty} A_i) = \sup_i \mu(A_i) = \lim_{i\to\infty} \mu(A_i)$.

We say that the triple $(X, \mathcal{F}, \mu)$ with a σ-algebra $\mathcal{F}$ and μ a measure on $\mathcal{F}$ is a *measure space*. In this book we will always assume, unless the opposite is stated, that μ is a *finite measure* that is $\mu : \mathcal{F} \to [0, \infty)$. By (1.1.6) it equivalently means that $\mu(X) < \infty$. If $\mu(X) = 1$, the triple $(X, \mathcal{F}, \mu)$ is called a*probability space* and μ a *probability measure*.

We say that $\phi \to I\!R$ is a measurable function, if $\phi^{-1}(J) \in \mathcal{F}$ for every interval $J \subset I\!R$ (compare Sec.2). We say that ϕ is μ-integrable if $\int |\phi|\, d\mu < \infty$. We write $\phi \in L^1(\mu)$. More generally, for every $1 \le p < \infty$ we write $(\int |\phi|^p\, d\mu)^{1/p} = ||\phi||_p$ and say ϕ belongs to $L^p(\mu) = L^p(X, \mathcal{F}, \mu)$. If $\inf_{\mu(E)=0} \sup_{X\setminus E} |\phi| < \infty$ we say $\phi \in L^\infty$ and denote the latter expression by $||\phi||_\infty$. $||\phi||_p, 1 \le p \le \infty$ are called L^p-norms of ϕ. We usually identify in this chapter functions which differ only on a set of μ-measure 0. $L^p(X, \mathcal{F}, \mu)$'s after these identifications are Banach spaces.

We say that a property $q(x)$, $x \in X$, is satisfied for μ almost every $x \in X$ (abbr: a.e.), or μ-a.e., if $\mu(\{x : q(x)$ is not satisfied$\}) = 0$. We can consider q as a subset of X with $\mu(X \setminus q) = 0$.

We shall often use in the book the following two facts.

Monotone Convergence Theorem. Suppose $\phi_1 \le \phi_2 \le ...$ is an increasing sequence of integrable, real-valued functions on a probability space $(X, \mathcal{F}, \mu)$. Then $\phi = \lim_{n\to\infty} \phi_n$ exists a.e. and $\lim_{n\to\infty} \int \phi_n\, d\mu = \int \phi\, d\mu$. (We allow $+\infty$'s here.)

and

Dominated Convergence Theorem. If $\phi_n, n \ge 1$ is a sequence of measurable real-valued functions on a probability space $(X, \mathcal{F}, \mu)$ and $|\phi_n| \le g$ for an integrable function g and $\phi_n \to \phi$ a.e., then ϕ is integrable and $\lim_{n\to\infty} \int \phi_n\, d\mu = \int \phi\, d\mu$.

Recall now that if $\mathcal{F}'$ is a sub-σ-algebra of $\mathcal{F}$ and $\phi : X \to I\!R$ is a μ-integrable function, then there exists a unique (mod 0) function usually denoted by $E(\phi|\mathcal{F}')$ such that $E(\phi|\mathcal{F}')$ is $\mathcal{F}'$-measurable and

(1.1.8)
$$\int_A E(\phi|\mathcal{F}')\, d\mu = \int_A \phi\, d\mu$$

for all $A \in \mathcal{F}'$. $E(\phi|\mathcal{F}')$ is called *conditional expectation value* of the function ϕ with respect to the σ-algebra $\mathcal{F}'$. Sometimes we shall use for $E(\phi|\mathcal{F}')$ the simplified notation $\phi_{\mathcal{F}'}$.

For $\mathcal{F}$ generated by a finite partition $\mathcal{A}$ (cf. Sec.3), one can think of $E(\phi|\sigma(\mathcal{A})$ as constant on each $A \in \mathcal{A}$ equal to the average $\int_A \phi \, d\mu / \mu(A)$.

The existence of $E(\phi|\mathcal{F}')$ follows from famous the Radon-Nikodym theorem, saying that if $\nu \ll \mu$, both measures defined on the same σ-algebra $\mathcal{F}'$, ($\nu \ll \mu$ means ν *absolutely continuous* with respect to μ, i.e. $\mu(A) = 0 \Rightarrow \nu(A) = 0$ for all $A \in \mathcal{F}'$), then there exists a unique (mod 0) $\mathcal{F}'$-measurable, μ-integrable function $\Phi = d\nu/d\mu : X \to \mathbb{R}^+$ such that for every $A \in \mathcal{F}'$

$$\int_A \Phi \, d\mu = \nu(A).$$

To deduce (1.1.8) we set $\nu(A) = \int_A \phi \, d\mu$ for $\mathcal{A} \in \mathcal{F}'$. The trick is that we restrict μ from $\mathcal{F}$ to $\mathcal{F}'$.

If $\phi \in L^p(X, \mathcal{F}, \mu)$ then $E(\phi|\mathcal{F}') \in L^p(X, \mathcal{F}', \mu)$ for all σ-algebras $\mathcal{F}'$ with L^p norms uniformly bounded. More precisely the operators $\phi \to E(\phi|\mathcal{F}')$ are linear projections from $L^p(X, \mathcal{F}, \mu)$ to $L^p(X, \mathcal{F}', \mu)$, with L^p-norms equal to 1 (see Exercise 0.).

We end this section with the following version of **Martingale Convergence Theorem**.

Theorem 1.1.1. If $(\mathcal{F}_n : n \geq 1)$ is either monotone increasing or monotone decreasing sequence of σ-algebras contained in $\mathcal{F}$, then for every $\phi \in L^p(\mu)$, $1 \leq p < \infty$

$$\lim_{n \to \infty} E(\phi|\mathcal{F}_n) = E(\phi|\mathcal{F}'), \quad \text{a.e. and in } \ L^p,$$

where $\mathcal{F}'$ is equal to either $\bigvee_{n=1}^{\infty} \mathcal{F}_n$ or $\bigcap_{n=1}^{\infty} \mathcal{F}_n$ respectively.

In the theorem above we denoted by $\bigvee_{n=1}^{\infty} \mathcal{F}_n$ the smallest σ-algebra containing $\bigcup_{n=1}^{\infty} \mathcal{F}_n$, the latter usually is not a σ-algebra, but only an algebra. Compare Sec.6 where complete σ-algebras of this form in Lebesgue space are considered.

Remark 1.1.2. For the existence of $\mathcal{F}'$ and the convergence in L^p no monotonicity is needed. It is sufficient to assume that for every $A \in \mathcal{F}$ the limit $\lim E(1\!\!1_A|\mathcal{F}_n)$ in measure μ exists.

(Recall that ψ_n is said to *converge in measure* ν to ψ if for every $\varepsilon > 0$, $\lim_{n \to \infty} \mu(\{x \in X : |\psi_n(x) - \psi(x)| \geq \varepsilon\}) \to 0$.)

In this book we denote by $1\!\!1_A$ the indicator function of A, namely equal to 1 on A and to 0 outside A.

We shall not provide here a proof of Theorem 1.1.1 in the full generality . Let us provide however a proof Theorem 1.1.1 (and Remark 1.1.2 in the case $\lim E(1\!\!1_A|\mathcal{F}_n) = 1\!\!1_A$) for the L^2-convergence for functions $\phi \in L^2(\mu)$. (This is the case sufficient for example to prove the important Lemma 1.8.6 later on in this chapter.)

For the increasing sequence $(\mathcal{F}_n)$ we have the equality $L^2(X, \mathcal{F}', \mu) = \overline{\bigcup_n L^2(X, \mathcal{F}_n, \mu)}$. Indeed, for every $B \in \mathcal{F}'$ there exists a sequence $B_n \in \mathcal{F}_n$, $n \geq 1$, such that $\mu(B \div B_n) \to 0$. ($B \div C = (B \setminus C) \cup (C \setminus B)$ is the *symmetric difference* of sets B and C.)

This follows for example from Carathéodory's argument, see the note Theorem 1.7.2. We have $\mu(B)$ equal to the outer measure of B constructed from μ restricted to the algebra $\bigcup_{n=1}^{\infty} \mathcal{F}_n$. In the Remark 1.1.2 case where we assumed $\lim E(1\!\!1_A|\mathcal{F}_n) = 1\!\!1_A$, this is immediate.

Hence $L^2(X, \mathcal{F}_n, \mu) \ni 1\!\!1_{B_n} \to 1\!\!1_B$ in $L^2(X, \mathcal{F}, \mu)$. Finally use the fact that every function $f \in L^2(X, \mathcal{F}', \mu)$ can be approximated in the space $L^2(X, \mathcal{F}', \mu)$ by the step functions, i.e. finite linear combinations of indicator functions. Therefore, since $E(\phi|\mathcal{F}_n)$ and $E(\phi|\mathcal{F}')$ are orthogonal projections of ϕ to $L^2(X, \mathcal{F}_n, \mu)$ and $L^2(X, \mathcal{F}', \mu)$ respectively (exercise) we obtain $E(\phi|\mathcal{F}_n) \to E(\phi|\mathcal{F}')$ in L^2.

For a decreasing sequence $\mathcal{F}_n$ use the equality $L^2(X, \mathcal{F}', \mu) = \bigcap_n L^2(X, \mathcal{F}_n, \mu)$.

§1.2 MEASURE PRESERVING ENDOMORPHISMS, ERGODICITY

Let $(X, \mathcal{F}, \mu)$ and $(X', \mathcal{F}', \mu')$ be measure spaces. A transformation $T : X \to X'$ is said to be *measurable* if $T^{-1}(A) \in \mathcal{F}$ for every $A \in \mathcal{F}'$. If moreover $\mu(T^{-1}(A)) = \mu'(A)$ for every $A \in \mathcal{F}'$, then T is called *measure preserving.* If $(X, \mathcal{F}, \mu) = (X', \mathcal{F}', \mu')$ we call T a *measure preserving endomorphisms*; we will say also that measure μ is T–invariant, or that T preserves μ.

If a measure preserving map T is invertible and the inverse T^{-1} is measurable, then clearly T^{-1} is also measure preserving. Therefore T is an *isomorphism* in the category of measure spaces. In the case of $(X, \mathcal{F}, \mu) = (X', \mathcal{F}', \mu')$ the transformation T is called an *automorphism.*

We shall prove now the following very useful fact in which the finitness of measure is a crucial assumption.

Theorem 1.2.1. (Poincaré Recurrence Theorem) If $T : X \to X$ is a measure preserving endomorphism, then for every mesurable set A

$$\mu\big(\{x \in A : T^n(x) \in A \ \text{ for infinitely many } n\text{'s}\}\big) = \mu(A).$$

Proof. Let

$$N = N(T, A) = \{x \in A : T^n(x) \notin A \ \forall n \geq 1\}.$$

We shall first show that $\mu(N) = 0$. Indeed, N is measurable since $N = A \cap \big(\bigcap_{n \geq 1} T^{-n}(X \setminus A)\big)$. If $x \in N$, then $T^n(x) \notin A$ for all $n \geq 1$ and, in particular, $T^n(x) \notin N$ which implies that $x \notin T^{-n}(N)$ and consequently $N \cap T^{-n}(N) = \emptyset$ for all $n \geq 1$. Thus, all the sets N, $T^{-1}(N)$, $T^{-2}(N), \ldots$ are mutually disjoint since if $n_1 \leq n_2$, then

$$T^{-n_1}(N) \cap T^{-n_2}(N) = T^{-n_1}(N \cap T^{-(n_2 - n_1)}(N)) = \emptyset.$$

Hence

$$1 \geq \mu\left(\bigcup_{n=0}^{\infty} T^{-n}(N)\right) = \sum_{n=0}^{\infty} \mu(T^{-n}(N)) = \sum_{n=0}^{\infty} \mu(N).$$

Therefore $\mu(N) = 0$. Fix now $k \geq 1$ and put

$$N_k = \{x \in A : T^n(x) \notin A \; \forall n \geq k\}.$$

Then $N_k \subset N(T^k, A)$ and therefore from what have been proved above it follows that $\mu(N_k) \leq \mu(N(T^k, A)) = 0$. Thus

$$\mu\big(\{x \in A : T^n(x) \in A \text{ for only finitely many } n\text{'s}\}\big) = 0.$$

The proof is finished. ♣

A measurable transformation $T : X \to X$ of a measure space $(X, \mathcal{F}, \mu)$ is said to be *ergodic* if for any measurable set A

$$\mu(T^{-1}(A) \div A) = 0 \quad \Rightarrow \quad \mu(A) = 0 \text{ or } \mu(X \setminus A) = 0$$

(Recall the notation $B \div C = (B \setminus C) \cup (C \setminus B)$.)

Note that we did not assume in the definition of ergodicity that μ is T-invariant (neither that μ is finite). Suppose that for every E of measure 0 the set $T^{-1}(E)$ is also of measure 0 (in Ch.4 we call this property of μ with respect to T, *backward quasi-invariant*). Then in the definition of ergodicity one can replace $\mu(T^{-1}(A) \div A) = 0$ by $T^{-1}(A) = A$. Indeed having A as in the definition one can define $A' = \bigcup_{n=0}^{\infty} \bigcap_{m=n}^{\infty} T^{-m}(A)$. Then $\mu(A') = \mu(A)$ and $T^{-1}(A') = A'$. If we assumed that the latter implies $\mu(A') = 0$ or $\mu(X \setminus A') = 0$, then $\mu(A) = 0$ or $\mu(X \setminus A) = 0$.

Let $\phi : X \to I\!\!R$ be a measurable function. For any $n \geq 1$ we define

$$S_n\phi = \phi + \phi \circ T + \ldots + \phi \circ T^{n-1} \tag{1.2.1}$$

Let $\mathcal{I} = \{A \in \mathcal{F} : \mu(T^{-1}(A) \div A) = 0\}$. We call it σ-algebra of T-invariant (mod 0) sets. Note that every $\psi : X \to I\!\!R$, measurable with respect to $\mathcal{I}$, is T-invariant (mod 0), namely $\psi \circ T = \psi$, but on a set of measure μ equal to 0. Indeed let $A = \{x \in X : \psi(x) \neq \psi \circ T(x)\}$ and suppose $\mu(A) > 0$. Hence there exists $a \in I\!\!R$ such that $A_a = \{x \in A : \psi(x) < a, \psi \circ T(x) > a\}$ and $\mu(A_a) > 0$. (or a similar A_a with reversed inequalities). Since $A_a \in \mathcal{I}$, there exists $E \subset A$ of measure 0, such that $T(A_a \setminus E) \subset A_a$, hence on $A_a \setminus E$ we have $\psi = \psi \circ T$. We arrived at a contradiction.

Theorem 1.2.2. (Birkhoff's Ergodic Theorem) If $T : X \to X$ is a measure preserving endomorphism of a probability space $(X, \mathcal{F}, \mu)$ and $\phi : X \to I\!\!R$ is an integrable function then

$$\lim_{n\to\infty} \frac{1}{n} S_n\phi(x) = E(\phi|\mathcal{I}) \text{ for } \mu\text{-a.e. } x \in X$$

We say that the *time average* exists for μ-almost every $x \in X$.

In particular Theorem 1.2.2 yields for T ergodic preserving μ, that

$$\lim_{n\to\infty} \frac{1}{n} S_n\phi(x) = \int \phi \, d\mu, \quad \text{for } \mu\text{-a.e. } x \tag{1.2.1a}$$

We say that the *time average equals the space average.*

If $\phi = 1\!\!1_A$, the indicator function of a measurable set A, then we deduce that for a.e. x the frequency of hitting A by the forward trajectory equals to the measure (probability) of A, namely $\lim_{n\to\infty} \#\{0 \le j < n : T^j(x) \in A\}/n$, is equal to $\mu(A)$.

This means for example that if we choose a point in X in a euclidean space at random its sufficiently long forward trajectory fills X with the density being approximately the density of μ with respect to the Lebesgue measure, provided μ is equivalent to the Lebesgue measure.

On the figure below, Fig.1.2a, for a randomly chosen backward trajectory $x_j, j = 0, 1, ..., n,, \ T(x_k) = x_{k-1}$, for $T(x) = 2x^2 - 1$ (see Example 0.9), for the interval $[-1, 1]$ divided into $k = 100$ equal pieces, the graph of the function $-1 + 2t/100 \mapsto 100 \cdot \#\{0 \le j < n : -1 + 2t/100 \le x_j(x) < -1 + 2(t+1)/100\}/n$ is plotted. It indeed resembles the graph of $1/\pi\sqrt{1-x^2}$, Fig.1.2b, which is the density of the invariant probability measure equivalent to the length measure.

FIGURE 1.1, The density of an invariant measure for $T(x) = 2x^2 - 1$.

As a corollary of Birkhoff's Ergodic Theorem one can obtain von Neumann's Ergodic Theorem. It says that if $\phi \in L^p(\mu)$ for $1 \le p < \infty$, then the convergence to $E(\phi|\mathcal{I})$ holds in L^p. It is not difficult, see for example [Wa].

Proof of Birkhoff's Ergodic Theorem. Let $f \in L^1(\mu)$ and $F_n = \max\{\sum_{i=0}^{k-1} f \circ T^i : 1 \le k \le n\}$, for $n = 1, 2, ...$ Then for every $x \in X$, $F_{n+1}(x) - F_n(T(x)) = f(x) - \min(0, F_n(T(x))) \ge f(x)$ and is monotone decreasing, since F_n is monotone increasing. Define

$$A = \{x : \sup_n \sum_{i=0}^{n} f(T^i(x)) = \infty\}$$

If $x \in A$ then $F_{n+1}(x) - F_n(T(x))$ monotonously decreases to $f(x)$ as $n \to \infty$. The Dominated Convergence Theorem then implies that

$$0 \le \int_A (F_{n+1} - F_n) \, d\mu = \int_A (F_{n+1} - F_n \circ T) \, d\mu \to \int_A f d\mu. \tag{1.2.2}$$

(We arrived at $\int_A f \, d\mu \ge 0$, which is a variant of so-called Maximal Ergodic Theorem.)

Notice that $\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \le F_n/n$, so outside A we have

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \le 0. \tag{1.2.3}$$

Therefore, if the conditional expectation value $f_{\mathcal{I}}$ of f is negative a.e., that is if $\int_C f d\mu = \int_C f_{\mathcal{I}} d\mu < 0$ for all $C \in \mathcal{I}$ with $\mu(C) > 0$, then, since by definition $A \in \mathcal{I}$, (1.2.2) implies that $\mu(A) = 0$, and hence (1.2.3) holds a.e.. Now if we let $f = \phi - \phi_{\mathcal{I}} - \varepsilon$, then $f_{\mathcal{I}} = -\varepsilon < 0$. Note that $\phi_{\mathcal{I}} \circ T = \phi_{\mathcal{I}}$ implies

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k = \Big(\frac{1}{n}\sum_{k=0}^{n-1} \phi \circ T^k\Big) - \phi_{\mathcal{I}} - \varepsilon.$$

So (1.2.3) yields

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \phi \circ T^k \le \phi_{\mathcal{I}} + \varepsilon \quad \text{a.e.}$$

Replacing ϕ by $-\phi$ gives

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \phi \circ T^k \ge \phi_{\mathcal{I}} - \varepsilon \quad \text{a.e.}$$

Thus $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \phi \circ T^k = \phi_{\mathcal{I}}$ a.e. ♣

Recall that at the end opposite to the absolute continuity (see Sec.1) there is the notion of singularity. Finite measures μ_1 and μ_2 on a σ-algebra $\mathcal{F}$ are called mutually *singular*, $\mu_1 \perp \mu_2$ if there exist disjoint sets $X_1, X_2 \in \mathcal{F}$ with $\mu_i(X_i) = 1$ for $i = 1, 2$.

Theorem 1.2.3. If $T : X \to X$ is a map measurable with respect to a σ-algebra $\mathcal{F}$ and if μ_1 and μ_2 are two different T-invariant probability ergodic measures on $\mathcal{F}$, then μ_1 and μ_2 are singular.

Proof. Since μ_1 and μ_2 are different, there exists a measurable set A such that

$$\mu_1(A) \neq \mu_2(A) \tag{1.2.2}$$

By Theorem 1.2.2 (Birkhoff's Ergodic Theorem) applied to μ_1 and μ_2 there exist sets $X_1, X_2 \in \mathcal{F}$ such that for every $i = 1, 2$ and every $x \in X_i$

$$\lim_{n\to\infty} \frac{1}{n} S_n 1\!\!1_A(x) = \mu_i(A)$$

and $\mu_i(X_i) = 1$. Thus in view of (1.2.2) the sets X_1 and X_2 are disjoint. The proof is finished. ♣

Proposition 1.2.4. If $T : X \to X$ is a measure preserving endomorphism of a probability space $(X, \mathcal{F}, \nu)$, then ν is ergodic if and only if there is no T-invariant probability measure on $\mathcal{F}$ absolutely continuous with respect to ν and different from ν.

Proof. Suppose that ν is ergodic and μ is a T-invariant probability measure on $\mathcal{F}$ with $\mu \ll \nu$. Then μ is also ergodic. Otherwise there would exist A such that $T^{-1}(A) = A$ and $\mu(A), \mu(X \setminus A) > 0$ so $\nu(A), \nu(X \setminus A) > 0$ so ν would not be ergodic. Hence by Theorem 1.2.3 $\mu = \nu$.

Suppose in turn that ν is not ergodic and let $A \in \mathcal{F}$ be a T-invariant set such that $0 < \nu(A) < 1$. Then the conditional measure on A is also T-invariant but simultaneously it is distinct from ν and absolutely continuous with respect to ν. The proof is finished. ♣

Observe now that the space $M(\mathcal{F})$ of probability measures on $\mathcal{F}$ is a convex set i.e. the convex combination $\alpha\mu + (1-\alpha)\nu$, $0 \le \alpha \le 1$, of two such measures is again in $M(\mathcal{F})$. The subspace $M(\mathcal{F}, T)$ of $M(\mathcal{F})$ consisting of T-invariant measures is also convex.

Recall that a point in a convex set is said to be *extreme* if and only if it cannot be represented as a convex combination of two distinct points with corresponding coefficient $0 < \alpha < 1$. We shall prove the following.

Theorem 1.2.5. The ergodic measures in $M(\mathcal{F}, T)$ are exactly the extreme points of $M(\mathcal{F}, T)$.

Proof. Suppose that $\mu, \mu_1, \mu_2 \in M(\mathcal{F}, T)$, $\mu_1 \neq \mu_2$ and $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ with $0 < \alpha < 1$. Then $\mu_1 \neq \mu$ and $\mu_1 \ll \mu$. Thus in view of Proposition 1.2.4 mesure μ is not ergodic.

Suppose in turn that μ is not ergodic and let $A \in \mathcal{F}$ be a T-invariant set such that $0 < \mu(A) < 1$. Recall that given $B \in \mathcal{F}$ with $\mu(B) > 0$ the *conditional measure* $A \to \mu(A|B)$ is defined by $\mu(A \cap B)/\mu(B)$. Thus the conditional measures $\mu(\cdot|A)$ and $\mu(\cdot|A^c)$ are distinct, T-invariant and $\mu = \mu(A)\mu(\cdot|A) + (1 - \mu(A)\mu(\cdot|A^c)$. Consequently μ is not en extreme point in $M(\mathcal{F}, T)$. The proof is finished. ♣

In Section 8 we shall formulate a theorem on decomposition into ergodic components, that will better clear the situation. This will correspond the Choquet Theorem in functional analysis, see Ch.2.1.

§1.3. ENTROPY OF PARTITION

Let $(X, \mathcal{F}, \mu)$ be a probability space. A *partition* of $(X, \mathcal{F}, \mu)$ is a subfamily (*a priori* may be uncountable) of $\mathcal{F}$ consisting of mutually disjoint elements whose union is X.
If $\mathcal{A}$ is a partition and $x \in X$ then the only element of $\mathcal{A}$ containing x is denoted by $\mathcal{A}(x)$ or, if $x \in A \in \mathcal{A}$, by $A(x)$.

If $\mathcal{A}$ and $\mathcal{B}$ are two partitions of X we define their *join*

$$\mathcal{A} \vee \mathcal{B} = \{A \cap B : A \in \mathcal{A},\ B \in \mathcal{B}\}$$

We write $\mathcal{A} \le \mathcal{B}$ if and only if $\mathcal{B}(x) \subset \mathcal{A}(x)$ for every $x \in X$, which in other words means that each element of the partition $\mathcal{B}$ is contained in an element of the partition $\mathcal{A}$ or equivalently $\mathcal{A} \vee \mathcal{B} = \mathcal{B}$. We sometimes say in this case, that $\mathcal{B}$ is *finer* than $\mathcal{A}$ or that $\mathcal{B}$ is a *refinement* of $\mathcal{A}$.

Now we introduce the notion of entropy of a countable (this word includes in this book: finite) partition and we collect its basic elementary properties. Define the function $k : [0,1] \to [0,\infty]$ putting

$$k(t) = \begin{cases} -t \log t & \text{for } t \in (0,1] \\ 0 & \text{for } t = 0 \end{cases} \tag{1.3.1}$$

Check that the function k is continuous. Let $\mathcal{A} = \{A_i : 1 \le i \le n\}$ be a countable partition of X, where n is a finite integer or ∞. In the sequel we shall usually write ∞.

The *entropy* of $\mathcal{A}$ is the number

$$\mathrm{H}(\mathcal{A}) = \sum_{i=1}^{\infty} -\mu(A_i) \log \mu(A_i) = \sum_{i=1}^{\infty} k(\mu(A_i)) \tag{1.3.2}$$

If $\mathcal{A}$ is infinite, $\mathrm{H}(\mathcal{A})$ may happen to be infinite too.

Define $I(x) = I(\mathcal{A})(x) := -\log \mu(\mathcal{A}(x))$. This is called an information function. Intuitively $I(x)$ is an information on an object x given by the experiment $\mathcal{A}$ in the logarithmic scale. Therefore the entropy in (1.3.2) is the integral (the average) of the information function.

Note that $\mathrm{H}(\mathcal{A}) = 0$ for $\mathcal{A} = \{X\}$ and that if $\mathcal{A}$ is finite, say consists of n elements, then $0 \le \mathrm{H}(\mathcal{A}) \le \log n$ and $\mathrm{H}(\mathcal{A}) = \log n$ if and only if $\mu(A_1) = \mu(A_2) = \ldots = \mu(A_n) = 1/n$. This follows from the fact that the logarithmic function is strictly concave.

In this section we deal with only one fixed measure μ. If however we need to consider more measures simultaneously (see for example Ch.2) we will rather use the notation $\mathrm{H}_\mu(A)$ for $\mathrm{H}(A)$.

Let $\mathcal{A} = \{A_i : i \ge 1\}$ and $\mathcal{B} = \{B_j : j \ge 1\}$ be two countable partitions of X. The *conditional entropy* $\mathrm{H}(\mathcal{A}|\mathcal{B})$ of $\mathcal{A}$ given $\mathcal{B}$ is defined as

$$\begin{aligned} \mathrm{H}(\mathcal{A}|\mathcal{B}) &= \sum_{j=1}^{\infty} \mu(B_j) \sum_{i=1}^{\infty} -\frac{\mu(A_i \cap B_j)}{\mu(B_j)} \log \frac{\mu(A_i \cap B_j)}{\mu(B_j)} \\ &= \sum_{i,j} -\mu(A_i \cap B_j) \log \frac{\mu(A_i \cap B_j)}{\mu(B_j)} \end{aligned} \tag{1.3.3}$$

The first equality, defining $\mathrm{H}(\mathcal{A}|\mathcal{B})$, can be viewed as follows: one considers each element B_j as a probability space with conditional measure $\mu(A|B_j) = \mu(A)/\mu(B_j)$ for $A \subset B_j$ and calculates the entropy of the partition of the set B_j into $A_i \cap B_j$. Then one averages the result over the space of B_j's. (This will be generalized in Def.1.8.3.)

For each x denote $-\log \mu((\mathcal{A}(x) \cap \mathcal{B}(x)|\mathcal{B}(x))$ by $I(x)$ or $I(\mathcal{A}|\mathcal{B})(x)$. The second equality in (1.3.3) can be rewritten as

$$\mathrm{H}(\mathcal{A}|\mathcal{B}) = \int_X I(\mathcal{A}|\mathcal{B})\, d\mu. \tag{1.3.3a}$$

Note by the way that if $\tilde{\mathcal{B}}$ is the σ-algebra consisting of all unions of elements of $\mathcal{B}$ (i.e. generated by $\mathcal{B}$, then $I(x) = -\log \mu((\mathcal{A}(x) \cap \mathcal{B}(x))|\mathcal{B}(x)) = -\log E(1\!\!1_{\mathcal{A}(x)}|\tilde{\mathcal{B}})(x)$, cf (1.1.8).

Note finally that for any countable partition $\mathcal{A}$ we have

$$\mathrm{H}(\mathcal{A}|\{X\}) = \mathrm{H}(\mathcal{A}). \tag{1.3.4}$$

.

Some futher basic properties of entropy of partitions are collected in the following.

Theorem 1.3.1. Let $(X, \mathcal{F}, \mu)$ be a probability space. If $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are countable partitions of X then:

(a) $$\mathrm{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) = \mathrm{H}(\mathcal{A}|\mathcal{C}) + \mathrm{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C})$$

(b) $$\mathrm{H}(\mathcal{A} \vee \mathcal{B}) = \mathrm{H}(\mathcal{A}) + \mathrm{H}(\mathcal{B}|\mathcal{A})$$

(c) $$\mathcal{A} \le \mathcal{B} \;\Rightarrow\; \mathrm{H}(\mathcal{A}|\mathcal{C}) \le \mathrm{H}(\mathcal{B}|\mathcal{C})$$

(d) $$\mathcal{B} \le \mathcal{C} \;\Rightarrow\; \mathrm{H}(\mathcal{A}|\mathcal{B}) \ge \mathrm{H}(\mathcal{A}|\mathcal{C})$$

(e) $$\mathrm{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) \le \mathrm{H}(\mathcal{A}|\mathcal{C}) + \mathrm{H}(\mathcal{B}|\mathcal{C})$$

(f) $$\mathrm{H}(\mathcal{A}|\mathcal{C}) \le \mathrm{H}(\mathcal{A}|\mathcal{B}) + \mathrm{H}(\mathcal{B}|\mathcal{C})$$

Proof. Let $\mathcal{A} = \{A_n : n \ge 1\}$, $\mathcal{B} = \{B_m : m \ge 1\}$, and $\mathcal{C} = \{C_l : l \ge 1\}$. Without loosing generality we can assume that all these sets are of positive measure.

(a) By (1.3.3) we have

$$\mathrm{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) = -\sum_{i,j,k} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(A_i \cap B_j \cap C_k)}{\mu(C_k)}$$

But

$$\frac{\mu(A_i \cap B_j \cap C_k)}{\mu(C_k)} = \frac{\mu(A_i \cap B_j \cap C_k)}{\mu(A_i \cap C_k)} \frac{\mu(A_i \cap C_k)}{\mu(C_k)}$$

unless $\mu(A_i \cap C_k) = 0$. But then the left hand side vanishes and we need not consider it. Therefore

$$\begin{aligned}
\mathrm{H}(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) &= -\sum_{i,j,k} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(A_i \cap C_k)}{\mu(C_k)} \\
&\quad - \sum_{i,j,k} \mu(A_i \cap B_j \cap C_k) \log \frac{\mu(A_i \cap B_j \cap C_k)}{\mu(A_i \cap C_k)} \\
&= -\sum_{i,k} \mu(A_i \cap C_k) \log \frac{\mu(A_i \cap C_k)}{\mu(C_k)} + \mathrm{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C}) \\
&= \mathrm{H}(\mathcal{A}|\mathcal{C}) + \mathrm{H}(\mathcal{B}|\mathcal{A} \vee \mathcal{C})
\end{aligned}$$

(b) Put $\mathcal{C} = \{X\}$ and apply (1.3.4) in (a).

(c) By (a)

$$\mathrm{H}(\mathcal{B}|\mathcal{C}) = \mathrm{H}(\mathcal{A}\vee\mathcal{B}|\mathcal{C}) = \mathrm{H}(\mathcal{A}|\mathcal{C}) + \mathrm{H}(\mathcal{B}|\mathcal{A}\vee\mathcal{C}) \geq \mathrm{H}(\mathcal{A}|\mathcal{C})$$

(d) Since the function k defined by (1.3.1) is strictly concave, we have for every pair i, j

$$k\left(\sum_l \frac{\mu(C_l\cap B_j)}{\mu(B_j)}\frac{\mu(A_i\cap C_l)}{\mu(C_l)}\right) \geq \sum_l \frac{\mu(C_l\cap B_j)}{\mu(B_j)} k\left(\frac{\mu(A_i\cap C_l)}{\mu(C_l)}\right) \tag{1.3.5}$$

But since $\mathcal{B}\leq\mathcal{C}$, we can write above $C_l\cap B_j = C_l$, hence the left hand side equals

$$k\left(\frac{\mu(A_i\cap B_j)}{\mu(B_j)}\right) = -\frac{\mu(A_i\cap B_j)}{\mu(B_j)}\log\frac{\mu(A_i\cap B_j)}{\mu(B_j)}$$

Thus multiplying both sides of (1.3.5) by $\mu(B_j)$ and summing over i and j we get

$$\begin{aligned}-\sum_{i,j}\mu(A_i\cap B_j)\log\frac{\mu(A_i\cap B_j)}{\mu(B_j)} &\geq -\sum_{i,j,l}\mu(C_l\cap B_j)\frac{\mu(A_i\cap C_l)}{\mu(C_l)}\log\frac{\mu(A_i\cap C_l)}{\mu(C_l)}\\ &= -\sum_{i,l}\mu(C_l)\frac{\mu(A_i\cap C_l)}{\mu(C_l)}\log\frac{\mu(A_i\cap C_l)}{\mu(C_l)}\end{aligned}$$

or equivalently $\mathrm{H}(\mathcal{A}|\mathcal{B}) \geq \mathrm{H}(\mathcal{A}|\mathcal{C})$.

Formula (e) follows immediately from (a) and (d) and formula (f) can proved by a straightforward calculation (its consequences are discussed in Exercise 1.9). ♣

§1.4. ENTROPY OF ENDOMORPHISM.

Let $(X, \mathcal{F}, \mu)$ be a probability space and let $T : X \to X$ be a measure preserving endomorphism of X. If $\mathcal{A} = \{A_i\}_{i\in I}$ is a partition of X then by $T^{-1}\mathcal{A}$ we denote the partition $\{T^{-1}(A_i)\}_{i\in I}$. Note that for any countable $\mathcal{A}$

$$\mathrm{H}(T^{-1}\mathcal{A}) = \mathrm{H}(\mathcal{A}) \tag{1.4.1}$$

For all $n \geq m \geq 0$ denote the partition $\bigvee_{i=0}^n T^{-i}\mathcal{A} = \mathcal{A}\vee T^{-1}(\mathcal{A})\vee ...\vee T^{-n}(\mathcal{A}) = \bigvee_{i=m}^n T^{-i}(\mathcal{A})$ by A_m^n. For $m = 0$ we shall sometimes use the notation $\mathcal{A}^n$.

Lemma 1.4.1. For any countable $\mathcal{A}$

$$\mathrm{H}(\mathcal{A}^n) = \mathrm{H}(\mathcal{A}) + \sum_{j=1}^n \mathrm{H}(\mathcal{A}|\mathcal{A}_1^j) \tag{1.4.2}$$

Proof. We prove this formula by induction. If $n = 0$ it is tautology. Suppose it is true for $n - 1 \geq 0$. Then with the use of Theorem 1.3.1(b) and (1.4.1) we obtain

$$\mathrm{H}(\mathcal{A}^n) = \mathrm{H}(\mathcal{A}_1^n \vee \mathcal{A}) = \mathrm{H}(\mathcal{A}_1^n) + \mathrm{H}(\mathcal{A}|\mathcal{A}_1^n) = \mathrm{H}(\mathcal{A}^{n-1}) + \mathrm{H}(\mathcal{A}|\mathcal{A}_1^n) = \mathrm{H}(\mathcal{A}) + \sum_{j=1}^{n} \mathrm{H}(\mathcal{A}|\mathcal{A}_1^j)$$

by the inductive assumption. Hence (1.4.2) holds for all n. ♣

Lemma 1.4.2. The sequences $\frac{1}{n+1}\mathrm{H}(\mathcal{A}^n)$ and $\mathrm{H}\big(\mathcal{A}|\mathcal{A}_1^n\big)$ are monotone decreasing to a limit $\mathrm{h}(T, \mathcal{A})$.

Proof. The sequence $\mathrm{H}\big(\mathcal{A}|\mathcal{A}_1^n\big)$, $n = 0, 1, \ldots$ is monotone decreasing, by Theorem 1.3.1 (d). Therefore the sequence of averages is also monotone decreasing to the same limit, furthermore it coincides with the limit of the sequence $\frac{1}{n+1}\mathrm{H}(\mathcal{A}^n)$ by (1.4.2). ♣

The limit $\frac{1}{n+1}\mathrm{H}(\mathcal{A}^n)$ whose existence has been shown in Lemma 1.4.2. is known as the *(measure–theoretic) entropy of T with respect to the partition* $\mathcal{A}$ and is denoted by $\mathrm{h}(T, \mathcal{A})$ or by $\mathrm{h}_\mu(T, \mathcal{A})$ if one wants to indicate the measure under consideration. Intuitively this means the limit rate of the growth of average (integral) information (in logarithmic scale), under consecutive experiments, for number of experiments tending to infinity.

Remark. To prove the existence of the limit $\frac{1}{n+1}\mathrm{H}(\mathcal{A}^n)$, instead of relying on (1.4.2) and the monotonicity we could use the estmate

$$a_{n+m} = \mathrm{H}(\mathcal{A}^{n+m-1}) \leq \mathrm{H}(\mathcal{A}^{n-1}) + \mathrm{H}(\mathcal{A}_n^{n+m-1}) = a_n + \mathrm{H}(\mathcal{A}^{m-1}) = a_n + a_m.$$

following from Theorem 1.3.1 (e) and from (1.4.1), and apply the following

Lemma 1.4.3. If $\{a_n\}_{n=1}^\infty$ is a sequence of real numbers such that $a_{n+m} \leq a_n + a_m$ for all $n, m \geq 1$ then $\lim_{n\to\infty} a_n$ exists and equals $\inf_n a_n/n$. The limit could be $-\infty$, but if the a_n's are bounded below, then the limit will be nonnegative.

Proof. Fix $m \geq 1$. Each $n \geq 1$ can be expressed as $n = km + i$ with $0 \leq i < m$. Then

$$\frac{a_n}{n} = \frac{a_{i+km}}{i + km} \leq \frac{a_i}{km} + \frac{a_{km}}{km} \leq \frac{a_i}{km} + \frac{ka_m}{km} = \frac{a_i}{km} + \frac{a_m}{m}$$

If $n \to \infty$ then also $k \to \infty$ and therefore $\limsup_{n\to\infty} \frac{a_n}{n} \leq \frac{a_m}{m}$. Thus $\limsup_{n\to\infty} \frac{a_n}{n} \leq \inf \frac{a_m}{m}$. Now the inequality $\inf \frac{a_m}{m} \leq \liminf_{n\to\infty} \frac{a_n}{n}$ finishes the proof. ♣.

Notice that there exists a subadditive sequence (i.e. satisfying $a_{n+m} \leq a_n + a_m$) such that the corresponding sequence a_n/n is not eventually decreasing. Indeed, it suffices to observe that each sequence consisting of 1's and 2's is subadditive and to consider such a sequence having infinitely many 1's and 2's. If for an $n > 1$ we have $a_n = 1$ and $a_{n+1} = 2$ we have $\frac{a_n}{n} < \frac{a_{n+1}}{n+1}$.

Exercise. Prove that Lemma 1.4.1 remains true under the weaker assumptions that there exists $c \in R$ such that $a_{n+m} \le a_n + a_m + c$ for all n and m.

The basic elementary properties of the entropy $\mathrm{h}(T, \mathcal{A})$ are collected in the next theorem below.

Theorem 1.4.4. If $\mathcal{A}$ and $\mathcal{B}$ are countable partitions of finite entropy then

(a) $\quad \mathrm{h}(T, \mathcal{A}) \le \mathrm{H}(\mathcal{A})$

(b) $\quad \mathrm{h}(T, \mathcal{A} \vee \mathcal{B}) \le \mathrm{h}(T, \mathcal{A}) + \mathrm{h}(T, \mathcal{B})$

(c) $\quad \mathcal{A} \le \mathcal{B} \ \Rightarrow \ \mathrm{h}(T, \mathcal{A}) \le \mathrm{h}(T, \mathcal{B})$

(d) $\quad \mathrm{h}(T, \mathcal{A}) \le \mathrm{h}(T, \mathcal{B}) + \mathrm{H}(\mathcal{A}|\mathcal{B})$

(e) $\quad \mathrm{h}(T, T^{-1}(\mathcal{A})) = \mathrm{h}(T, \mathcal{A})$

(f) $\quad$ If $k \ge 1$ then $\mathrm{h}(T, \mathcal{A}) = \mathrm{h}\big(T, \mathcal{A}^k\big)$

(g) $\quad$ If T is invertible and $k \ge 1$ then $\mathrm{h}(T, \mathcal{A}) = \mathrm{h}\big(T, \bigvee_{i=-k}^{k} T^i(\mathcal{A})\big)$

The standard proof (see for example [Wa]) based on Theorem 1.3.1 and formula (1.3.2) is left for the reader as an exercise. Let us prove only (d).

$$\mathrm{h}(T, \mathcal{A}) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}(\mathcal{A}^{n-1}) \le \lim_{n\to\infty} \frac{1}{n}\Big(\mathrm{H}(\mathcal{A}^{n-1}|\mathcal{B}^{n-1}) + \mathrm{H}(\mathcal{B}^{n-1})\Big)$$

$$\le \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} H(T^{-j}(\mathcal{A})|\mathcal{B}^{n-1}) + \lim_{n\to\infty} \frac{1}{n}\mathrm{H}(\mathcal{B}^{n-1})$$

$$\le \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} H(T^{-j}(\mathcal{A})|T^{-j}(\mathcal{B})) + \mathrm{h}(T, \mathcal{B}) \le \mathrm{H}(\mathcal{A}|\mathcal{B}) + \mathrm{h}(T, \mathcal{B}).$$

Here is one more useful fact, stronger than Th.1.4.4 (c):

Theorem 1.4.5. If $T : X \to X$ is a measure preserving endomorphism of a probability space $(X, \mathcal{F}, \mu)$ and $\mathcal{A}$ and $\mathcal{B}_m, m = 1, 2, ...$ are countable partitions of finite entropy, and $\mathrm{H}(\mathcal{A}|\mathcal{B}_m) \to 0$ as $m \to \infty$, then $\mathrm{h}(T, \mathcal{A}) \le \liminf_{m\to\infty} \mathrm{h}(T, \mathcal{B}_m)$. In particular, for $\mathcal{B}_m := \mathcal{B}^m = \bigvee_{j=0}^{m} T^{-j}(\mathcal{B})$, one obtains $\mathrm{h}(T, \mathcal{A}) \le \mathrm{h}(T, \mathcal{B})$.

Proof. By Theorem 1.4.4 (d), for every positive integer m,

$$\mathrm{h}(T, \mathcal{A}) = \mathrm{H}(\mathcal{A}|\mathcal{B}_m) + \mathrm{h}(T, \mathcal{B}_m).$$

Letting $m \to \infty$ this yields the first part of the assertion. For $\mathcal{B}_m = \mathcal{B}^m$, one can substitute in place of the last summand $\mathrm{h}(T, \mathcal{B}^m) = \mathrm{h}(T, \mathcal{B})$, by Theorem 1.4.4(f). ♣

The *(measure-theoretic) entropy of the endomorphism* $T : X \to X$ is defined as

$$\mathrm{h}_\mu(T) = \mathrm{h}(T) = \sup_{\mathcal{A}}\{\mathrm{h}(T, \mathcal{A})\} \tag{1.4.3}$$

where the supremum is taken over all finite (or countable of finite entropy) partitions of X. See Exercise 12.

It is clear from the definition that the entropy of T is an isomorphism invariant.

Later on (see Th.1.8.7, Remark 1.8.7", Corollary 1.8"' and Exercise 1.9') we shall discuss the cases where $\mathrm{H}(\mathcal{A}|\mathcal{B}_n) \to 0$ for every $\mathcal{A}$ (finite or of finite entropy). This will allow us to write $\mathrm{h}_\mu(T) = \lim_{m\to\infty} \mathrm{h}(T, \mathcal{B}_m)$ or $\mathrm{h}(T) = \mathrm{h}(T, \mathcal{B})$.
Let us end this Section with the following useful

Theorem 1.4.6. If $T : X \to X$ is a measure preserving endomorphism of a probability space $(X, \mathcal{F}, \mu)$ then

(a) $\qquad \mathrm{h}(T^k) = k\mathrm{h}(T)$ for all $k \geq 1$

(b) $\qquad$ If T is invertible then $\mathrm{h}(T^{-1}) = \mathrm{h}(T)$

Proof. (a) Fix $k \geq 1$. Since

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{j=0}^{n-1} T^{-kj}\Big(\bigvee_{i=0}^{k-1} T^{-i}\mathcal{A}\Big)\Big) = \lim_{n\to\infty} \frac{k}{nk}\mathrm{H}\Big(\bigvee_{i=0}^{nk-1} T^{-i}\mathcal{A}\Big) = k\mathrm{h}(T, \mathcal{A})$$

we have $\mathrm{h}\big(T^k, \bigvee_{i=0}^{k-1} T^{-i}\mathcal{A}\big) = k\mathrm{h}(T, \mathcal{A})$. Therefore

$$k\mathrm{h}(T) = k \sup_{\mathcal{A}\text{ finite}} \mathrm{h}(T, \mathcal{A}) = \sup_{\mathcal{A}} \mathrm{h}\Big(T^k, \bigvee_{i=0}^{k-1} T^{-i}\mathcal{A}\Big) \leq \sup_{\mathcal{B}} \mathrm{h}(T^k, \mathcal{B}) = \mathrm{h}(T^k) \tag{1.4.4}$$

On the other hand by Theorem 1.4.4(c) we get $\mathrm{h}(T^k, \mathcal{A}) \leq \mathrm{h}\big(T^k, \bigvee_{i=0}^{k-1} T^{-i}\mathcal{A}\big) = k\mathrm{h}(T, \mathcal{A})$ and therefore $\mathrm{h}(T^k) \leq k\mathrm{h}(T)$. The result follows from this and (1.4.4).

(b) In view of (1.4.1) for all finite partitions $\mathcal{A}$ we have

$$\mathrm{H}\Big(\bigvee_{i=0}^{n-1} T^{i}\mathcal{A}\Big) = \mathrm{H}\Big(T^{-(n-1)}\bigvee_{i=0}^{n-1} T^{i}\mathcal{A}\Big) = \mathrm{H}\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{A}\Big)$$

This finishes the proof. ♣

§1.5. SHANNON-MCMILLAN-BREIMAN THEOREM.

Let $(X, \mathcal{F}, \mu)$ be a probability space, $T : X \to X$ be a measure preserving endomorphism of X and $\mathcal{A}$ be a countable finite entropy partition of X.

Lemma 1.5.1. (maximal inequality) For each $n = 1, 2, ...$ let $f_n = I(\mathcal{A}|\mathcal{A}_1^n)$ and $f^* = \sup_{n\geq 1} f_n$. Then for each λ and each $A \in \mathcal{A}$

$$\mu\{x \in A : f^*(x) > \lambda\} \leq e^{-\lambda}.$$

Proof. For each $A \in \mathcal{A}$ and $n = 1, 2, ...$ let $f_n^A = -\log E(1\!\!1_A|\mathcal{A}_1^n)$. Of course $f_n = \sum_{A\in\mathcal{A}} 1\!\!1_A f_n^A$. Denote

$$B_n^A = \{x : f_1^A(x), ..., f_{n-1}^A(x) \leq \lambda, f_n^A(x) > \lambda\}.$$

Since $B_n^A \in \mathcal{F}(\mathcal{A}_1^n)$, the σ-algebra generated by $\mathcal{A}_1^n$,

$$\mu(B_n^A \cap A) = \int_{B_n^A} 1\!\!1_A \, d\mu = \int_{B_n^A} E(1\!\!1_A|\mathcal{A}_1^n) \, d\mu = \int_{B_n^A} e^{-f_n^A} \, d\mu \leq e^{-\lambda}\mu(B_n^A).$$

Therefore

$$\mu(\{x \in A : f^*(x) > \lambda\}) = \sum_{n=1}^{\infty} \mu(B_n^A \cap A) \leq e^{-\lambda} \sum_{n=1}^{\infty} \mu(B_n^A) \leq e^{-\lambda}.$$

♣

Corollary 1.5.2. The function f^* is integrable with integral bounded by $\mathrm{H}(\mathcal{A}) + 1$.

Proof. Of course $\mu\{x \in A : f^* > \lambda\} \leq \mu(A)$, so $\mu(\{x \in A : f^* > \lambda\}) \leq \min\{\mu(A), e^{-\lambda}\}$. So by Lemma 1.5.1

$$\int_X f^* \, d\mu = \sum_{A\in\mathcal{A}} \int_A f^* \, d\mu = \sum_{A\in\mathcal{A}} \int_0^{\infty} \mu\{x \in A : f^* > \lambda\} \, d\lambda$$

$$\leq \sum_{A\in\mathcal{A}} \int_0^{\infty} \min\{\mu(A), e^{-\lambda}\} \, d\lambda = \sum_{A\in\mathcal{A}} \Big(\int_0^{-\log\mu(A)} \mu(A) \, d\lambda + \int_{-\log\mu(A)}^{\infty} e^{-\lambda} \, d\lambda \Big)$$

$$= \sum_{A\in\mathcal{A}} \Big(-\mu(A)(\log\mu(A)) + \mu(A) \Big) = \mathrm{H}(\mathcal{A}) + 1.$$

♣

Corollary 1.5.3. f_n converge a.e. and in L^1.

Proof. $E(1\!\!1_A|\mathcal{A}_1^n)$ is a martingale to which we can apply Theorem 1.1.1. This gives convergence a.e., hence convergence a.e. of each f_n^A, hence f_n. Now convergence in L^1 follows from Corollary 1.5.2. and Dominated Convergence Theorem ♣

Theorem 1.5.4. (Shannon-McMillan-Breiman) Suppose that $\mathcal{A}$ is a countable partition of finite entropy. Then there exist limits

$$f = \lim_{n\to\infty} I(\mathcal{A}|\mathcal{A}_1^n) \text{ and } f_{\mathcal{I}}(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \text{ for a.e. } x$$

and

$$\lim_{n\to\infty} \frac{1}{n+1} I(\mathcal{A}^n) = f_{\mathcal{I}} \text{ a.e. and in } L^1. \tag{1.5.1}$$

Furthermore

$$\mathrm{h}(T,\mathcal{A}) = \lim_{n\to\infty}\frac{1}{n+1}\mathrm{H}(\mathcal{A}^n) = \int f_{\mathcal{I}}\,d\mu = \int f\,d\mu. \tag{1.5.2}$$

The limit f will gain a new interpretation in (1.8.6), in the context of Lebesgue spaces, where the notion of information function I will be generalized.

Proof. First note that $f_n = I(\mathcal{A}|\mathcal{A}_1^n)$ converge to an integrable f by Corollary 1.5.3. (Caution: though integrals of f_n decrease to the entropy, Lemma 1.4.3, it is usually not true that f_n decrease.) Hence the a.e. convergence of time averages to $f_{\mathcal{I}}$ a.e. holds by Birkhoff's Ergodic Theorem. It will suffice to prove (1.5.1) since then (1.5.2), the second equality, holds by integration and the last equality by Birkhoff's Ergodic Theorem, the convergence in L^1.

Let us now establish some identities (compare Lemma 1.4.3). Let $\{\mathcal{A}_n : n \geq 0\}$ be a sequence of countable partitions. Then we have

$$I\left(\bigvee_{i=0}^n \mathcal{A}_i\right) = I\left(\mathcal{A}_0|\bigvee_{i=1}^n \mathcal{A}_i\right) + I\left(\bigvee_{i=1}^n \mathcal{A}_i\right) = I\left(\mathcal{A}_0|\bigvee_{i=1}^n \mathcal{A}_i\right) + I\left(\mathcal{A}_1|\bigvee_{i=2}^n \mathcal{A}_i\right) + .. + I(\mathcal{A}_n).$$

In particular, it follows from the above formula that for $\mathcal{A}_i = T^{-i}\mathcal{A}$, we have

$$\begin{aligned} I(\mathcal{A}^n) &= I(\mathcal{A}|\mathcal{A}_1^n) + I(T^{-1}\mathcal{A}|\mathcal{A}_2^n) + \ldots + I(T^{-n}\mathcal{A}) \\ &= I(\mathcal{A}|\mathcal{A}_1^n) + I(\mathcal{A}|\mathcal{A}_1^{n-1})\circ T + \ldots I(\mathcal{A})\circ T^n \\ &= f_n + f_{n-1}\circ T + f_{n-2}\circ T^2 + \ldots + f_0 \circ T^n, \end{aligned}$$

where $f_k = I(\mathcal{A}|\mathcal{A}_1^k)$, $f_0 = I(\mathcal{A})$. Now

$$\left|\frac{1}{n+1} I(\mathcal{A}^n) - f_{\mathcal{I}}\right| \leq \left|\frac{1}{n+1}\sum_{j=0}^n (f_{n-i}\circ T^i - f\circ T^i)\right| + \left|\frac{1}{n+1}\sum_{j=0}^n f\circ T^i - f_{\mathcal{I}}\right|.$$

Since by Birkhoff's Ergodic Theorem the latter term converges to zero both almost everywhere and in L^1, it suffices to prove that for $n \to \infty$

$$\frac{1}{n+1}\sum_{i=0}^n g_{n-i}\circ T^i \to 0 \text{ a.e. and in } L^1. \tag{1.5.3}$$

where $g_k = |f - f_k|$.

Now, since T is measure preserving, for every $i \geq 0$

$$\int g_{n-i} \circ T^i d\mu = \int g_{n-i} d\mu.$$

Thus $\frac{1}{n}\sum_{i=0}^{n} \int g_{n-i} \circ T^i \, d\mu = \frac{1}{n}\sum_{i=0}^{n} \int g_{n-i} \, d\mu \to 0$, since $f_k \to f$ in L^1 by Corollary 1.5.3. Thus we established the L^1 convergence in (1.5.3).

Now, let $G_N = \sup_{n>N} g_n$. Of course G_N is monotone decreasing and since $g_n \to 0$ a.e. (Corollary 1.5.3) we get $G_N \searrow 0$ a.e.. Moreover, by Corollary 1.5.2, $G_0 \leq \sup_n f_n + f \in L_1$.

For arbitrary $N < n$ we have

$$\begin{aligned}
\frac{1}{n+1}\sum_{i=0}^{n} g_{n-i} \circ T^i &= \frac{1}{n+1}\sum_{i=0}^{n-N-1} g_{n-i} \circ T^i + \frac{1}{n+1}\sum_{i=n-N}^{n} g_{n-i} \circ T^i \\
&\leq \frac{1}{n+1}\sum_{i=0}^{n-N-1} G_N \circ T^i + \frac{1}{n+1}\sum_{i=n-N}^{n} G_0 \circ T^i.
\end{aligned}$$

Hence, for $K_N = G_0 + G_0 \circ T + \ldots + G_0 \circ T^N$

$$\limsup_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} g_{n-i} \circ T^i \leq (G_N)_{\mathcal{I}} + \limsup_{n\to\infty} \frac{1}{n+1} K_N \circ T^{n-N} = (G_N)_{\mathcal{I}} \quad \text{a.e.,}$$

where $(G_N)_{\mathcal{I}} = \lim_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} G_N \circ T^i$ by Birkhoff's Ergodic Theorem.

Now $(G_N)_{\mathcal{I}}$ decreases with N because G_N decreases, and

$$\int (G_N)_{\mathcal{I}} \, d\mu = \int G_N d\mu \to 0$$

because G_N are non-negative uniformly bounded by $G_0 \in L^1$ and tend to 0 a.e..

Hence $(G_N)_{\mathcal{I}} \to 0$ a.e.. Therefore

$$\limsup_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} g_{n-i} \circ T^i \to 0 \quad \text{a.e.}$$

establishing the missing a.e. convergence in (1.5.3). ♣

As an immediate consequence of (1.5.1) and 1.5.2) for T ergodic, along with $f_{\mathcal{I}} = \int f_{\mathcal{I}} \, d\mu$, we get the following:

Theorem 1.5.5 (Shannon-McMillan-Breiman, ergodic case) If $T : X \to X$ is ergodic and $\mathcal{A}$ is a countable partition of finite entropy, then

$$\lim_{n\to\infty} \frac{1}{n} I(\mathcal{A}^{n-1})(x) = \mathrm{h}_\mu(T, \mathcal{A}). \quad \text{for a.e. } x \in X$$

The left hand side can be viewed as a *local entropy* at x. The Theorem says that at a.e. x the local entropy exists and is equal to the entropy (compare comments after (1.3.2) and Lemma 1.4.2).

§1.6. LEBESGUE SPACES, MEASURABLE PARTITIONS AND CANONICAL SYSTEMS OF CONDITIONAL MEASURES.

Let $(X, \mathcal{F}, \mu)$ be a probability space. We consider only *complete* measures (probabilities), namely such that every subset of a set of measure 0 is measurable. If a measure is not complete we can always consider its completion, namely to include in the completion of $\mathcal{F}$ all sets A such that there exists $B \in \mathcal{F}$ with $A \div B$ contained in a set in $\mathcal{F}$ of measure 0. Consider $\mathcal{A}$, an arbitrary partition of X, not necessarily countable nor consisting of measurable sets. By $\tilde{\mathcal{A}}$ we denote the sub σ-algebra of $\mathcal{F}$ consisting of those sets in $\mathcal{F}$ that are unions of whole elements (fibres) of $\mathcal{A}$. Note that $\tilde{\mathcal{A}} \supset \sigma(\mathcal{A})$ defined in Sec.1 (in case $\mathcal{A} \subset \mathcal{F}$) but the inclusion can be strict. Obviously $\tilde{\mathcal{A}} \supset \{\emptyset, X\}$.

Definition 1.6.0. The partition $\mathcal{A}$ is called *measurable* if it satisfies the following separation property.

(1.6.1) There exists a sequence $\mathbf{B}= \{B_n : n \geq 1\} \subset \tilde{\mathcal{A}}$ such that for any two $A_1, A_2 \in \mathcal{A}$ with $A_1 \neq A_2$ there is an integer $n \geq 1$ such that either

$$A_1 \subset B_n \text{ and } A_2 \subset X \setminus B_n$$

or

$$A_2 \subset B_n \text{ and } A_1 \subset X \setminus B_n$$

Since each element of the measurable partition $\mathcal{A}$ can be represented as an intersection of countably many elements B_n or their complements, each element of $\mathcal{A}$ is measurable. Let us stress however that the measurability of all elements of $\mathcal{A}$ is not sufficient for $\mathcal{A}$ to be a measurable partition (see Exercise 1). The sequence $\mathbf{B}$ is called a *basis* for $\mathcal{A}$.

Remark 1.6.0a. A popular definition of an uncountable measurable partition $\mathcal{A}$ is that there exists a sequence of finite partitions (recall that this means: finite partitions into measurable sets) $\mathcal{A}_n, n = 0, 1, ...$, such that $\mathcal{A} = \bigvee_{n=0}^{\infty} \mathcal{A}_n$. Here (unlike later on) the join $\bigvee$ is considered in the set-theoretic sense, i.e. as $\{A_{n_1} \cap A_{n_2} \cap ... : A_{n_i} \in \mathcal{A}_{n_i}, i = 1, ...\}$. Clearly it is equivalent to (1.6.1).

Notice that for any measurable map $T : X \to X'$ between probability measure spaces, if $\mathcal{A}$ is a measurable partition of X', then $T^{-1}(\mathcal{A})$ is a measurable partition of X.

Now we pass to the very useful class of probability spaces: *Lebesgue spaces*.

Definition 1.6.1. We call a sequence $\mathbf{B}= (B_n : n \geq 1) \subset \mathcal{F}$, *basis* of $(X, \mathcal{F}, \mu)$ if the two following conditions are satisfied:

(i) (1.6.1) holds for $\mathcal{A} = \varepsilon$, the partition into points;

(ii) for any $A \in \mathcal{F}$ there exists a set $C \in \sigma(\mathbf{B})$ such that $C \supset A$ and $\mu(C \setminus A) = 0$.

(Recall, Sec.1, that $\sigma(\mathbf{B})$ denotes the smallest σ-algebra containing all $B_n \in\mathbf{B}$. Rohlin used the name Borel σ-algebra.) ,

$(X, \mathcal{F}, \mu)$ satisfying (i) and (ii) for a basis $\mathbf{B}$ is called *separable.*

Now let $\varepsilon = \pm 1$ and $B_n^{(\varepsilon)} = B_n$ if $\varepsilon = 1$ and $B_n^{(\varepsilon)} = X \setminus B_n$ if $\varepsilon = -1$. To any sequence of numbers $\varepsilon_n, n = 1, 2, ...$ there corresponds the intersection $\bigcap_{n=1}^{\infty} B_n^{(\varepsilon_n)}$. By (i) every such intersection contains no more than one point.

The space $(X, \mathcal{F}, \mu)$ is said to be *complete* with respect to a basis $\mathbf{B}$ if all the intersections $\bigcap_{n=1}^{\infty} B_n^{(\varepsilon_n)}$ are non-empty. The space $(X, \mathcal{F}, \mu)$ is said to be *complete* (mod 0) with respect to a basis $\mathbf{B}$ if X can be included as a subset of full measure into a certain measure space $(\overline{X}, \overline{\mathcal{F}}, \overline{\mu})$ which is complete with respect to its own basis $\overline{\mathbf{B}} = (\overline{B}_n)$ satisfying $\overline{\mathrm{B}}_n \cap X = B_n$ for all n.

It turns out that a space which is complete (mod 0) with respect to its one bases is also complete (mod 0) with respect to its every other basis.

Definition 1.6.2. The space $(X, \mathcal{F}, \mu)$ complete (mod 0) with respect to one of its bases is called *Lebesgue space.*

Exercise. If $(X_1, \mathcal{F}_1, \mu_1)$ and $(X_2, \mathcal{F}_2, \mu_2)$ are two probability spaces with complete measures, such that $X_1 \subset X_2, \mu_2(X_2 \setminus X_1) = 0$ and $\mathcal{F}_1 = \mathcal{F}_2|_{X_1}, \mu_1 = \mu_2|_{\mathcal{F}_1}$ (where $\mathcal{F}_2|_{X_1} := \{A \cap X_1 : A \in \mathcal{F}_2\}$), then the first space is Lebesgue iff the second is.

It is not difficult to check that (see Exercise 3) that $(X, \mathcal{F}, \mu)$ is a Lebesgue space if and only if $(X, \mathcal{F}, \mu)$ is isomorphic to the unit interval (equipped with classical Lebesgue measure) together with countably many atoms.

Theorem 1.6.3. Assume that $T : X \to X'$ is a measurable injective map from a Lebesgue space $(X, \mathcal{F}, \mu)$ onto a separable space $(X', \mathcal{F}', \mu')$ and pre-images of the sets of mesure 0 (or positive) are of measure 0 (resp. positive). Then the space $(X', \mathcal{F}', \mu')$ is Lebesgue and T^{-1} is a measurable map.

Remark that in particular a measurable, measure preserving, injective map between Lebesgue spaces is an isomorphism. If $X = X', \mathcal{F} \supset \mathcal{F}', \mathcal{F} \neq \mathcal{F}'$ and $X', \mathcal{F}', \mu'$ is separable, then the above implies that $(X, \mathcal{F}, \mu)$ is not Lebesgue.

Let now $(X, \mathcal{F}, \mu)$ be a Lebesgue space and $\mathcal{A}$ be a measurable partition of X. We say that a property holds for all almost all atoms of $\mathcal{A}$ if and only if the union of atoms for which it is satisfied is measurable, of full measure. The following fundamental theorem holds:

Theorem 1.6.4. For almost all $A \in \mathcal{A}$ there exists a Lebesgue space $(A, \mathcal{F}_A, \mu_A)$ such that the following conditions are satisfied:

(1.6.2) If $B \in \mathcal{F}$, then $B \cap A \in \mathcal{F}_A$ for almost all $A \in \mathcal{A}$.

(1.6.3) The function $X \to [0, 1]$, $x \mapsto \mu_{A(x)}(B \cap A(x))$ is $\mathcal{F}$-measurable for all $B \in \mathcal{F}$, where $A(x)$ is the element of $\mathcal{A}$ containing x.

(1.6.4) $\mu(B) = \int_X \mu_{A(x)}(B \cap A(x))\, d\mu(x)$

Remark 1. One can consider the quotient (factor) space $(X/\mathcal{A}, \mathcal{F}_{\mathcal{A}}, \mu_{\mathcal{A}})$ with $X/\mathcal{A}$ being just $\mathcal{A}$ and with $\mathcal{F}_{\mathcal{A}} = p(\tilde{\mathcal{A}})$ and $\mu_{\mathcal{A}}(B) = \mu(p^{-1}(B))$ for the projection $p(x) = A(x)$. It can be proved that the factor space is again a Lebesgue space. Then $x \mapsto \mu_A(x)(B \cap A(x))$ is $\mathcal{F}_{\mathcal{A}}$-measurable and the property 1.6.4 can be rewritten in the form

$$\mu(B) = \int_{X/\mathcal{A}} \mu_A(B \cap A)\, d\mu_{\mathcal{A}}(A). \tag{1.6.5}$$

Remark 2. If partition $\mathcal{A}$ is finite or countable, then the measures μ_A are just the conditional measures given by the formulas $\mu_A(B) = \mu(A \cap B)/\mu(A)$.

Remark 3. (1.6.4) can be rewritten for every μ-integrable ϕ, or non-negative μ-measurable ϕ if we allow $+\infty$-ies, as

$$\int \phi\, d\mu = \int_X \Big(\int_{A(x)} \phi|_{A(x)}\, d\mu_{A(x)} \Big)\, d\mu(x).$$

This is a version of the Fubini Theorem.

The family of measures $\{\mu_A : A \in \mathcal{A}\}$ is called the *canonical system of conditional measures* with respect to the partition $\mathcal{A}$. It is unique (mod 0) in the sense that any other system μ'_A coincides with it for almost all atoms of $\mathcal{A}$.

The method of construction of the system μ_A is *via* conditional expectations values with respect to the σ-algebra $\tilde{\mathcal{A}}$. Having chosen a basis (B_n) of the Lebesgue space $(X, \mathcal{F}, \mu)$, for every finite intersection

$$B = \bigcap_i B_{n_i}^{(\varepsilon_{n_i})} \tag{1.6.6}$$

one considers $\phi_B := E(1\!\!1_B|\mathcal{A})$, that can be considered as a function on the factor space $X/\mathcal{A}$, unique on a.e. $A \in \mathcal{A}$ such that for all $Z \in \tilde{\mathcal{A}}$

$$\mu(B \cap Z) = \int_{p(Z)} \phi_B(A)\, d\mu_{\mathcal{A}}(A).$$

Clearly $(B_n \cap A)$ is a basis for all A. It is not hard to prove that for a.e. A, for each B from our countable family (1.6.6), $\phi_B(A)$ as a function of B generates Lebesgue space on A. Uniqueness of ϕ_B yields additivity.

Theorem 1.6.5. If $T : X \to X'$ is a measurable map of a Lebesgue space $(X, \mathcal{F}, \mu)$ onto a Lebesgue space $(X', \mathcal{F}', \mu')$, then the induced map from $(X/\zeta, \mathcal{F}_\zeta, \mu_\zeta)$ for $\zeta = T^{-1}(\varepsilon)$, to $(X', \mathcal{F}', \mu')$ is an isomorphism.

Proof. This immediately follows from the fact that the quotient is a Lebesgue space and from Theorem 1.6.2. ♣

In what follows we consider partitions (mod 0), i.e. we identify two partitions if they coincide, restricted to a measurable subset of full measure. For these classes of equivalence we use the same notation $\leq, \geq$ as in Section 4. . They define a partial order. If $\mathcal{A}_\tau$ is a family of measurable partitions of a measure space (unlike in previous Sections the family may be uncountable), then by its product $\mathcal{A} = \bigvee_\tau \mathcal{A}_\tau$ we mean the measurable partition $\mathcal{A}$ defined by

(i) $\mathcal{A} \geq \mathcal{A}_\tau$ for every τ;

(ii) if $\mathcal{A}' \geq \mathcal{A}_\tau$ for every τ and $\mathcal{A}'$ is measurable, then $\mathcal{A}' \geq \mathcal{A}$.

Similarly, replacing $\geq$ by $\leq$, we define the intersection $\bigwedge_\tau \mathcal{A}_\tau$.

The product and intersection exist in a Lebesgue space (i.e. the partially ordered structure is complete). They of course generalize the notions of Section 4. Clearly for a countable family of measurable partitions $\mathcal{A}_\tau$ the above $\bigvee$ and the set-theoretic one coincide (the assumption the space is Lebesgue and the reasoning (mod 0) is not needed). In Exercise 7 we give some examples.

There is a natural one-to-one correspondence between the measurable partitions (mod 0) of a Lebesgue space $(\mathcal{A}, \mathcal{F}, \mu)$ and the complete σ-subalgebras of $\mathcal{F}$, i.e. such σ-algebras $\mathcal{F}' \subset \mathcal{F}$ that the measure μ restricted to $\mathcal{F}'$ is complete. This correspondence is defined by the assignment to each $\mathcal{A}$ the σ-algebra $\mathcal{F}(\mathcal{A})$ of all sets which coincide (mod 0) with the sets of $\tilde{\mathcal{A}}$ (defined at the beginning of this Section). To operations on the measurable partitions (mod 0) correspond operations on the corresponding σ-algebras. Namely , if $\mathcal{A}_\tau$ is a family of measurable partitions (mod 0), then

$$\mathcal{F}(\bigvee_\tau \mathcal{A}_\tau) = \bigvee_\tau \mathcal{F}(\mathcal{A}_\tau), \qquad \mathcal{F}(\bigwedge_\tau \mathcal{A}_\tau) = \bigwedge_\tau \mathcal{F}(\mathcal{A}_\tau).$$

Here $\bigwedge_\tau \mathcal{F}(\mathcal{A}_\tau) = \bigcap_\tau \mathcal{F}(\mathcal{A}_\tau)$ is the set-theoretic intersection of the σ-algebras, while $\bigvee_\tau \mathcal{F}(\mathcal{A}_\tau)$ is the set-theoretic intersection of all the σ-algebras which contain all $\mathcal{F}(\mathcal{A}_\tau)$.

For a monotone increasing (decreasing) sequence of measurable partitions $\mathcal{A}_n$ and $\mathcal{A} = \bigvee_n \mathcal{A}_n$ ($\mathcal{A} = \bigwedge_n \mathcal{A}_n$ respect.) we write $\mathcal{A}_n \nearrow \mathcal{A}$ (or $A_n \searrow \mathcal{A}$). In the language of measurable partitions of a Lebesgue space the Martingale Theorem 1.1.1 can be expressed as follows:

Theorem 1.6.6. If $\mathcal{A}_n \nearrow \mathcal{A}$ or $\mathcal{A}_n \searrow \mathcal{A}$, then for every integrable function f, μ a.s. $E(f|\mathcal{A}_n) \to E(f|\mathcal{A})$, where for $\mathcal{A}$ any measurable partition one writes $E(f|\mathcal{A})(x) := \int f|_{\mathcal{A}(x)}\, d\mu_{\mathcal{A}(x)}$.

Proof. By the definition of canonical system of conditional measures and the definition of conditional expectation value we have for every measurable partition $\mathcal{A}$ the identity $E(f|\mathcal{A}) = E(f|\mathcal{F}(\mathcal{A}))$. ♣

§1.7 ROHLIN NATURAL EXTENSION

We shall prove here following very useful (see Ch.8.9)

Theorem 1.7.1. For every measure preserving endomorphism T of a Lebesgue space $(X, \mathcal{F}, \mu)$ there exists a Lebesgue space $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ with measure preserving transformations $\pi_n : \tilde{X} \to X, n \leq 0$ satisfying $T \circ \pi_{n-1} = \pi_n$, which is an inverse limit of the system $... \xrightarrow{T} X \xrightarrow{T} X$.

Recall that in *category theory* [Lang, Ch.I], for a sequence (system) of objects and morphisms $... \xrightarrow{M_{n-1}} O_n \xrightarrow{M_n} ...... \xrightarrow{M_0} O_0$ an object O equipped with morphisms $\pi_n : O \to O_n$ is called an *inverse limit* if $M_n \circ \pi_{n-1} = \pi_n$ and for every other O' equipped with morphisms $\pi'_n : O' \to O_n$ satisfying $M_n \circ \pi'_{n-1} = \pi'_n$ there exists a unique morphism $M : O \to O$ such that $\pi_n \circ M = \pi'_n$ for every $n \leq 0$. In particular, for $\pi'_n := M_n \circ \pi_{n-1} : O \to O_n$, there exists $M : O \to O$ such that $\pi_n \circ M = \pi'_n = M_n \circ \pi_{n-1}$ for every n. It is easy to see that M is an automorphism.

Here objects are probability spaces or probability spaces with complete probabilities, and morphisms are measure preserving transformations or measure preserving transformations up to sets of measure 0. (We have thus multiple meaning of Theorem 1.7.1.)

Thus Theorem 1.7.1 produces a measure preserving automorphism $\tilde{T} : \tilde{X} \to \tilde{X}$ satisfying

$$\pi_n \circ \tilde{T} = T \circ \pi_{n-1} \tag{1.7.1}$$

for every $n \leq 0$. This automorphism is called Rohlin's *natural extension* of T.

In the proof of the Theorem we shall use the following

Theorem 1.7.2 (On Extension of Measure). Every probability measure ν (σ-additive) on an algebra $\mathcal{G}_0$ of subsets of a set X can be uniquely extended to a measure on the σ-algebra $\mathcal{G}$ generated by $\mathcal{G}_0$

This Theorem can be proved with the use of the famous construction by Carathéodory [Carathéodory, Ch.V], namely by the construction of the outer measure: $\nu_e(A) = \inf \nu(B) : B \in \mathcal{G}_0, A \subset B$ for every $A \subset X$.

We say that A is measurable (in Carathéodory's sense) if for every $E \subset X$ the outer measure ν_e satisfies $\nu_e(E) = \nu_e(E \cap A) + \nu_e(E \setminus A)$. The family of these sets appears to be a σ-algebra containing $\mathcal{G}_0$, hence containing $\mathcal{G}$.

For a general definition of outer measure and sketch of the theory see Ch.6.

Proof of Theorem 1.7.1. We start with producing inverse limit in the set-theoretic category: Consider for $\boldsymbol{Z}^-$, the set of all non-positive integers, the space

$$\tilde{X} = \{(x_n)_{n \in \boldsymbol{Z}^-} : T(x_n) = x_{n+1} \ \forall n < 0\}. \tag{1.7.2}$$

and $\pi_i : \tilde{X} \to X$ the projection to the i-th coordinate, $\pi_i((x_n)_{n\in \boldsymbol{Z}}) = x_i$

Now provide $\tilde{X}$ with a σ-algebra $\tilde{\mathcal{F}}$ and probability measure $\tilde{\mu}$, so that $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ becomes the inverse limit.

Consider $\mathcal{G}_n = \pi_n^{-1}(\mathcal{F})$. Note that this is an increasing sequence of σ-algebras with growing $|n|$ because $\pi^{-1}(A) = \pi_{n-1}^{-1}(T^{-1}(A))$ for every $A \in \mathcal{F}$. Write $\tilde{\mathcal{F}}_0 = \bigcup_{n\le 0} \mathcal{G}_n$. This is an algebra. For every $A \in \mathcal{F}$ and $n \le 0$ define $\tilde{\mu}(\pi_n^{-1}(A)) := \mu(A)$. This is well-defined because if $C = \pi_n^{-1}(A_1) = \pi_m^{-1}(A_2)$ for $A_1, A_2 \in \mathcal{F}$ and $n < m$ then $A_1 = T^{-(m-n)}(A_2)$. Since T preserves μ, we have $\mu(A_1) = \mu(A_2)$.

The next step is to observe that $\tilde{\mu}$ is σ-additive on the algebra $\tilde{\mathcal{F}}_0$. For that we use the assumption $(X, \mathcal{F}, \mu)$ is a Lebesgue space[1]. We just assume that X is a full Lebesgue measure subset of the unit interval, with classical Lebesgue measure and atoms, and the σ-algebra of Lebesgue measurable sets $\mathcal{F}$, see Exercise 3. Now it is sufficient to prove that for every decreasing sequence $C_i \in \tilde{\mathcal{F}}_0, i = 1, 2, ...$ if $\bigcap_i C_i = \emptyset$ then $\tilde{\mu}(C_i) \to 0$. Suppose to the contrary that there exists $\varepsilon > 0$ such that $\tilde{\mu}(C_i) \ge \varepsilon$ for every i. Passing to a subsequence and reindexing we can write $C_n = \pi_{-n}^{-1}(C'_n), n = 1, 2, ...$ We construct compact sets $D'_n \subset C'_n$ such that $\mu(C'_n \setminus D'_n) \le \varepsilon 2^{-(n+1)}$ and $T|_{D'_n}$ is continuous for all n (Lusin's Theorem, [Halmos, Sec.55]).

Write $\Pi = \prod_{-\infty}^{0} X = \{(x_n)_{n\in \boldsymbol{Z}^-} : x_n \in X\}$ for the cartesian product of the countable number of exemplars of X with the product topology (compact by Tichonov's Theorem). Define $\tilde{X}^n := \{(x_i)_{i\in \boldsymbol{Z}^-} : T(x_i) = x_{i+1} \forall n \le i < 0\}$. Of course $\tilde{X} \subset \tilde{X}^n \subset \Pi$. Denote by π^n the projection from $\tilde{X}^n$ to the n-th coordinate.

Then the sets $D_n = \bigcap_{i=1}^{n} (\pi^i)^{-1}(D'_i)$ are compact and decreasing. They are non-empty because $\mu(\pi^n(D_n)) \ge \varepsilon/2$ by the construction, for all n. Therefore $\bigcap_n D_n$ is non-empty (Cantor theorem). Notice finally that $\bigcap_n D_n \subset \bigcap C_n$ so the latter set is non-empty. We have proved that $\tilde{\mu}$ is σ-additive on $\tilde{\mathcal{F}}_0$.

The final step of the construction is the extension of $\tilde{\mu}$ to the σ-algebra $\tilde{\mathcal{F}}$ generated by $\tilde{\mathcal{F}}_0$. It exists (and is unique) due to Theorem 1.7.2.

If we work in the category of complete measures we define the σ-algebra $\tilde{\mathcal{F}}$ as the completion (by subsets of sets of measure 0) of the σ-algebra generated by $\tilde{\mathcal{F}}_0$.

Thus the probability space $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ has been constructed. We leave checking that it is indeed the inverse limit to the reader.

Let us prove that the probability space $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ with completed $\tilde{\mu}$ is Lebesgue space. Let (B_l) be a basis of $(X, \mathcal{F}, \mu)$. Denote by π_n the projection of Π to the n-th coordinate for all n. (We use the same symbol as for projections from $\tilde{X} \subset \Pi$ before. Recall also that projections from intermediate domains have been denoted by π^n.) Then clearly the family $\pi_n^{-1}(B_l)$ is a basis of the partition ε in Π. The restrictions of $\pi_n^{-1}(B_l)$ to $\tilde{X}$ generate the σ-algebra $\tilde{\mathcal{F}}$ on $\tilde{X}$ discussed before (in the sense of Def.1.6.1 (ii)), because (B_l) generates $\mathcal{F}$. We define

$$\tilde{\mu}_{\Pi,n}\Big(\bigcap_{i=-n}^{0} \pi_i^{-1}(C_i)\Big) := \mu\Big(\bigcap_{i=-n}^{0} T^{-(i-n)}(C_i)\Big),$$

[1] this "detail" has been ovelooked in [CFS]

for $C_i = \bigcap_{l=1}^{l(i)} B_l^{(\varepsilon_{l,i})}$ for all $n \le i \le 0$ and ± 1 sequences $\varepsilon_{l,i}$ for $l = 1, 2, ..., l(i)$. It is easy to see that the sequence $\tilde{\mu}_{\Pi,n}$ is compatible on algebras: finite unions of $\bigcap_{i=-n}^{0} \pi_l^{-1}(C_i)$, namely $\tilde{\mu}_{\Pi,n+1}$ extends $\tilde{\mu}_{\Pi,n}$. (One says that this is a *compatible family of of finite-dimensional probability distributions.*)But Π is compact, hence $\tilde{\mu}_{\Pi,n}$ are σ-additive on the union of these algebras, hence extend to a measure (σ-additive) $\tilde{\mu}_\Pi$ on the σ-algebra $\tilde{\mathcal{F}}_\Pi$ generated by them (Kolmogoroff Theorem, see bibliographical notes). The restriction of $\tilde{\mu}_\Pi$ to $\tilde{X}$ coincides with $\tilde{\mu}$ on $\tilde{\mathcal{F}}$ by the uniqueness in Theorem 1.7.2. The restriction of $\tilde{\mathcal{F}}_\Pi$ is a σ-algebra so it contains $\tilde{\mathcal{F}}$. We shall know these σ-algebras coincide if we verify that $\tilde{X}$ is $\tilde{\mu}_\Pi$-measurable, i.e. $\tilde{X} \in \tilde{\mathcal{F}}_\Pi$.

Thus the assertion to be proved is that $\tilde{X} \in \tilde{\mathcal{F}}_\Pi$ and that $\tilde{X}$ is of full measure $\tilde{\mu}_\Pi$. This will prove that $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ is complete (mod 0) with respect to (B_l) restricted to $\tilde{X}$, hence it is Lebesgue space.

Recall that $\tilde{X} = \bigcap_n \tilde{X}^n$ and note that by Lusin Theorem for each n there exist compact sets $D_{n,i} \subset \tilde{X}^n, i = 1, 2, ...$ such that $\tilde{\mu}_\Pi\Big(\Pi \setminus \bigcup_i D_{n,i}\Big) = 0$. Compact sets are measurable as their complementary open sets are countable unions of cylinders. ♣

Remark 1. $\tilde{X}$ can be interpreted as the space of all backward trajectories for T. The map $\tilde{T} : \tilde{X} \to \tilde{X}$ can be defined by the formula

$$\tilde{T}((x_n)_{n \in \boldsymbol{Z}^-}) = (..., x_{-2}, x_{-1}, x_0, T(x_0)). \tag{1.7.3}$$

$\tilde{X}$ could be defined in (1.7.2) as the space of full trajectories $\{(x_n)_{n \in \boldsymbol{Z}}; T(x_n) = x_{n+1}\}$. Then (1.7.3) is the *shift to the left.*

The formula (1.7.3) holds because $\tilde{T}$ defined by it, satisfies (1.7.1), and there holds uniqueness of $\tilde{T}$ satisfying (1.7.1)

Remark 2. Alternatively to Lusin Theorem argument above, we could find for $\tilde{X}^n$ sets $E_{n,i} \supset \tilde{X}^n$, with $\tilde{\mu}_\Pi(E_{n,i} \setminus \tilde{X}^n) \to 0$, which are unions of cylinders $\bigcap_{i=-n}^{0} \pi_i^{-1}(C_i)$. This agrees with the following general fact:

If a sequence of sets Σ generates a σ-algebra $\mathcal{G}$ with a mesure ν on it (see Def.1.6.2 (ii)) then for every $A \in \mathcal{G}$ there exists $C \supset A$ with $\nu(C \setminus A) = 0$ such that $C \in \Sigma'_{d\sigma\delta}$, i.e. C is a countable intersection of countable unions of finite intersections of sets belonging to Σ or their complements. Exercise: Prove this general fact, using Caratheodory's outer measure constructed on measurable sets.

The construction via Lusin theorem presents $\tilde{X}$ as $\Sigma'_{d\sigma\delta\sigma\delta}$ set up to measure 0 (as compact sets are in $\Sigma'_{d\sigma\delta}$). So it is not the most economic.

Remark 3. Another way to prove Theorem 1.7.1 is to construct first $(\Pi, \tilde{F}_\Pi, \tilde{\mu}_{\mathrm{P}})$ on the infinite cartesian product Π, and next $(\tilde{X}, \tilde{F}, \tilde{\mu})$ as the restriction of the first probability space to $\tilde{X}$. We have chosen a different way in order to avoid in the construction the correctness of the definition of $\tilde{\mu}_n$'s in the finite products and the compatibility. We needed it only to prove that the inverse limit is Lebesgue.

We end this section with another version of Theorem 1.7.1. Let us start with

Definition 1.7.3. Suppose that T is an automorphism of a Lebesgue space $(X, \mathcal{F}, \mu)$. Let ζ be a measurable partition. Assume it is forward invariant, namely $T(\zeta) \geq \zeta$, equivalently $T^{-1}(\zeta) \leq \zeta$. Then ζ is said to be *exhausting* if $\bigvee_{n\geq 0} T^n(\zeta) = \varepsilon$.

Theorem 1.7.4. For every measure preserving endomorphism T of a Lebesgue space $(X, \mathcal{F}, \mu)$ there exists a Lebesgue space $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ with an automorphism $\tilde{T}$, with a forward invariant exhausting measurable partition ζ, such that $(X, \mathcal{F}, \mu) = (\tilde{X}/\zeta, \tilde{\mathcal{F}}_\zeta, \tilde{\mu}/\zeta)$ the factor space, cf.Sec.6, Remark 1, and T is factor of $\tilde{T}$, namely $T \circ p = p \circ \tilde{T}$ for the projection $p : \tilde{X} \to X$.

Proof. Take $(\tilde{X}, \tilde{\mathcal{F}}, \tilde{\mu})$ and $\tilde{T}$ from Theorem 1.7.1. Set $\zeta := \pi_0^{-1}(\varepsilon)$. By (1.7.1) and $T^{-1}(\varepsilon) \leq \varepsilon$ we get $\tilde{T}^{-1}(\zeta) \leq \zeta$.

If $\varepsilon' = \bigvee_{n\geq 0} T^n(\zeta)$ is not the partition of $\tilde{X}$ into points, then $\tilde{T}/\varepsilon'$ is an automorphism of $(\tilde{X}/\varepsilon', \tilde{\mathcal{F}}_{\varepsilon'}, \tilde{\mu}_{\varepsilon'})$. Moreover if we denote by p' the projection from $\tilde{X}$ to $\tilde{X}/\varepsilon'$ then we can write $\pi_{-n} = \pi'_{-n} \circ p'$ for some maps π'_{-n} for every $n \geq 0$. By the definition of inverse limit p' must have an inverse which is impossible.

The last part, that $\bigvee_{n\geq 0} T^n(\zeta)$ is the partition of $\tilde{X}$ into points, has also an immediate, not category theory, proof following directly from the form of $\tilde{X}$ in (1.7.2). Indeed for $n \geq 0$ $T^n(\zeta)$ at $\tilde{x} = (..., x_{-2}, x_{-1}, x_0$ is the n-th image of ζ at $\tilde{T}^{-n}(\tilde{x})$ i.e. at $(..., x_{-n-1}, x_{-n})$. So it is equal to $\{(..., x'_{-n-1}, x'_{-n}, ..., x_0) \in \tilde{X} : x_{-n} = x_{-n})\}$. Intersecting for $n \to \infty$ we obtain $\{\tilde{x}\}$.

§1.8 GENERALIZED ENTROPY, CONVERGENCE THEOREMS.

This section contains generalizations of entropy notions introduced in Section 3 to the case of all measurable partitions. The triple $(X, \mathcal{F}, \mu)$ is assumed to be a Lebesgue space.

Definition 1.8.1. If $\mathcal{A}$ is a measurable partition of X then its (generalized) entropy is defined as follows:

$$\mathrm{H}(\mathcal{A}) = \infty \text{ if } \mathcal{A} \text{ is not a countable partition (mod 0);}$$
$$H(\mathcal{A}) = \sum_{A\in\mathcal{A}} -\mu(A) \log \mu(A) \text{ if } \mathcal{A} \text{ is a countable partition (mod 0).}$$

Lemma 1.8.2. If $\mathcal{A}_n$ and $\mathcal{A}$ are measurable partitions of X and $\mathcal{A}_n \nearrow \mathcal{A}$, then $\mathrm{H}(\mathcal{A}_n) \nearrow \mathrm{H}(\mathcal{A})$.

Proof. Write $\mathrm{H}(\mathcal{A}) = \int I(\mathcal{A})\, d\mu$ where $I(\mathcal{A})(x) = -\log \mu(\mathcal{A}(x))$ is the information function (compare Sec.4, we set $\log 0 = -\infty$, hence $I(\mathcal{A})(x) = \infty$ if $\mu(\mathcal{A}(x)) = 0$, here). Write the same for $\mathcal{A}_n$. As $\mu(\mathcal{A}_n(x)) \searrow \mu(\mathcal{A}(x))$ for a.e. x, the convergence in the Lemma follows from Monotone Convergence Theorem. ♣

Definition 1.8.3. If $\mathcal{A}$ and $\mathcal{B}$ are two measurable partitions of X, then the (generalized) *conditional entropy* $\mathrm{H}(\mathcal{A}|\mathcal{B})$ of partition $\mathcal{A}$ subject to $\mathcal{B}$ is defined by the following integral

$$\mathrm{H}_\mu(\mathcal{A}|\mathcal{B}) = \int_{X/\mathcal{B}} \mathrm{H}_{\mu_B}(\mathcal{A}|B)\, d\mu_{\mathcal{B}}(B) \tag{1.8.1}$$

where $\mathcal{A}|B$ is the partition $\{A \cap B : A \in \mathcal{A}\}$ of B and μ_B form a canonical system of conditional measures (Sec. 7). Choose a sequence of finite partitions $\mathcal{A}_n \nearrow \mathcal{A}$ (see Remark 1.6.0). The conditional entropy $\mathrm{H}_{\mu_B}(\mathcal{A}_n|B)$ is measurable as a function of B in the factor space $(X/\mathcal{B}, \mathcal{F}_\mathcal{B}, \mu_\mathcal{B})$, hence of course as a function on $(X, \mathcal{F}, \mu)$, since it is a finite sum of measurable functions $B \mapsto -\mu_B(A \cap B) \log \mu_B(A \cap B)$. Since $\mathcal{A}_n|B \nearrow \mathcal{A}|B$ for a.e. B, we obtain, by using Lemma 1.8.2, that $\mathrm{H}_{\mu_B}(\mathcal{A}_n|B) \to \mathrm{H}_{\mu_B}(\mathcal{A}|B)$. Hence $\mathrm{H}_{\mu_B}(\mathcal{A}|B)$ is measurable, so our definition of $\mathrm{H}_\mu(\mathcal{A}|\mathcal{B})$ makes sense (we allow ∞'s here).

Of course (1.8.1) can be also written in the form

$$\int_X \mathrm{H}_{\mu_{\mathcal{B}(x)}}(\mathcal{A}|\mathcal{B}(x))\, d\mu(x), \tag{1.8.2}$$

with $\mathrm{H}_{\mu_B}(\mathcal{A}|B)$ understood as constant function on each B (compare (1.6.4) versus (1.6.5)).

As in Sec.3 we can write

$$\mathrm{H}_\mu(\mathcal{A}|\mathcal{B}) = \int_X I(\mathcal{A}|\mathcal{B})\, d\mu, \tag{1.8.3}$$

where $I(\mathcal{A}|\mathcal{B})$ is the *conditional information function*:
$I(\mathcal{A}|\mathcal{B})(x) := -\log \mu_{\mathcal{B}(x)}(\mathcal{A}(x) \cap \mathcal{B}(x))$.

Indeed $I(\mathcal{A}|\mathcal{B})$ is non-negative and μ-measurable as $\lim_{n\to\infty} I(\mathcal{A}_n|\mathcal{B})$ (a.e.), so (1.8.3) follows from (1.6.5a).

Lemma 1.8.4. If $\{\mathcal{A}_n : n \geq 1\}$ and $\mathcal{A}$ are measurable partitions, $\mathcal{A}_n \searrow \mathcal{A}$ and $\mathrm{H}(\mathcal{A}_1) < \infty$ then $\mathrm{H}(\mathcal{A}_n) \searrow \mathrm{H}(\mathcal{A})$.

Proof. The proof is similar to Proof of Lemma 1.8.2. ♣

Theorem 1.8.5. If $\mathcal{A}, \mathcal{B}$ are measurable partitions and $\{\mathcal{A}_n : n \geq 1\}$ is an increasing (decreasing and $\mathrm{H}(\mathcal{A}_1|\mathcal{B}) < \infty$) sequence of measurable partitions converging to $\mathcal{A}$, then

$$\lim_{n\to\infty} \mathrm{H}(\mathcal{A}_n|\mathcal{B}) = \mathrm{H}(\mathcal{A}|\mathcal{B}) \tag{1.8.4}$$

and the convergence is respectively monotone.

Proof. Applying Lemmas 1.8.2 and 1.8.4 we get the monotone convergence $\mathrm{H}_{\mu_B}(\mathcal{A}_n|B) \to \mathrm{H}_{\mu_B}(\mathcal{A}|B)$ for almost all $B \in X/\mathcal{B}$. Thus the integrals in the Definition 1.8.3 converge by the Monotone Convergence Theorem. ♣

Theorem 1.8.6. If $\mathcal{A}, \mathcal{B}$ are measurable partitions and $\{\mathcal{B}_n : n \geq 1\}$ is a decreasing (increasing and $\mathrm{H}(\mathcal{A}|\mathcal{B}_1) < \infty$) sequence of measurable partitions converging to $\mathcal{B}$, then

$$\lim_{n\to\infty} \mathrm{H}(\mathcal{A}|\mathcal{B}_n) = \mathrm{H}(\mathcal{A}|\mathcal{B}) \tag{1.8.5}$$

and the convergence is respectively monotone.

Proof 1. Assume first that $\mathcal{A}$ is finite (or countable with finite entropy). Then the a.e. convergence $I(\mathcal{A}|\mathcal{B}_n) \to I(\mathcal{A}|\mathcal{B})$ follows from Martingale Convergence Theorem (more precisely from Theorem 1.6.6), applied to $f = 1\!\!1_A$, the indicator function for each $A \in \mathcal{A}$.

Now it is sufficient to prove $\sup_n I(\mathcal{A}|\mathcal{B}_n) \in L^1$ in order to use Dominated Convergence Theorem (compare Corollary 1.5.3) and (1.8.3). One can repeat Proofs of Lemma 1.5.1 (for increasing $\mathcal{B}_n$) and Corollary 1.5.2.

The monotonicity of the sequence $\mathrm{H}(\mathcal{A}|\mathcal{B}_n)$ relies on Theorem 1.3.d. However for infinite $\mathcal{B}_n$ one needs to approximate $\mathcal{B}_n$ by finite (or finite entropy) partitions. For details see [Rohlin 1967, Sec.5.12].

For $\mathcal{A}$ measurable, represent $\mathcal{A}$ as $\lim_{j\to\infty} \mathcal{A}_j$ for an increasing sequence of finite partitions $\mathcal{A}_j, j = 1, 2, ..$, next refer to Th.1.8.5. In the case of decreasing $\mathcal{B}_n$ the proof is straightforward. In the case of increasing $\mathcal{B}_n$ use

$$\mathrm{H}(\mathcal{A}|\mathcal{B}_n) - \mathrm{H}(\mathcal{A}_j|\mathcal{B}_n) = H(\mathcal{A}|(\mathcal{A}_j \vee \mathcal{B}_n)) \le H(\mathcal{A}|(\mathcal{A}_j \vee \mathcal{B}_1)) \le \mathrm{H}(\mathcal{A}|\mathcal{B}_1) - \mathrm{H}(\mathcal{A}_j|\mathcal{B}_1).$$

This implies that the convergence as $j \to \infty$ is uniform with respect to n, hence in the limit $\mathrm{H}(\mathcal{A}|\mathcal{B}_n) \to H(\mathcal{A}|\mathcal{B})$. ♣

Proof 2. For $\mathcal{A}$ finite (or countable with finite entropy) there is a simpler way to prove (1.8.5). We have for every $A \in \mathcal{A}$ by Theorem 1.1.1, the convergence in L^2 applied to $E(1\!\!1_A|\mathcal{F}(\mathcal{B}_n))$, hence the convergence in measure μ of $\mu_{\mathcal{B}_n(x)}(A \cap \mathcal{B}_n(x))$. By the continuity of the function $k(t) = -t\log t$, see Sec.3, this implies the convergence in measure μ

$$k(\mu_{\mathcal{B}_n(x)}(A \cap \mathcal{B}_n(x))) \to k(\mu_{\mathcal{B}(x)}(A \cap \mathcal{B}(x))).$$

(We do not assume $x \in A$ here.) Summing over $A \in \mathcal{A}$ we obtain the convergence $\mathrm{H}_{\mu_{\mathcal{B}_n(x)}}(\mathcal{A}|\mathcal{B}_n(x)) \to \mathrm{H}_{\mu_{\mathcal{B}(x)}}(\mathcal{A}|\mathcal{B}(x))$ in measure μ. These functions are uniformly bounded by $\log \#\mathcal{A}$ (or by $\mathrm{H}(\mathcal{A})$) and non-negative, hence we get the convergence in L^1 and in consequence, due to (1.8.2), we obtain (1.8.5). (Note that we have not used the a.e. convergence in Th.1.1.1, but only the convergence in L^2 proved there.) ♣

Observe that we can rewrite now the definition of the entropy $\mathrm{h}_\mu(T, \mathcal{A})$ from Section 1.5 as

$$\mathrm{h}_\mu(T, \mathcal{A}) = \mathrm{H}(\mathcal{A}|\mathcal{A}^-), \quad \text{where } \mathcal{A}^- := \bigvee_{n=1}^{\infty} T^{-n}(\mathcal{A}). \tag{1.8.6}$$

A countable partition $\mathcal{B}$ is called *a countable generator* for an endomorphism of a Lebesgue space if $\mathcal{B}^m \nearrow \varepsilon$. Due to Theorem 1.8.6 we obtain the following facts useful in computing the entropy for concrete examples.

Theorem 1.8.7. (a) If $\mathcal{B}_m$ is a sequence of finite partitions of a Lebesgue space, such that $\mathcal{B}_m \nearrow \varepsilon$, then, for any endomorphism T of the space, $\mathrm{h}(T) = \lim_{m\to\infty} \mathrm{h}(T, \mathcal{B}_m)$.

(b) If $\mathcal{B}$ is a countable generator of finite entropy for an endomorphism T of a Lebesgue space, then $\mathrm{h}(T) = \mathrm{h}(T, \mathcal{B})$.

Proof. By Theorem 1.8.6 for every finite $\mathcal{A}$ we have $\lim_{m\to\infty} \mathrm{H}(\mathcal{A}|\mathcal{B}_m) = \mathrm{H}(\mathcal{A}|\varepsilon) = 0$, hence in view of Theorem 1.4.5, instead of $\sup_{\mathcal{A}}$ in the definition of $\mathrm{h}(T)$, it is sufficient in (1.4.3) to consider $\lim_{m\to\infty} \mathrm{h}(T,\mathcal{B}_m)$. This proves (a). Theorem 1.4.5 together with the definition of the generator prove also (b). ♣

Remark 1.8.7'. For T an automorphism one considers two-sided countable (in particular finite) generator: $\bigvee_{n=-\infty}^{\infty} T^n(\mathcal{B}) = \varepsilon$. Then, as in the one-sided case, $\mathrm{H}(\mathcal{B})$ finite implies $\mathrm{h}(T) = \mathrm{h}(T,\mathcal{B})$.

Remark 1.8.7". In both Theorem 1.8.6 and Theorem 1.8.7(a) the assumption on the monotonicity of $\mathcal{B}_m$ can be weakened. Assume for example that $\mathcal{A}$ is finite and $\mathcal{B}_m \to \varepsilon$ in the sense that for every measurable Y, $E(1\!\!1_Y|\mathcal{B}_m) \to 1\!\!1_Y$ in measure, as in Remark 1.1.2. Then $\mathrm{H}(\mathcal{A}|\mathcal{B}_m) \to 0$, hence $\mathrm{h}(T) = \lim_{m\to\infty} \mathrm{H}(T,\mathcal{B}_m)$.
Indeed for $\mathrm{H}(\mathcal{A}|\mathcal{B}_m) \to 0$ just repeat Proof 2 of Theorem 1.8.6. The convergence in measure μ of $\mu_{\mathcal{B}_n(x)}(A\cap\mathcal{B}_n(x)))$ to $\mu_{\varepsilon(x)}(A\cap\varepsilon(x)))$ writes as $E(1\!\!1_A|\mathcal{B}_n) \to 1\!\!1_A$, which has just been assumed.

Corollary 1.8.7"'. If X is a compact metric space and $\mathcal{F}$ the σ-algebra of Borel sets (generated by open sets), then if $\sup_{B\in\mathcal{B}_m}\big(\mathrm{diam}(B)\big) \to 0$ as $m\to\infty$, then $\mathrm{h}(T) = \lim_{m\to\infty} \mathrm{H}(T,\mathcal{B}_m)$.
Proof. It is sufficient to check $E(1\!\!1_A|\mathcal{B}_m) \to 1\!\!1_A$ in measure. First note that for every $\delta > 0$ there exist an open set U and closed set K such that $K\subset A\subset U$ and $\mu(U\setminus K)\le\delta$. This property is called *regularity* of our measure μ and is true for every finite measure on the σ-algebra of Borel sets for a metric space (compactness is not needed here). It can be proved by Caratheodory's argument, compare Proof of Th.1.1.1. Namely we construct the outer measure with the help of open sets, as in the sketch of the proof of theorem 1.7.2 (where we used $\mathcal{G}_0$) and notice that since each closed set is an intersection of a decreasing sequence of open sets we will have the same outer measure if in the construction of outer measure we use the algebra generated by open sets. Now we can refer to Theorem 1.7.2.
Next, due to compactness of X, hence K, for m large enough the set $A' := \bigcup\{B\in\mathcal{B}_m : B\cap K\ne\emptyset\}$ contains K and is contained in U, hence $\mu(A\div A')\le\delta$. This implies that

$$\int_X |E(1\!\!1_A|\mathcal{B}_m) - 1\!\!1_A|\,d\mu =$$

$$\int_{X\setminus(A\cup A')} E(1\!\!1_A|\mathcal{B}_m)\,d\mu + \int_{A\div A'} |E(1\!\!1_A|\mathcal{B}_m) - 1\!\!1_A|\,d\mu + \int_{A\cap A'} 1 - E(1\!\!1_A|\mathcal{B}_m)\,d\mu \le$$

$$\frac{\delta}{\mu(X\setminus A')}\mu(X\setminus(A\cup A')) + \delta + \Big(1 - \frac{\mu(A\cap A')}{\mu(A')}\Big)\mu(A\cap A') \le 3\delta.$$

Hence $\mu\{x : |E(1\!\!1_A|\mathcal{B}_m) - 1\!\!1_A| \ge \sqrt{3\delta}\} \le \sqrt{3\delta}$. ♣

For a simpler proof, omitting Theorem 1.8.6, see Exercise 1.9'.

We end this Section with the *theorem on decomposition into ergodic components* and the adequate entropy formula. Compare this with Choquet representation theorem: Th. 2.1.11, and Th. 2.1.13.

Let T be a measure preserving endomorphism of a Lebesgue space. A measurable partition $\mathcal{A}$ is said to be T-invariant if $T(A) \subset A$ for almost every $A \in \mathcal{A}$. The induced map $T_A = T|_A : A \to A$ is a measurable endomorphism of the Lebesgue space $(A, \mathcal{F}_{\mathcal{A}}, \mu_A)$. One calls T_A a *component* of T.

Theorem 1.8.8. (a) There exists a smallest T-invariant measurable partition $\mathcal{A}$ (mod 0) (called the *ergodic decomposition*). Almost all of its components are ergodic.

(b) $h(T) = \int_{X/\mathcal{A}} h(T_A)\, d\mu_{\mathcal{A}}(A)$.

Proof. We shall not prove here the part (a). Let us mention only that the ergodic decomposition partition corresponds (see Sec.6) to the completion of $\mathcal{I}$, the σ-subalgebra of $\mathcal{F}$ consisting of T invariant sets in $\mathcal{F}$ (compare Theorem 1.2.2).

To prove the part (b) notice that for every T-invariant measurable partition $\mathcal{A}$, for every finite partition ξ and almost every $A \in \mathcal{A}$, writing ξ_A for the partition $\{s \cap A : s \in \xi\}$, we obtain

$$\mathrm{h}(T_A, \xi_A) = H(\xi_A | \xi_A^-) = \int_A I_{\mu_A}(\xi_A | \xi_A^-)\, d\mu_A.$$

Notice next that the latter information function is equal a.e. to $I_\mu(\xi|\xi^- \vee \mathcal{A})$ restricted to A. Hence

$$\int_{X/\mathcal{A}} h(T_A)\, d\mu_{\mathcal{A}}(A) = \int_{X/\mathcal{A}} d\mu_{\mathcal{A}} \int_A I_{\mu_A}(\xi_A|\xi_A^-)\, d\mu_A =$$

$$\int_X I_\mu(\xi|\xi^- \vee \mathcal{A})\, d\mu = \mathrm{H}(\xi|\xi^- \vee \mathcal{A}) = \mathrm{h}(T, \xi)$$

The latter equality follows from an approximation of $\mathcal{A}$ by finite T-invariant partitions $\eta \nearrow \mathcal{A}$ and from

$$\mathrm{H}(\xi|\xi^- \vee \eta) = \mathrm{H}(\xi \vee \eta | \xi^- \vee \eta^-) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}((\xi \vee \eta)^n) =$$

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{H}(\xi^n \vee \eta) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}(\xi^n) = H(T, \xi).$$

Let now ξ_n be a sequence of finite partitions such that $\xi_n \nearrow \varepsilon$. Then $\mathrm{h}(T, \xi_n) \nearrow \mathrm{h}(T)$ and $\mathrm{h}(T_A, (\xi_n)_A) \nearrow \mathrm{h}(T_A)$. So $\mathrm{h}(T, \xi_n) = \int_{X/\mathcal{A}} h(T_A, \xi_n)\, d\mu_{\mathcal{A}}(A)$ and Lebesgue monotone convergence theorem prove (b) ♣

§1.9 COUNTABLE TO ONE MAPS, JACOBIAN AND ENTROPY OF ENDOMORPHISMS .

We start with a formulation of a deep theorem by Rohlin:

Theorem 1.9.1. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are two measurable partitions of a Lebesgue space $(X, \mathcal{F}, \mu)$ such that $\mathcal{A}|_B$ is countable (mod 0 with respect to μ_B) for almost every $B \in \mathcal{B}$. Then there exists a countable partition $\gamma = \{\gamma_1, \gamma_2, ...\}$ of X (mod 0) such that such that each $\gamma_j \in \gamma$ intersects almost every B at not more than one point, which is then an atom of μ_B, in particular

$$\mathcal{A} \vee \mathcal{B} = \gamma \vee \mathcal{B} \pmod 0.$$

Furthermore, if $\mathrm{H}(\mathcal{A}|\mathcal{B}) < \infty$, then γ can be chosen so that

$$\mathrm{H}(\gamma) < H(\mathcal{A}|\mathcal{B}) + 3\sqrt{H(\mathcal{A}|\mathcal{B})} < \infty.$$

Definition 1.9.2. Let $(X, \mathcal{F}, \mu)$ be a Lebesgue space. Let $T : X \to X$ be a measurable endomorphism. We say that T is *essentially countable to one* if the measures μ_A of a canonical system of conditional measures for the partition $\mathcal{A} := T^{-1}(\varepsilon)$ are purely atomic (mod 0 with respect to μ_A), for almost all A. We say that T is *countable to one* if we can omit the phrase "mod 0 with respect to μ_A" above.

Lemma 1.9.3. If T is essentially countable to one and preserves μ then there exists a measurable $Y \subset X$ of full measure such that $T(Y) \subset Y$ and

1. $T^{-1}(x) \cap Y$ for a.e. $x \in Y$ is countable, moreover it consists only of atoms of the conditional measure $\mu_{T^{-1}(x)}$;
2. $T(B)$ is measurable if $B \subset Y$ is measurable;
3. $T|_Y$ is forward quasi-invariant, namely $\mu(B) = 0$ for $B \subset Y$ implies $\mu(T(B)) = 0$.

Proof. Let Y' be the union of atoms mentioned in Definition 1.9.2.. We can write, due to Theorem 1.9.1, $Y' = \bigcup_j \gamma_j$, so Y' is measurable. Set $Y = \bigcap_{n=0}^{\infty} T^{-n}(Y')$. Denote the partition $T^{-1}(\varepsilon)$ in Y by ζ. Property 1. follows from the construction. To prove 2. we use the fact that $(Y/\zeta, \mathcal{F}_\zeta, \mu_\zeta)$ is a Lebesgue space and the factor map $T_\zeta : Y_\zeta \to X$ is an automorphism (Th.1.6.5). So, for measurable $B \subset Y$, the set

$$\{A \in \zeta : \mu_A(B \cap A) \neq 0\} = \{A \in \zeta : A \cap B \neq \emptyset\} \tag{1.9.1}$$

is measurable by (1.6.3) and therefore its image under T_ζ, equal to $T(B)$, is measurable. If $\mu(B) = 0$, then the set in (1.9.1) has measure μ_ζ equal to 0, hence as T_ζ is isomorphism we obtain that $T(B)$ is measurable, of measure 0. ♣

The key property in the above proof is the equality (1.9.1). Without assuming that μ_A are purely atomic there could existed B of measure 0 with $C := \{A \in \zeta : \mu_A(B \cap A) \neq 0\}$ not measurable in $\mathcal{F}_\zeta$.

To have such a situation just consider a non-measurable $C \subset Y/\zeta$. Consider the disjoint union $D := C \cup Y$ and denote the embedded C by C'. Finally, defining measure on D, put $\mu(C') = 0$ and μ on the embedded Y. Define $T(c') = T(C)$ for $C \ni c$ and c' being the image of c under the abovementioned embedding. Thus C' is measurable, of measure 0, whereas $T(C')$ is not measurable because C is not measurable and T_ζ is isomorphism.

Definition 1.9.4. Let $(X, \mathcal{F}, \mu)$ and $(X', \mathcal{F}', \mu')$ be probability measure spaces. Let $T : X \to X'$ be a measurable homomorphism. We say that a real, nonnegative, measurable function J is a *weak Jacobian* if there exists E of measure 0 such that for every measurable $A \subset X \setminus E$ on which T is injective, the set $T(A)$ is measurable and $\mu(T(A)) = \int_A J\,d\mu$. We say J is *strong Jacobian* if the above holds without assuming $A \subset X \setminus E$.

Notice that if T is forward quasi-invariant, namely $(\mu(A) = 0) \Rightarrow (\mu'(T(A)) = 0)$, then automatically weak Jacobian is strong Jacobian.

Proposition 1.9.5. Let $(X, \mathcal{F}, \mu)$ be Lebesgue space and $T : X \to X$ be a measurable, essentially countable to one, endomorphism. Then there exists a weak Jacobian J. It is unique (mod 0). For T restricted to Y (from Lemma 1.9.3.) J is strong Jacobian.

Proof. Consider the partition $\gamma = \{\gamma_1, \gamma_2, ...\}$ given by Theorem 1.9.1. Then for each j the map $T|_{\gamma_j \cap Y}$ is injective. Moreover by Lemma 1.9.3 $T|_{\gamma_j \cap Y}$ maps measurable sets onto measurable sets and is forward quasi-invariant. Therefore J exists on each $\gamma_j \cap Y$ by Radon-Nikodym theorem.

By the presentation of each $A \subset Y$ as $\bigcup_{j=1}^\infty A \cap \gamma_j$ the function J satisfies the assertion of the Proposition. The uniqueness follows from the uniqueness of Jacobian in Radon-Nikodym theorem on each $\gamma_j \cap Y$.

Theorem 1.9.6. Let $(X, \mathcal{F}, \nu)$ be a Lebesgue space. Let $T : X \to X$ be a ν preserving endomorphism, essentially countable to one. Then its Jacobian, strong on Y defined in Lemma 1.9.3, weak on X, has logarithm equal to $I(\varepsilon | T^{-1}(\varepsilon))$. ($I$ stands for the information function, see Sections 1.4 and 1.8)

Proof. Consider already T restricted to Y. Let $Z \subset Y$ be an arbitrary measurable set such that T is 1–to–1 on it. For each $y \in Y$ denote by $A(y)$ the element of $\zeta = T^{-1}(\varepsilon)$ containing y. We obtain

$$
\begin{aligned}
\nu\big(T(Z)\big) = \nu\Big(T^{-1}\big(T(Z)\big)\Big) &= \int_{T^{-1}\big(T(Z)\big)} \left(\int_{A(y)} 1\!\!1 \, d\nu_{A(y)} \right) d\nu(y) = \\
&= \int_{T^{-1}\big(T(Z)\big)} \left(\int_{A(y)} 1\!\!1_Z(y)/\nu_{A(y)}\{y\}) \, d\nu_{A(y)} \right) d\nu(y) = \\
&= \int_{T^{-1}\big(T(Z)\big)} 1\!\!1_Z(y)/\nu_{A(y)}\{y\}) \, d\nu(y) = \int_Z 1/\nu_{A(y)}\{y\}) \, d\nu(y)
\end{aligned}
$$

♣

Theorem 1.9.6 gives rise to the so called Rohlin formula:

Theorem 1.9.7. Let $(X, \mathcal{F}, \mu)$ be a Lebesgue space. Let $T : X \to X$ be a μ preserving endomorphism, essentially countable to one. Suppose that on each component A of the ergodic decomposition (cf. Th.1.8.8) the restriction T_A has a countable generator of finite entropy. Then for the Jacobian J

$$h_\mu(T) = \mathrm{H}(\varepsilon|T^{-1}(\varepsilon)) = \int I(\varepsilon|T^{-1}(\varepsilon))\, d\mu = \int \log J\, d\mu.$$

Proof. The third equality follows from Theorem 1.9.6, the second one is the definition of the conditional entropy, see Sec. 8. To prove the first equality we can assume, due to Theorem 1.8.8, that T is ergodic. Then, for ζ, a countable generator of finite entropy, with the use of Theorems 1.8.5 and 1.8.6, we obtain

$$\mathrm{H}(\varepsilon|T^{-1}(\varepsilon) = \mathrm{H}(\varepsilon|\zeta^-) = \lim_{n\to\infty} \mathrm{H}(\zeta^n|\zeta^-) = \mathrm{H}(\zeta|\zeta^-) = \mathrm{h}(T, \zeta) = \mathrm{h}(T).\clubsuit$$

Remark. The existence of a countable generator is a general, not very difficult, fact, namely the following holds:

Theorem 1.9.8. Let $(X, \mathcal{F}, \mu)$ be Lebesgue space. Let $T : X \to X$ be a μ-preserving aperiodic endomorphism, essentially countable to one. Then there exists a *countable generator*, namely a countable partition ζ such that $\zeta^- = \varepsilon$ (mod 0).

Aperiodic means there exists no B of positive measure and a positive integer n so that $T^n|_B =$id. For the proof see [Rohlin, 1967, Sec.10.12-13] or [Parry]. To construct ζ one uses the partition γ found for ε and $T^{-1}(\varepsilon)$ according to Theorem 1.9.1 and so-called *Rohlin towers*.
The existence of a generator with finite entropy is in fact equivalent to $H(\varepsilon|\varepsilon^-) = \mathrm{h}(T) < \infty$. The proof of the implication to the right is contained in Proof of Th.1.9.7. The reverse implication, the construction of the partition, is not easy, it uses in particular the estimate in Th.1.9.1.
The existence of a generator with finite entropy is a strong property. It may fail even for *exact* endomorphisms, see Sec.10 and Exercise 13. Neither its existence implies exactness, Exercise 13. To the contrary, for automorphisms, two-sided generators, even finite, always exist, provided the map is aperiodic.

§1.10. MIXING PROPERTIES.

In this section we examine briefly some mixing properties of a measure preserving endomorphism, stronger than ergodicity. A measure preserving endomorphism is said to be *weakly mixing* if and only if for every two measurable sets A and B

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\mu(T^{-j}(B) \cap A) - \mu(A)\mu(B)| = 0$$

To see that a weakly mixing transformation is ergodic, suppose that $T^{-1}(B) = B$. Then $T^{-k}(B) = B$ for all $k \geq 0$ and consequently for every n, $\frac{1}{n}\sum_{j=0}^{n-1} |\mu(T^{-j}(B) \cap A) - \mu(A)\mu(B)| = |\mu(B) - \mu(B)^2|$. Thus $\mu(B) - \mu(B)^2 = 0$ and therefore $\mu(B) = 0$ or 1.

A measure preserving endomorphism is said to be *mixing* if and only if for every two measurable sets A and B

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) - \mu(A)\mu(B) = 0$$

Clearly, every mixing transformation is weakly mixing. The property equivalent to the mixing property is the following: for every square integrable functions f, g

$$\lim_{n\to\infty} \int f(g \circ T^n)\, d\mu = \int f\, d\mu \int g\, d\mu.$$

Indeed, the former property follows from the latter one if we substitute the indicator functions $1\!\!1_A, 1\!\!1_B$ in place of f, g. To prove the opposite implication notice that with the help of Hölder inequality it is sufficient to restrict to simple functions $f = \sum_i a_i 1\!\!1_{A_i}, g = \sum_j a_j 1\!\!1_{A_j}$ for finite partitions (A_i) and (B_j). Then

$$\left| \int f(g \circ T^n)\, d\mu - \int f\, d\mu \int g\, d\mu \right| = \left| \sum_{i,j} a_i b_j (\mu(A_i \cap T^{-n}(B_j)) - \mu(A_i)\mu(B_j)) \right| \to 0$$

because every summand converges to 0 as $n \to \infty$.

In the sequel we will deal also with stronger mixing properties. An endomorphism is called *K-mixing* if for every measurable set A and every finite partition $\mathcal{A}$

$$\lim_{n\to\infty} \sup_{B \in \mathcal{F}(\mathcal{A}_n^\infty)} |\mu(A \cap B) - \mu(A)\mu(B)| = 0,$$

Recall that $\mathcal{F}(\mathcal{A}_n^\infty)$ for $n \geq 0$ means the complete σ-algebra assigned to the partition $\mathcal{A}_n^\infty = \bigvee_{j=n}^\infty T^{-j}(\mathcal{A})$. The following theorem provides us with alternative definitions of the K-mixing property in case T is an automorphism.

Theorem 1.10.1. If $T : X \to X$ is a measure-preserving automorphism of a Lebesgue space, then the following conditions are equivalent:

(a) T is K-mixing.

(b) For every finite partition $\mathcal{A}$ $\mathrm{Tail}(\mathcal{A}) := \bigwedge_{n=0}^\infty \bigvee_{k=n}^\infty T^{-k}(\mathcal{A})$ is equal to the trivial partition $\nu = \{X\}$

(c) For every finite partition $\mathcal{A} \neq \nu$, $\mathrm{h}_\mu(T, \mathcal{A}) > 0$ (T has *completely positive entropy*)

(d) There exists a forward invariant *exhausting* measurable partition α (i.e. satisfying $T^{-1}(\alpha) \leq \alpha$, $T^n(\alpha) \nearrow \varepsilon$, see Def. 1.7.4) such that $T^{-n}(\alpha) \searrow \nu$.

The property $\mathrm{Tail}(\mathcal{A}) = \nu$ is a version of the 0-1 Law. An automorphism satisfying (d) is usually called: K-*automorphism*. The symbol K comes from the name: Kolmogorov. Each partition satisfying the properties of α in (d) is called K-partition.

Remark: the properties (a)-(c) make sense for endomorphisms and they are equivalent (proofs are the same as for automorphisms). Moreover they hold for an endomorphis iff they hold for its natural extension.

Proof. (a part of) To show the reader what is the Theorem about let us prove at least some implications:

(a)⇒(b) Let $A \in \mathcal{F}(\mathrm{Tail}(\mathcal{A}))$ for a finite partition $\mathcal{A}$. Then $A \in \mathcal{F}(\bigvee_{k=n}^{\infty} T^{-k}(\mathcal{A}))$ for every n. Hence, by K-mixing, $\mu(A \cap A) - \mu(A)\mu(A) = 0$ and therefore $\mu(A) = 0$ or 1.

(b)⇒(c) Suppose $\mathrm{h}(T, \mathcal{A}) = 0$ for a finite partition $\mathcal{A}$. Then $\mathrm{H}(\mathcal{A}|\mathcal{A}^-) = 0$, hence $I(\mathcal{A}|\mathcal{A}^-) = 0$ a.s. (see Sec.8), hence $\mathcal{A} \leq \mathcal{A}^-$. Hence

$$\bigvee_{k=0}^{\infty} T^{-k}(\mathcal{A}) = \bigvee_{k=1}^{\infty} T^{-k}(\mathcal{A}) \quad \text{and} \quad \bigvee_{k=m}^{\infty} T^{-k}(\mathcal{A}) = \bigvee_{k=n}^{\infty} T^{-k}(\mathcal{A})$$

for every $m, n \geq 0$. So $\bigwedge_{n=0}^{\infty} \bigvee_{k=n}^{\infty} T^{-k}(\mathcal{A}) = \bigvee_{k=0}^{\infty} T^{-k}(\mathcal{A})$. The latter partition is $\mathrm{Tail}\mathcal{A}$, so it is equal to ν by (b). But it is finer than $\mathcal{A}$, hence $\mathcal{A} = \nu$. So each finite partition different from ν, the trivial one, has positive entropy.

(b)⇒(d) (in case there exists a finite two-sided generator $\mathcal{B}$, i.e. $\bigvee_{n=-\infty}^{\infty} T^n(\mathcal{B}) = \varepsilon$) $\xi = \bigvee T_{n=0}^{\infty} T^{-n}(\mathcal{B})$ is exhausting. ♣

Let us finish the Section with the following useful:

Definition 1.10.2. A measure preserving endomorphism is said to be *exact* if

$$\bigwedge_{n=0}^{\infty} T^{-n}(\varepsilon) = \nu,$$

(Remind that ε is the partition into points and ν is the trivial partition $\{X\}$.)

Exercise: Prove that exactness is equivalent to the property: $\mu_e(T^n(A)) \to 1$ for every A of positive measure (μ_e is outer measure), or to the property: $\mu(T^n(A)) \to 1$ provided $\mu(A) > 0$ and the sets $T^n(A)$ are measurable.

The property *exact* implies the natural extension is a K-automorphism (in Theorem 1.10.1(d) set for α the lift of ε). The converse is of course false. Non-one atom space automorphisms are not exact. Observe however that if T is an automorphism and α is a measurable partition satisfying (d), then the factor map T/α on X/α is exact. Exercise: prove that T is the natural extension of T/α. Remind finally (Sec. 9) that even for exact endomorphisms $\mathrm{h}(\varepsilon|T^{-1}(\varepsilon))$ can be strictly less than $\mathrm{h}(T)$.

§1.11. PROBABILITY LAWS AND BERNOULLI PROPERTY.

For $(X, \mathcal{F}, \mu)$ a probability space (or whenever it is needed: a Lebesgue space). Let f and g be real square-integrable functions on X. For every positive integer n the n-th *correlation* of the pair f, g, is the number

$$C_n(f,g) := \int f \cdot (g \circ T^n)\, d\mu - \int f\, d\mu \int g\, d\mu.$$

provided the above integrals exist. Notice that due to the T-invariance of μ we can also write

$$C_n(f,g) = \int (f - Ef)\big((g - Eg) \circ T^n\big)\, d\mu,$$

where we write $Ef = \int f\, d\mu$ and $Eg = \int g\, d\mu$

Let $g : X \to R$ be a square-integrable function. The limit

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n} \int \Big(\sum_{j=0}^{n-1} g \circ T^j - nEg\Big)^2 d\mu \tag{1.11.1}$$

is called *asymptotic variance* or *dispersion*, provided it exists.

Write $g_0 = g - Eg$. Then we can rewrite the above as $\sigma^2 = \lim_{n\to\infty} \frac{1}{n} \int \Big(\sum_{j=0}^{n-1} g_0 \circ T^j\Big)^2 d\mu$.

Another useful expression for the asymptotic variance is the following

$$\sigma^2(g) = \int g_0^2\, d\mu + 2 \sum_{j=1}^{\infty} \int g_0 \cdot (g_0 \circ T^j)\, d\mu. \tag{1.11.2}$$

The convergence of the series of correlations $C_n(g,g)$ in (1.11.2) easily implies that $\sigma^2(g)$ from this formula is equal to σ^2 defined in (1.11.1), compare the computation in the proof of Theorem 1.11.3 later on.

We say that the *law of iterated logarithm*, LIL, is satisfied for g if $\sigma^2(g)$ exists (i.e. the above series converges) and

$$\limsup_{n\to\infty} \frac{\sum_{j=0}^{n-1} g \circ T^j - nEg}{\sqrt{n \log\log n}} = \sqrt{2\sigma^2} \quad \mu- \text{ almost surely }. \tag{1.11.3}$$

(μ almost surely, a.s., means μ almost everywhere, a.e. This is the probability theory language.)
We say that the *central limit theorem*, CLT, holds, if

$$\mu\left(\left\{x \in X : \frac{\sum_{j=0}^{n-1} g \circ T^j - nEg}{\sqrt{n}} < r\right\}\right) \to \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{r} e^{-t^2/2\sigma^2}\, dt. \tag{1.11.4}$$

For $\sigma \neq 0$ the convergence is for all r, for $\sigma^2 = 0$ the convergence holds for $r \neq 0$ and on the right hand side one sets 0 for $r < 0$ and 1 for $r > 0$.

The LIL and CLT for $\sigma^2 \neq 0$ are often, and this is the case in Theorem 1.11.1 below, a consequence of the *almost sure invariance principle*, ASIP, which says that the sequence of random variables g, $g \circ T$, $g \circ T^2$, *centered at the expectation value* i.e. provided $Eg = 0$, is "approximated with the rate $n^{1/2-\varepsilon}$" for an $\varepsilon > 0$, depending on δ in Theorem 1.11.1 below, by a martingale difference sequence and a respective Brownian motion.

Theorem 1.11.1. Let $(X, \mathcal{F}, \mu)$ be a probability space and T an endomorphism preserving μ. Let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra. Write $\mathcal{G}_m^n := \bigvee_{j=m}^n T^{-j}(\mathcal{G})$ (notation from §.1.6) for $m \leq n \leq \infty$ and suppose that the following property called *ϕ-mixing* holds:
There exists a sequence $\phi(n), n = 0, 1, ..$ of positive numbers satisfying

$$\sum_{n=1}^{\infty} \phi^{1/2}(n) < \infty \tag{1.11.5}$$

such that for every $A \in \mathcal{G}_0^m$ and $B \in \mathcal{G}_n^\infty$, $0 \leq m \leq n$ we have

$$|\mu(A \cap B) - \mu(A)\mu(B)| \leq \phi(n - m)\mu(A). \tag{1.11.6}$$

Finally consider a $\mathcal{G}_0^\infty$ measurable function $g : X \to I\!\!R$ such that

$$\int |g|^{2+\delta}\, d\mu < \infty \quad \text{for some } \delta > 0,$$

and that for all $n \geq 1$

$$\Big(\int |h - E(h|\mathcal{G}_0^n)|^{2+\delta}\Big)\Big)^{-(2+\delta)} \leq Kn^{-s}, \qquad K > 0, s > 0 \text{ large enough.} \tag{1.11.7}$$

(A concrete formula for s can be given, depending on δ.)

Then g satisfies CLT and LIL.

LIL for $\sigma^2 \neq 0$ is a special case, for $\psi(n) = \sqrt{2 \log\log n}$, of the following: for every real positive non-decreasing function ψ one has, provided $\int g\, d\mu = 0$,

$$\mu\{x \in X : \sum_{j=0}^{n} g(T^j(x)) > \psi(n)\sqrt{\sigma^2 n} \text{ for infinitely many } n\} = 0 \text{ or } 1$$

according as $\int_1^\infty \frac{\psi(t)}{t} \exp(-\frac{1}{2}\psi^2(t))\, dt$ converges or diverges.

As we already remarked, this Theorem, for $\sigma^2 \neq 0$, is a consequence of ASIP and the similar conclusions for the standard Brownian motion. We do not give the proofs here. For ASIP and further references see [Philipp, Stout, Ch.4,7]. Let us discuss only the existence of

σ^2. It follows from the following consequence of (1.11.6): For α, β square integrable real functions on X, α measurable in $\mathcal{G}_0^m$ and β measurable in $\mathcal{G}_n^\infty$ we have

$$\left|\int \alpha\beta \, d\mu - E\alpha \, E\beta\right| \leq 2(\phi(n-m))^{1/2}\|\alpha\|_2\|\beta\|_2. \tag{1.11.8}$$

The proof of this inequality is not difficult, but tricky, with the use of Hölder inequality, see [Ibragimov] or [Billingsley, 1968]. It is sufficient to work with the functions $\alpha = \sum_i a_i 1\!\!1_{A_i}, \beta = \sum_j a_j 1\!\!1_{A_j}$ for finite partitions (A_i) and (B_j), as with mixing in Sec. 10. Note that if instead of (1.11.6) we have stronger:

$$|\mu(A\cap B) - \mu(A)\mu(B)| \leq \phi(n-m)\mu(A)\mu(B), \tag{1.11.9}$$

as will happen in Ch.3, then we very easily obtain in (1.11.7) the estimate by $\phi(n-m)\|\alpha\|_1\|\beta\|_1$, by the computation the same as for mixing in Sec.10.

We may assume that g is centered at the expectation value. Write $g = k_n + r_n = E(g|\mathcal{G}_0^{[n/2]}) + (g - E(g|\mathcal{G}_0^{[n/2]})$. We have

$$\left|\int g(g\circ T^n)\, d\mu\right| \leq$$

$$\left|\int k_n(k_n\circ T^n)\, d\mu\right| + \left|\int k_n(r_n\circ T^n)\, d\mu\right| + \left|\int r_n(k_n\circ T^n)\, d\mu\right| + \left|\int r_n(r_n\circ T^n)\, d\mu\right| \leq$$

$$2(\phi(n-[n/2]))^{1/2}\|k_n\|_2^2 + 2\|k_n\|_2\|r_n\|_2 + \|r_n\|_2^2 \leq$$

$$2(\phi(n-[n/2]))^{1/2}\|k_n\|_2^2 + 2K[n/2]^{-s}\|k_n\|_2 + K[n/2]^{-2s},$$

the first summand estimated according to (1.11.8). For $s > 1$ we obtain convergence of the series of correlations.

Let us go back to the discussion of the ϕ-mixing. If $\mathcal{G}$ is associated to a finite partition that is a generator, ϕ-mixing with $\phi(n) \to 0$ as $n \to \infty$ implies K-mixing (see Sec.10). Indeed B is the same in both definitions, whereas A in K-mixing can be approximated by sets belonging to $\mathcal{G}_0^m$. We leave details to the reader.

Intuitively both notions mean that any event B in remote future weakly depends on the present state A, i.e. $|\mu(B) - \mu(B|A)|$ is small.

In applications $\mathcal{G}$ will be usually associated to a finite or countable partition.

In Theorems 1.11.1, the case $\sigma^2 = 0$ is easy. It relies on Theorem 1.11.3 below. Let us first introduce the following fundamental

Definition 1.11.2. Two functions $f, g : X \to I\!\!R$ (or $\mathbb{C}$) are said to be *cohomologous* in a space $\mathcal{K}$ of real (or complex) -valued functions on X (or f is called cohomologous to g), if there exists $h \in \mathcal{K}$ such that

$$f - g = h\circ T - h. \tag{1.11.10}$$

If f, g are defined mod 0, then (1.11.10) is understood a.s.. This formula is called a *cohomology equation.*

Theorem 1.11.3. Let f be a square integrable function on a probability space $(X, \mathcal{F}, \mu)$, centered at the expectation value. Assume that $\sum_{n=0}^{\infty} n|\int f \cdot (f \circ T^n)\, d\mu| < \infty$. Then the following three conditions are equivalent:

(a) $\sigma^2(f) = 0$;

(b) All the sums $S_n = S_n f = \sum_{j=0}^{n-1} f \circ T^j$ have the norm in L^2 (the space square integrable functions) bounded by the same constant;

(c) f is cohomologous to 0 in the space $\mathcal{H} = L^2$.

Proof. (c)$\Rightarrow$(a) follows immediately from (1.11.1) after substituting $f = h \circ T - h$. Let us prove (a)$\Rightarrow$(b): Write C_j for the correlation $\int f \cdot (f \circ T^j)\, d\mu, j = 0, 1, \ldots.$ Then

$$\int |S_n|^2\, d\mu = nC_0^2 + 2\sum_{j=1}^{n}(n-j)C_j$$

$$= n\Big(C_0^2 + 2\sum_{j=1}^{\infty} C_j\Big) - 2n \sum_{j=n+1}^{\infty} C_j - 2\sum_{j=1}^{n} j \cdot C_j = n\sigma^2 - I_n - II_n.$$

Since $I_n \to 0$ and II_n stays bounded as $n \to \infty$ and $\sigma^2 = 0$, we deduce that all the sums S_n are uniformly bounded in L^2.

(b)$\Rightarrow$(c): $f = h \circ T - h$ for any h a limit in weak*-topology of the bounded sequence $\frac{1}{n} S_n$. We leave the easy computation to the reader. (This computation will be provided in detail in the similar situation of Bogolyubov-Krylov Theorem, in 2.1.14.). ♣

Now Theorem 1.11.1 for $\sigma^2 = 0$ follow from (c), which gives $\sum_{j=0}^{n-1} f \circ T^j = h \circ T^n - h$, with the use of Borel-Cantelli lemma.

Remark.. Th.1.11.1 in the two-sided case: where g depens on $\mathcal{G}_j = T^j(\mathcal{G})$ for $j = \ldots, -1, 0, 1, \ldots$ for an automorphism T, also holds. In 1.11.8 one should replace $\mathcal{G}_0^n$ by $\mathcal{G}_{-n}^n$

Given two finite partitions $\mathcal{A}$ and $\mathcal{B}$ of a probability space and $\varepsilon \geq 0$ we say that $\mathcal{B}$ is *ε-independent* of $\mathcal{A}$ if there is a subfamily $\mathcal{A}' \subset \mathcal{A}$ such that $\mu(\bigcup \mathcal{A}') > 1 - \varepsilon$ and for every $A \in \mathcal{A}'$

$$\sum_{B \in \mathcal{B}} \left| \frac{\mu(A \cap B)}{\mu(A)} - \mu(B) \right| \leq \varepsilon. \tag{1.11.11}$$

Given an ergodic measure preserving endomorphism $T : X \to X$ of a Lebesgue space, a finite partition $\mathcal{A}$ is called *weakly Bernoulli* (abbr. WB) if for every $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that the partition $\bigvee_{j=n}^{s} T^{-j}(\mathcal{A})$ is ε-independent of the partition $\bigvee_{j=0}^{m} T^{-j}(\mathcal{A})$ for every $0 \leq m \leq n \leq s$ such that $n - m \geq N$.

Of course in the definition of ε-independence we can consider any measurable (maybe uncountable) partition $\mathcal{A}$ and write conditional measures $\mu_A(B)$ in (1.11.11). Then for T an automorphism we can replace in the definition of WB $\bigvee_{j=n}^{s} T^{-j}(\mathcal{A})$ by $\bigvee_{j=0}^{s-n} T^{-j}(\mathcal{A})$ and $\bigvee_{j=0}^{m} T^{-j}(\mathcal{A})$ by $\bigvee_{j=-n}^{m-n} T^{-j}(\mathcal{A})$ and set $n = \infty, n - m \geq N$. WB in this formulation becomes one more version of weak dependence of present (and future) from remote past. If $\varepsilon = 0$ and $N = 1$ then all partitions $T^{-j}(\mathcal{A})$ are mutually independent (recall that $\mathcal{A}, \mathcal{B}$ are called independent if $\mu(A \cap B) = \mu(A)\mu(B)$ for every $A \in \mathcal{A}, B \in \mathcal{B}$.). We say then that $\mathcal{A}$ is Bernoulli. If $\mathcal{A}$ is a generator (two-sided generator), then clearly T on $(X, \mathcal{F}, \mu)$ is isomorphic to one-sided (two-sided) Bernoulli shift of $\sharp\mathcal{A}$ symbols, see Chapter 0, Examples 0.8. The following famous theorem of Friedman and Ornstein holds:

Theorem 1.11.4. If $\mathcal{A}$ is a weakly Bernoulli finite two-sided generating partition of X for an automorphism T, then T is isomorphic to a two-sided Bernoulli shift.

Of course the standard Bernoulli partition (in particular the number of its states) in the above Bernoulli shift can be different from the image under the isomorphism of the WB partition.

Bernoulli shift above is unique in the sense that each two two-sided Bernoulli shifts of the same entropy are isomorphic [O].

Note that ϕ-mixing in the sense (1.11.9), with $\phi(n) \to 0$, for $\mathcal{G}$ associated to a finite partition $\mathcal{A}$, implies weak Bernoulli.

Central Limit Theorem is a much weaker property than LIL. We want to end this Section with a useful abstract theorem that allows us to deduce CLT for g without specifying $\mathcal{G}$. This Theorem similarly as Theorem 1.11.1 can be proved with the use of an approximation by a martingale difference sequence.

Theorem 1.11.5. Let $(X, \mathcal{F}, \mu)$ be a probability space and $T : X \to X$ an automorphism preserving μ. Let $\mathcal{F}_0 \subset \mathcal{F}$ be a σ-algebra such that $T^{-1}(\mathcal{F}_0) \subset F_0$. Denote $\mathcal{F}_n = T^{-n}(\mathcal{F}_0)$ for all integer $n = ..., -1, 0, 1, ...$ Let g be a real square integrable function. If

$$\sum_{n \geq 0} \|E(g|\mathcal{F}_n)\|_2 + \|g - E(g|\mathcal{F}_{-n})\|_2 < \infty,$$

then g satisfies CLT.

Exercises.

0. Prove that for any two σ-algebras $\mathcal{F} \ni \mathcal{F}'$ and ϕ an $\mathcal{F}$-measurable function, the conditional expectation value operator $L^p(X, \mathcal{F}, \mu) \ni \phi \to E(\phi|\mathcal{F}')$ has norm 1 in L^p, for every $1 \leq p \leq \infty$. (Hint: Prove that $E((\vartheta \circ |\phi|)|\mathcal{F}') \geq \vartheta \circ E((|\phi|)|\mathcal{F}')$ for convex ϑ, in particular for $t \mapsto t^p$.)

1. Let T be an ergodic automorphism of a probability non-atomic measure space and $\mathcal{A}$ its partition into orbits $\{T^n(x), n = ..., -1, 0, 1...\}$. Prove that $\mathcal{A}$ is not measurable.

Suppose we do not assume ergodicity of T. What is the largest measurable partition, smaller than the partition into orbits? (Hint: Th.1.8.8.)

2. Prove that the following partitions of measure spaces are not measurable:

(a) Let $T : S^1 \to S^1$ be a mapping of the unit circle with Haar measure defined by $T(z) = e^{2\pi i\alpha}z$ for an irrational α. $\mathcal{P}$ is the partition into orbits;

(b) T is the automorphism of the 2-dimensional torus $\mathbb{R}^2/\mathbb{Z}^2$, given by a hyperbolic integer matrix of determinant 1. Let $\mathcal{P}$ be the partition into stable, or unstable, lines (i.e. straight lines parallel to an eigenvector of the matrix);

(c) Let $T : S^1 \to S^1$ be defined by $T(z) = z^2$. Let $\mathcal{P}$ be the partition into *grand orbits*, i.e. equivalence classes of the relation $x \sim y$ iff $\exists m, n \geq 0$ such that $T^m(x) = T^n(y)$.

3. Prove that every Lebesgue space is isomorphic to the unit interval equipped with the Lebesgue measure together with countably many atoms.

4. Prove that every separable complete metric space with a measure on the σ-algebra containing all open sets, minimal among complete measures, is Lebesgue space.

Hint: [Rohlin 1949, 2.7].

5. Let $(X, \mathcal{F}, \mu)$ be a Lebesgue space. Then $Y \subset X, \mu_e(Y) > 0$ is measurable iff $(Y, \mathcal{F}_Y, \mu_Y)$ is Lebesgue, where μ_e is the outer measure, $\mathcal{F}_Y = \{A \cap Y : A \in \mathcal{F}\}$ and $\mu_Y(A) = \frac{\mu_e(A\cap Y)}{\mu_e(Y)}$.

Hint: If $\mathbf{B}$=(B_n) is a basis for $(X, \mathcal{F}, \mu)$, then $B'_n = B_n \cap Y)$ is a basis for $(Y, \mathcal{F}_Y, \mu_Y)$. Add to Y one point for each sequence $(B'_n)^{\varepsilon_n}$ whose intersection is missing in Y and in the space $\tilde{Y}$ obtained in such a way generate complete measure space $(\tilde{Y}, \tilde{\mathcal{F}}, \tilde{\mu})$ from the extension $\tilde{\mathbf{B}}$ of the basis (B'_n). Borel sets with respect to $\mathbf{B}$ in X correspond to Borel sets with respect to $\tilde{\mathbf{B}}$ and sets of μ measure 0 correspond to sets of $\tilde{\mu}$ measure 0. So measurability of Y implies $\tilde{\mu}(\tilde{Y} \setminus Y) = 0$.

6. Prove Th.1.6.3.

Hint: In the case both spaces are unit intervals with standard Lebesgue measure, consider all intervals J' with rational ends. $J = T^{-1}(J')$ is contained in a Borel set B_J with $\mu(B_J \setminus J) = 0$. Remove from X a Borel set of measure 0 containing $\bigcup_J (B_J \setminus J)$. Then T becomes a Borel map, hence it is a Baire function, hence due to the injectivity it maps Borel sets to Borel sets.

7. (a) Consider the unit square $[0,1]\times[0,1]$ equipped with Lebesgue measure. For each $x \in [0,1]$ let $\mathcal{A}_x$ be the partition into points (x', y) for $x' \neq x$ and the interval $\{x\} \times [0,1]$. What is $\bigwedge_x \mathcal{A}_x$? Let $\mathcal{B}_x$ be the partition into the intervals $\{x'\} \times [0,1]$ for $x' \neq x$ and the points $\{(x,y) : y \in [0,1]\}$. What is $\bigwedge_x \mathcal{B}_x$?

(b) Find two measurable partitions $\mathcal{A}, \mathcal{A}'$ of a Lebesgue space such that their set-theoretic intersection (i.e. the largest partition such that $\mathcal{A}, \mathcal{A}'$ are finer than it) is not measurable.

8. Find an example of $T: X \to X$ an endomorphism of a probability space $(X, \mathcal{F}, \mu)$, injective and onto, such that for the system $... \xrightarrow{T} X \xrightarrow{T} X$, natural extension does not exist.

Hint: Set X the unit circle and T irrational rotation. Let A be a set consisting of exactly one point in each T-orbit. Set $B = \bigcup_{j \geq 0} T^j(A)$. Notice that B is not Lebesgue measurable and that the outer measure of B is 1 (use *unique ergodicity* of T, i.e. that (1.2.1a) holds for every x)

Let $\mathcal{F}$ be the σ-algebra consisting of all the sets $C = B \cap D$ for D Lebesgue measurable, set $\mu(C) = \mathrm{Leb}(D)$, and of $C \subset X \setminus B$, set then $\mu(C) = 0$. Note that $\bigcap_{n \geq 0} T^n(B) = \emptyset$ and in the set-theoretic inverse limit the set $\pi_{-n}^{-1}(B) = \pi_0^{-1}(T^n(B))$ would be of measure 1 for every $n \geq 0$.

9. (a) Prove that in a Lebesgue space $d(\mathcal{A}, \mathcal{B}) := H(\mathcal{A}|\mathcal{B}) + H(\mathcal{B}|\mathcal{A})$ is a metric in the space Z of countable partitions (mod 0) of finite entropy. Prove that the metric space (Z, d) is separable and complete.

(b) Prove that if T is an endomorphism of the Lebesgue space, then the function $\mathcal{A} \to \mathrm{h}(T, \mathcal{A})$ is continuous for $\mathcal{A} \in Z$ with respect to the above metric d.

Hint: $|\mathrm{h}(T, \mathcal{A}) - \mathrm{h}(T, \mathcal{B})| \leq \max\{\mathrm{H}(\mathcal{A}|\mathcal{B}), \mathrm{H}(\mathcal{B}|\mathcal{A})\}$. Compare Proof of Th.1.4.5.

9'. (a) Let $d_0(\mathcal{A}, \mathcal{B}) := \sum_i \mu(A_i \div B_i)$ for partitions of a probability space into r measurable sets $\mathcal{A} = \{A_i, i = 1, ..., r\}$ and $\mathcal{B} = \{B_i, i = 1, ..., r\}$. Prove that for every r and every $d > 0$ there exists $d_0 > 0$ such that if $\mathcal{A}, \mathcal{B}$ are partitions into r sets and $d_0(\mathcal{A}, \mathcal{B}) < d_0$, then $d(\mathcal{A}, \mathcal{B}) < d$

(b) Using (a) give a simple proof of Corollary 1.8.7"'. (Hint: Given an arbitrary finite $\mathcal{A}$ construct $\mathcal{B} \leq \mathcal{B}_m$ so that $d_0(\mathcal{A}, \mathcal{B})$ be small for m large. Next use (a) and Theorem 1.4.4.d).

10. Prove that there exists a finite generator for every T, a continuous positively expansive map of a compact metric space (see the definition of *positively expansive* in Ch.2, Sec.5).

11. Compute the entropy $\mathrm{h}(T)$ for Markov shifts.

12. Prove that the entropy $\mathrm{h}(T)$ defined either as supremum of $\mathrm{H}(T, \mathcal{A})$ over finite partitions, or over countable partitions of finite entropy, or as $\sup \mathrm{H}(\xi|\xi^-)$ over all measurable partitions ξ that are forward invariant (i.e. $T^{-1}(\xi) \leq \xi$) is the same.

13. Let T be an endomorphism of the 2-dimensional torus $I\!R^2/\mathbf{Z}^2$, given by an integer matrix of determinant larger than 1 and with eigenvalues λ_1, λ_2 such that $|\lambda_1| < 1$ and $|\lambda_2| > 1$.

Let S be the endomorphism of $I\!R^2/\mathbf{Z}^2$ being the cartesian product of $S_1(x) = 2x$ (mod 1) on the circle $I\!R/\mathbf{Z}$ and of $S_2(y) = y + \alpha$ (mod 1), the rotation by an irrational angle α.

Which of the maps T, S is exact? Which has a countable generator of finite entropy?

(Answer: T does not have the generator, but it is exact. The latter holds because for each small parallelepiped P spanned by the eigendirections there exists n such that $T^n(P)$ covers the torus with multiplicity bounded by a constant not depending on P. See ??? S is not exact, but it is ergodic and has a generator.)

14. Prove that if the definition of partition $\mathcal{A}$ ε-independent of partition $\mathcal{B}$ is replaced by $\sum_{A\in\mathcal{A},B\in\mathcal{B}} |\mu(A\cap B) - \mu(A)\mu(B)|$, then the definition of weakly Bernoulli is equivalent to the old one. (Note that now the expression is symmetric with respect to $\mathcal{A}, \mathcal{B}$.)

Bibliographical notes:

For the Martingale Convergence Theorem see for example [Doob], [Billingsley, 1979], [Petersen] or [Stroock]. Its standard proof goes via a *maximal function*. We followed this way in Proof of Shannon, McMillan, Breiman Theorem in Sec.5, L.1.5.1, where we relied on [Petersen] and [Parry]. Remark 1.1.2 is taken from [Neveu,Ch.4.3], see for example [Hoover] for a more advanced theory.

Standard proofs of Birkhoff's Ergodic Theorem also use the idea of maximal function. This concerns in particular the extremaly simple proof in Sec. 2, which has been taken from [KH].

For the material of Sec.6 and related exercises see [Rohlin, 1949]. It is also written in an elegant and a very concise way in [Cornfeld, Fomin, Sinai].

The consideration in Sec.7 leading to the extension of the compatible family $\tilde{\mu}_{\Pi,n}$ to $\tilde{\mu}_\Pi$ is known as Kolmogoroff Theorem on the existence of stochastic process. First, one verifies σ-additivity of a measure on an algebra, next uses the Extension Theorem 1.7.2. Our proof of σ-additivity of $\tilde{\mu}$ on $\tilde{X}$ via Lusin theorem is also a variant of Kolmogoroff's proof. The proofs of σ-additivity on algebras depend unfortunately on topologocal concepts. Halmos wrote [Halmos, p. 212]: "this peculiar and somewhat undesirable circumstance appears to be unavoidable" Indeed the σ-additivity may be not true, see [Halmos, p.214]. Our example of non-existence of natural extension, Exercise 8, is in the spirit of Halmos' example. There might be troubles even with extending a measure from cylinders in product of two measure spaces, see [MR] for a counterexample. On the other hand product measures extend to generated σ-algebras without any additional assumptions [Halmos], [Billingsley, 1979].

For Th.1.8.1: the existence of a countable γ such that $\mathcal{A} \vee \mathcal{B} = \gamma \vee \mathcal{B}$, see [Rohlin, 1949]; for the estimate that follows, see for example [Rohlin 1967] or [Parry]. The simple proof of Th.1.8.6 via convergence in measure has been taken from [Rohlin 1967] and [Wa]. Proof of Th.1.8.8 (b) is taken from [Rohlin, 1967, sec.8.10-11 and 9.8].

For Th.1.9.6 see [Parry, L.10.5]; our proof is different. For the construction of generator and 2-sided generator see again [Rohlin 1967],[Parry] or [CFS]. The same are references to the theory of measurable invariant partitions: exhausting and extremal, and to Pinsker partition, which we omitted because we do not need these notions further in the book, but which are fundamental to understand deeper the measure-theoretic entropy theory. Finally we encourage the reader to become acquainted with spectral theory in relation to mixing properties [CFS].

Th. 1.11.1 can be found in [Philipp, Stout]. See also [PUZ, Part 1.]. For (1.11.8) see [Ibragimov, 1.1.2] or [Billingsley, 1968]. For Th.1.11.3 see [Leonov], [Ibragimov, 1.5.2] or [PUZ, Part 1., L.1]. Theorem 1.11.5 can be found in [Gordin]

References.

[Billingsley, 1965] P. Billingsley: Ergodic Theory and Information. Wiley, New York, 1965.

[Billingsley, 1979] P. Billingsley: Probability and Mesure. Wiley, New York, 1979.

[CFS] I.P.Cornfeld, S.V.Fomin, Ya.G.Sinai: Ergodic Theory. Springer-Verlag, New York, 1982.

[Carathéodory] C. Carathéodory: Vorlesungen über reelle Funktionen. Leipzig-Berlin, 1927.

[Doob] J.L. Doob: Stochastic Processes, Wiley, New York, 1953

[Gordin]

[Hoover] D.N. Hoover: Convergence in distribution ond Skorokhod Convergence for the general theory of processes. Probab. Th. Rel. Fields 89 (1991), 239-259.

[Ibragimov] I.A. Ibragimov: Some limit theorems for stationary processes. Theor. Prob. Appl. 7 (1962), 349-382.

[KH] A. Katok, B. Hasselblatt: Introduction to the Modern Theory of Dynamical Systems. Cambridge University Press, 1995

[Kolmogoroff] A.N. Kolmogoroff: Gundbegriffe der Wahrscheinlichkeitsrechnung. Berlin 1933.

[Lang] S. Lang: Algebra. Addison-Wesley, 1970.

[Leonov] V.P. Leonov: On the dispersion of time-dependent means of stationary stochastic process. Theor. Prob. Appl. 6 (1961), 93-101.

[MR] E. Marczewski, C. Ryll-Nardzewski:

[Neveu] Jacques Neveu: Bases Mathématiques du Calcul des Probabilités. Masson et Cie, Paris, 1964.

[Parry] W.Parry: Entropy and Generators in Ergodic Theory. W.A.Benjamin, Inc., New York, 1969.

[Petersen] K. Petersen: Ergodic Theory. Cambridge University Press, 1983.

[Philipp, Stout] W.Philipp, W.Stout: Almost sure invariance principles for partial sums of weakly dependent random variables. Mem.AMS 161 (1975).

[PUZ] F. Przytycki, M. Urbański, A. Zdunik: Harmonic, Gibbs and Hausdorff measures for holomorphic maps. Part 1. Anals of Math. 130 (1989), 1-40. Part 2. Studia Math. 97 (1991), 189-225.

[Rohlin, 1949] V.A.Rohlin: On the fundamental ideas of measure theory. Mat.Sb. 67.1 (1949), 107-150. Amer. Math. Soc. Trans. (1)10 (1962), 1-54.

[Rohlin, 1967] V.A.Rohlin: Lectures on the entropy theory of transformations with invariant measure. Uspehi Mat. Nauk 22.5 (1967), 3-56. Russian Math. Surveys, 22.5 (1967), 1-52.

[Stroock] D.W.Stroock: Probability Theory, an Analytic View. Cambridge University Press, 1993.

[Wa] - P. Walters: An Itroduction to Ergodic Theory. Springer-Verlag, New York, 1982.

CHAPTER 2
ERGODIC THEORY ON COMPACT METRIC SPACES

Dec.5, 1999

last revised Feb 24, 2002

In the previous Chapter a measure preserved by a measurable map T was given *a priori*. Here a continuous mapping T of a topological compact space is given and we look for various measures preserved by T. Given a real continuous function ϕ on X we try to maximize the functional *measure theoretical entropy +integral*, i.e. $\mathrm{h}_\mu(T) + \int \phi\, d\mu$. Supremum over all probability measures on the Borel σ-algebra happens to be *topological pressure*, similar to P in the prototype lemma on the finite space or $P(\alpha)$ for ϕ_α on the Cantor set discussed in Introduction. We discuss equilibria, namely measures on which supremum is attained. This Chapter provides an introduction to the theory called *thermodynamical formalism*, which will be the main technical tool in this book. We shall continue to introduce thermodynamical formalism in more specific situations in Chapter 4.

§2.1 INVARIANT MEASURES FOR CONTINUOUS MAPPINGS

We recall in this Section basic facts from functional analysis to study the space of measures and invariant measures. We recall Riesz representation theorem, weak* topology, Schauder fixed point theorem. We recall also Krein-Milman theorem on extremal points and its stronger form: Choquet representation theorem. This gives a variant of Ergodic Decomposition Theorem from Chapter 1.

Let X be a topological space. The *Borel* σ-algebra $\mathcal{B}$ in subsets of X is defined as generated by open subsets of X. We call every probability measure on Borel σ-algebra for X, a *Borel probability measure* on X. We denote the set of all such measures by $M(X)$.

Denote by $C(X)$ the Banach space of real continuous functions on X with the supremum norm: $|\phi| := \sup_{x\in X} |\phi(x)|$. Sometimes we shall use the notation $||\phi||_\infty$, introduced in Ch.1.1 in $L^\infty(\mu)$, though it is compatible only if μ is positive on open sets.

Note that each Borel probability measure μ on X induces a bounded linear functional F_μ on $C(X)$ defined by the formula

$$F_\mu(\phi) = \int \phi\, d\mu. \tag{2.1.1}$$

One can extend the notion of measure and consider σ-additive set function, another name : *signed measure*. Just in definition of measure in Ch.1.1 consider $\mu : \mathcal{F} \to [-\infty, \infty)$

or $\mu : \mathcal{F} \to (-\infty, \infty]$ and keep the notation $(X, \mathcal{F}, \mu)$ from Ch.1. The set of signed measures is a linear space. On the set of finite signed measures, namely with the range $I\!R$, one can introduce the following total variation norm:

$$v(\mu) := \sup \sum_{i=1}^{n} |\mu(A_i)|$$

supremum taken over all finite sequences of disjoint sets in $\mathcal{F}$.

It is easy to prove that every finite signed measure is bounded and that it has finite total variation. It is also not difficult to prove the following

Theorem 2.1.1. (Hahn-Jordan decomposition). For every signed measure on a σ-algebra $\mathcal{F}$ there exists $A_\mu \in \mathcal{F}$ and two measures μ^+ and μ^- such that $\mu = \mu^+ - \mu^-$, μ^- is zero on all measurable subsets of A_μ, μ^- is zero on all measurable subsets of $X \setminus A_\mu$.

Notice that $v(\mu) = \mu^+(X) + \mu^-(X)$.)

A measure (or signed measure) is called *regular*, if for every $A \in \mathcal{F}$ and $\varepsilon > 0$ there exist $E_1, E_2 \in \mathcal{F}$ such that $\mathrm{cl} E_1 \subset A \subset \mathrm{int} E_2$ and for every $C \in \mathcal{F}$, $C \subset E_2 \setminus E_1$ we have $|\mu(C)| < \varepsilon$.

If X is a topological space, denote the space of all regular finite signed measures as above with the total variation norm by $\mathrm{rca}(X)$. The abbreviation rca replaces *regular countably additive*.

If $\mathcal{F} = \mathcal{B}$ Borel σ-algebra and X is metrizable, regularity holds for every finite signed measure. It can be proved by Carathéodory's outer measure argument, compare Proof of Corollary 1.8.7"'.

Denote by $C(X)^*$ the space of all bounded linear functionals on $C(X)$. This is called the *dual space*. *Bounded* means here bounded on the unit ball in $C(X)$, which is equivalent to *continuous*. The space $C(X)^*$ is equipped with the norm $||F|| = \sup\{F(\phi) : \phi \in C(X), |\phi| \leq 1\}$, in which it is a Banach space.

There is a natural order in $\mathrm{rca}(X)$: $\nu_1 \leq \nu_2$ iff $\nu_2 - \nu_1$ is a measure.

Also in the space $C(X)^*$ one can distinguish positive functionals, similarly to measures in signed measures, as those which are non-negative on the set of functions $C^+(X) := \{\phi \in C(X) : \phi(x) \geq 0$ for every $x \in X\}$. This gives the order: $F \leq G$ for $F, G \in C(X)^*$ iff $G - F$ is positive.

Remark that $F \in C(X)^*$ is positive iff $||F|| = F(1\!\!1)$, where $1\!\!1$ is the function on X identically equal to 1. Also for every bounded linear operator $F : C\!\!\!\!C(X) \to C(X)$ which is positive, namely $F(C^+(X)) \subset C^+(X)$, we have $||F|| = |F(1\!\!1)|$.

Remark that (2.1.1) transforms measures to positive linear functionals.

The following fundamental theorem of F. Riesz says more about the transformation $\mu \mapsto F_\mu$ in (2.1.1) (see [DS, pp. 373,380] for the history of this theorem):

Theorem 2.1.2 (Riesz representation theorem). If X is a compact Hausdorff space, the transformation $\mu \mapsto F_\mu$ defined by (2.1.1) is an isometric isomorphism between the Banach space $C(X)^*$ and $\mathrm{rca}(X)$. Furthermore this isomorphism preserves order.

In the sequel we shall often write μ instead of F_μ and *vice versa* and $\mu(\phi)$ or $\mu\phi$ instead of $F_\mu(\phi)$ or $\int \phi\, d\mu$.

Notice that in Theorem 2.1.2 the hard part is the existence, i.e. that for every $F \in C(X)^*$ there exists $\mu \in \mathrm{rca}(X)$ such that $F = F_\mu$. The uniqueness is just the following:

Lemma 2.1.3. If μ and ν are finite regular Borel signed measures on a compact Hausdorff space X, such that $\int \phi\, d\mu = \int \phi\, d\nu$ for each $\phi \in C(X)$, then $\mu = \nu$.

Proof. This is an exercise on the use the regularity of μ and ν. Let $\eta := \mu - \nu = \eta^+ - \eta^-$ in Hahn-Jordan decomposition. Suppose that that $\mu \neq \nu$. Then η^+ (or η^-) is non-zero, say $\eta^+(X) = \eta^+(A_\eta) = \varepsilon > 0$, where A_η is the set defined in Th.2.1.1. Let E_1 be a closed set and E_2 an open set, such that $E_1 \subset A_\eta \subset E_2$ and $\eta^-(E_2 \setminus A_\eta) < \varepsilon/3$, $\eta^+(A_\eta \setminus E_1) < \varepsilon/3$. There exists $\phi \in C(X)$ with values in $[0,1]$ identically 1 on E_1 and 0 on $X \setminus E_2$. Then $\int \phi\, d\eta = \int_{E_1} \phi\, d\eta + \int_{E_2 \setminus A_\eta} \phi\, d\eta + \int_{A_\eta \setminus E_1} \phi\, d\eta \geq \eta^+(E_1) - \varepsilon/3 \geq \eta^+(A_\eta) - 2\varepsilon/3 > 0$. ♣

The space $C(X)^*$ can be equipped with weak* topology. In the case X is *metrizable*, i.e. if there exists a metric on X such that the topology induced by this metric is the original topology on X, weak* topology is characterized by the property that a sequence $\{F_n : n = 1, 2, \ldots\}$ of functionals in $C(X)^*$ converges to a functional $F \in C(X)^*$ if and only if

$$\lim_{n \to \infty} F_n(\phi) = F(\phi) \tag{2.1.2}$$

for every function $\phi \in C(X)$.

If we do not assume X is metrizable, weak* topology is defined as the smallest one in which all elements of $C(X)$ are continuous on $C(X)^*$ (recall that $\phi \in C(X)$ acts on $F \in C(X)^*$ by $F(f)$). One says *weak** to distinguish this topology from the *weak topology* where one considers all continuous functionals on $C(X)^*$, not only those represented by $f \in C(X)$. This discussion of topologies concerns of course every Banach space B and its dual B^*.

Using the bijection established by Riesz representation theorem we can move the weak* topology from $C(X)^*$ to $\mathrm{rca}(X)$ and restrict it to $M(X)$. The topology on $M(X)$ obtained in this way is usually called *weak* topology* on the space of probability measures (sometimes one omits * to simplify the language and notation but one still has in mind weak*, unless stated otherwise). In view of (2.1.2) if X is metrizable this topology is characterized by the property that a sequence $\{\mu_n : n = 1, 2, \ldots\}$ of measures in $M(X)$ converges to a measure $\mu \in M(X)$ if and only if

$$\lim_{n\to\infty} \mu_n(\phi) = \mu(\phi) \tag{2.1.3}$$

for every function $\phi \in C(X)$. Such convergence of measures will be called *weak* convergence* or *weak convergence* and can be also characterized as follows.

Theorem 2.1.4 Suppose that X is metrizable (we do not assume compactness here). A sequence $\{\mu_n : n = 1, 2, \ldots\}$, of Borel probability measures on X converges weakly to a measure μ if and only if $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ for every Borel set A such that $\mu(\partial A) = 0$.

Proof. Suppose that $\mu_n \to \mu$ and $\mu(\partial A) = 0$. Then there exist sets $E_1 \subset \text{int}A, E_2 \supset \text{cl}A$ such that $\mu(E_2 \setminus E_1) = \varepsilon$ is arbitrarily small. Indeed metrizability of X implies that every open set, in particular int$A$, is union of a sequence of closed sets and every closed set is an intersection of a sequence of open sets. For example $\text{int}A = \bigcup_{n=1}^{\infty}\{x \in X : \inf_{z\notin \text{int}A} \rho(x, z) \geq 1/n\}$ for a metric ρ.

Next, there exist $f, g \in C(X)$ with range in the unit interval $[0, 1]$ such that f is identically 1 on E_1, 0 on $X \setminus \text{int}A$, g identically 1 on cl$A$ and 0 on $X \setminus E_2$. Then $\mu_n(f) \to \mu(f)$ and $\mu_n(g) \to \mu(g)$. As $\mu(E_1) \leq \mu(f) \leq \mu(g) \leq \mu(E_2)$ and $\mu_n(f) \leq \mu_n(A) \leq \mu_n(g)$ we obtain

$$\mu(E_1) \leq \mu(f) = \lim_{n\to\infty} \mu_n(f) \leq \liminf_{n\to\infty} \mu_n(A)$$

$$\leq \limsup_{n\to\infty} \mu_n(A) \leq \lim_{n\to\infty} \mu_n(g) = \mu(g) \leq \mu(E_2).$$

As also $\mu(E_1) \leq \mu(A) \leq \mu(E_2)$ we obtain, letting $\varepsilon \to 0$, $\lim_{n\to\infty} \mu_n(A) = \mu(A)$.

Proof in the opposite direction follows from the definition of integral: approximate uniformly an arbitrary continuous function f by simple functions $\sum_{i=1}^{k} \varepsilon_i 1\!\!1_{E_i}$ where $E_i = \{x \in X : \varepsilon_i \leq f(x) < \varepsilon_{i+1}\}$, for an increasing sequence $\varepsilon_i, i = 1, ..., k$ such that $\varepsilon_i - \varepsilon_{i-1} < \varepsilon$ and $\mu(f^{-1}(\{\varepsilon_i\})) = 0$, with $\varepsilon \to 0$. This is possible to find such numbers ε_i because only countably many sets $f^{-1}(a)$ for $a \in I\!\!R$ can have non-zero measure. ♣

Example 2.1.5 The assumption $\mu(\partial A) = 0$ is substantial. Let X be the interval $[0, 1]$. Denote by δ_x the Dirac measure concentrated at the point x, which is defined by the following formula:

$$\delta_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$

for all sets $A \in \mathcal{B}$.

Consider non-atomic probability measures μ_n supported respectively on the ball $B(x, \frac{1}{n})$. The sequence μ_n converges weakly to δ_x but does not converge on $\{x\}$.

Of particular importance is the following

Theorem 2.1.6. The space $M(X)$ is compact in weak* topology.

This theorem follows immediately from compactness in weak* topology of any subset of $C(X)^*$ closed in weak* topology, which is bounded in the standard norm of the dual space $C(X)^*$ (compare for example [DS, V.4.3], where this result is proved for all spaces dual to Banach spaces) and from the way we introduced the weak topology on $M(X)$.

It turns out (see [DS, V.5.1]) that if X is compact metrizable, the space $C(X)^*$ with weak* topology is metrizable, hence in particular $M(X)$ is metrizable.

Let now $T : X \to X$ be a continuous transformation of X. The mapping T is measurable with respect to the Borel σ-algebra. In the very begining of Chap.1.2 we have defined T-invariant meaures μ to satisfy the condition $\mu = \mu \circ T^{-1}$. It means that Borel probability T-invariant meaures are exactly fixed points of the transformation $T_* : M(X) \to M(X)$ defined by the formula $T_*(\mu) = \mu \circ T^{-1}$. It easily follows from the definitions that T_* is continuous.

We denote the set of all T-invariant measures in $M(X)$ by $M(X,T)$. This notation is consistent with the notation from Chapter 1.2. We omit here σ-algebra $\mathcal{F}$ because it is always Borel $\mathcal{B}$.

Noting that $\int \phi \, d(\mu \circ T^{-1}) = \int \phi \circ T \, d\mu$ for any $\mu \in M(X)$ and any integrable function ϕ (Prop. 1.2.0), it follows from Lemma 2.1.3 that a Borel probability measure μ is T-invariant if and only if for every continuous function $\phi : X \to I\!R$

$$\int \phi \, d\mu = \int \phi \circ T \, d\mu. \tag{2.1.4}$$

In order to look for fixed points for T_* one can apply the following very general result whose proof (and the definition of locally convex topological vector spaces, abbreviation: LCTVS) can be found for example in [DS] or [Edwards].

Theorem 2.1.7. (Schauder-Tychonoff theorem [DS, V.10.5) If K is a non-empty compact convex subset of an LCTVS then any continuous transformation $H : K \to K$ has a fixed point.

Assume from now on that X is compact, metrizable. To apply Schauder-Tychonoff theorem consider the LCTVS $C(X)^*$ with weak* topology and $K \subset C(X)^*$, being the image of $M(X)$ under the identification between measures and functionals, given by Riesz representation theorem. Move also T_* to K with the use of this identification. Note that the resulting continuous linear operator, denote it also by T_*, conjugate to $\phi \mapsto \phi \circ T$, restricted to K, is continuous also in the weak* topology. This is an easy fact about conjugate operators. We obtain

Theorem 2.1.8. (Bogolyubov–Krylov theorem [Walters, 6.9.1]) There exists a Borel probability measure μ invariant under T.

Thus, our $M(X,T)$ is non-empty. It is also weak* compact, since it is closed as the set of fixed points for a continuous transformation.

As an immediate consequence of this theorem and Theorem 1.8.8 (Ergodic Decomposition Theorem), we get the following:

Corollary 2.1.9. There exists a Borel ergodic probability measure μ invariant under T.

We shall use the notation $M_e(X,T)$ for the set of all ergodic measures in $M(X,T)$. Write also $\mathcal{E}(M(X,T))$ for the set of extreme points in $M(X,T)$.

Thus, in view of Theorem 1.2.5 and Corollary 2.1.9, we know that $M_e(X,T) = \mathcal{E}(M(X,T)) \neq \emptyset$.

In fact Corollary 2.1.9 can be obtained in a more elementery way without using Theorem 1.8.8. Namely it now immediately follows from Theorem 1.2.5 and the following

Theorem 2.1.10. (Krein–Milman theorem on extremal points [DS, V.8.4]) If K is a nonempty compact convex subset of an LCTVS then the set $\mathcal{E}(K)$ of extreme points of K is nonempty and moreover K is the closure of the convex hull of $\mathcal{E}(K)$.

Below we state Choquet representation theorem which is stronger than Krein-Milman theorem. It corresponds to the Ergodic Decomposition Theorem (Th. 1.8.8). We formulate it in $C(X)^*$ with weak* topology as in [Walters, p.153]. The reader can find a general LCTVS version for example in [Edwards]. We rely here also on [Ruelle, Appendix A.5], where the reader can find further references.

Theorem 2.1.11. Choquet representation theorem. Let K be a nonempty compact convex set in $M(X)$ with weak* topology. Then for every $\mu \in K$ there exists a "mass distribution" i.e. measure $\alpha_\mu \in M(\mathcal{E}(K))$ such that

$$\mu = \int m \, d\alpha_\mu(m).$$

This integral converges in weak* topology which means that for every $f \in C(X)$

$$\mu(f) = \int m(f) \, d\alpha_\mu(m). \tag{2.1.5}$$

Notice that we have had already the formula analogous to (2.1.5) in Theorem 2.1.6.

Notice that Krein-Milman theorem follows from Choquet representation theorem because one can weakly approximate α_μ by measures on $\mathcal{E}(K)$ with finite support (finite linear combinations of Dirac measures).

Example. 2.1.12. For $K = M(X)$ we have $\mathcal{E}(K) = \{\text{Dirac measures on } X\}$. Then $\alpha_\mu\{\delta_x : x \in A\} = \mu(A)$ for every $A \in \mathcal{B}$ defines a Choquet representation for every $\mu \in M(X)$. (Exercise)

Choquet theorem asserts the existence of α_μ satisfying (2.1.5) but not uniqueness, which is usually not true. A compact closed set K with the uniqueness of α_μ satisfying (2.1.5), for every $\mu \in M(K)$ is called *symplex*.

Theorem 2.1.13. $K = M(X)$ or $K = M(X,T)$ for every continuous $T : X \to X$ is a symplex.

Proof in the case of $K = M(X)$ is very easy, see Example 2.1.12. A proof for $K = M(X,T)$ is also not hard. The reader can look in [Ruelle, A.5.5]. Proof relies on the fact that two different measures $\mu_1, \mu_2 \in \mathcal{E}(M(X,T))$ are singular (see Theorem 1.2.3). Observe that $||\mu_1 - \mu_2|| = 2$.

One proves in fact that for every $\mu_1, \mu_2 \in M(X,T)$, $||\alpha_{\mu_1} - \alpha_{\mu_2}|| = ||\mu_1 - \mu_2||$.

Let us go back to Schauder-Tychonoff theorem (Th 2.1.7). We shall use it in this book later, in Chapter 4 Sec.2, for maps different from T_*. Just Bogolyubov-Krylov theorem proved above with the help of Theorem 2.1.7, has a different more elementary proof due to the fact that T_* is affine. A general theorem on the existence of a fixed point for a family of commuting continuous affine maps on K is called Markov-Kakutani theorem , [DS, V.10.6], [Walters, 6.9]).

2.1.14. An alternative proof of Theorem 2.1.8. Take an arbitrary $\nu \in M(X)$ and consider the sequence

$$\mu_n = \mu_n(\nu) = \frac{1}{n}\sum_{j=0}^{n-1} T_*^j(\mu)$$

In view of Theorem 2.1.4 it has a weakly convergent subsequence, say $\{\mu_{n_k} : k = 1, 2, \ldots\}$. Denote its limit by μ. We shall show that μ is T-invariant.

We have

$$T_*(\mu_{n_k}) = T_*(\frac{1}{n_k}\sum_{j=0}^{n_k-1} T_*^j(\nu)) = (\frac{1}{n_k}\sum_{j=0}^{n_k-1} T_*^{j+1}(\nu))$$

So for every $\phi \in C(X)$ we have

$$|\mu(\phi) - T_*(\mu(\phi))| = |\lim_{k\to\infty}\Big(\mu_{n_k}(\phi) - T_*(\mu_{n_k})(\phi)|\Big) \le$$

$$\lim_{k\to\infty}\frac{1}{n_k}|\nu(\phi) - T_*^{n_k}(\nu)(\phi)| \le \lim_{k\to\infty}\frac{2}{n_k}|\phi| = 0.$$

This in view of Lemma 2.1.3 finishes the proof. ♣

Remark. If in the above proof we consider $\nu = \delta_x$, a Dirac measure, then $T_*^j(\delta_x) = \delta_{T^j(x)}$ and $\mu_n(\phi) = \frac{1}{n}\sum_{j=0}^{n-1}\phi(T^j(x))$. If we have *a priori* $\mu \in M(X,T)$ then

$$\mu_n(\delta_x) = \frac{1}{n}\sum_{j=0}^{n-1}\delta_{T^j(x)}$$

is weakly convergent for μ-a.e.$x \in X$ by Birkhoff ergodic theorem.

Remark. Recall that in Birkhof ergodic theorem (Chapter 1), for $\mu \in M(X,T)$ for every integrable f one considers $\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\phi(T^j(x))$ for a.e. x. This "almost every" depends on f. If X is compact, as in this Chapter, one can reverse the order of quantificators for continuous functions.

Namely there exists $\Lambda \in \mathcal{B}$ such that $\mu(\Lambda) = 1$ and for every $f \in C(X)$ and $x \in \Lambda$ the limit $\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1} f(T^j(x))$ exists.

Remark. We could take in 2.1.14 an arbitrary sequence $\nu_n \in M(X)$ and take $\mu_n := \mu_n(\nu_n)$. This gives a general method of constructing measures in $M(X,T)$, see for example Proof of Variational principle in Section 4. (This point of view is taken from [Walters]).

We end this Section with the following Lemma useful in the sequel.

Lemma 2.1.15. For every finite partition P of the space $(X, \mathcal{B}, \mu)$ where X as above is a metrizable compact space, $\mathcal{B}$ is Borel σ-algebra and $\mu \in M(X,T)$, if $\sum_{A\in \mathrm{P}}\mu(\partial A) = 0$, then the entropy $\mathrm{H}_\nu(\mathrm{P})$ is a continuous function of $\nu \in M(X,T)$ at μ. The entropy $\mathrm{h}_\nu(T,\mathrm{P})$ is upper semicontinuous at μ.

Proof. The continuity of $\mathrm{H}_\nu(\mathrm{P})$ follows immediately from Theorem 2.1.4. This applied to the partitions $\bigvee_{i=1}^{n-1} T^{-i}\mathrm{P}$ gives the upper semicontinuity of $\mathrm{h}_\nu(T,\mathrm{P})$ as the limit of the decreasing sequence of continuous functions $\frac{1}{n}\mathrm{H}_\nu(\bigvee_{i=1}^{n-1} T^{-i}\mathrm{P})$. See Lemma 1.4.3. ♣

§2.2 TOPOLOGICAL PRESSURE AND TOPOLOGICAL ENTROPY

This section is of topological character and no measure is involved. We introduce and examine here some basic topological invariants coming from thermodynamic formalism of statistical physics.

Let $\mathcal{U} = \{A_i\}_{i\in I}$ and $\mathcal{V} = \{B_j\}_{j\in J}$ be two covers of the compact metric space X considered in the previous section. We define the new cover $\mathcal{U} \vee \mathcal{V}$ putting

$$\mathcal{U} \vee \mathcal{V} = \{A_i \cap B_j : \; i \in I, \, j \in J\} \tag{2.2.1}$$

and we write

$$\mathcal{U} \prec \mathcal{V} \iff \forall_{j\in J} \, \exists_{i\in I} \;\; B_j \subset A_i \tag{2.2.2}$$

Let, as in the previous section, $T : X \to X$ be a continuous transformation of X. Let $\phi : X \to I\!\!R$ be a continuous function, frequently called it potential and let $\mathcal{U}$ be a finite, open cover of X. For every integer $n \geq 1$ we set

$$\mathcal{U}^n = \mathcal{U} \vee T^{-1}(\mathcal{U}) \vee \ldots \vee T^{-(n-1)}(\mathcal{U}),$$

for every set $Y \subset X$

$$S_n\phi(Y) = \sup\{\sum_{k=0}^{n-1} \phi \circ T^k(x) : x \in Y\}$$

and for every $n \geq 1$

$$Z_n(\phi, \mathcal{U}) = \inf_{\mathcal{V}}\{\sum_{U\in\mathcal{V}} \exp S_n\phi(U)\} \tag{2.2.3}$$

where $\mathcal{V}$ ranges over all covers of X contained (in the sense of inclusion) in $\mathcal{U}^n$. The quantity $Z_n(\phi, \mathcal{U})$ is sometimes called the partition function.

Lemma 2.2.1. The limit $\mathrm{P}(\phi, \mathcal{U}) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(\phi, \mathcal{U})$ exists and moreover it is finite. In fact $\mathrm{P}(\phi, \mathcal{U}) \geq -||\phi||_\infty$.

Proof. Fix $m, n \geq 1$ and consider arbitrary covers $\mathcal{V} \subset \mathcal{U}^m$, $\mathcal{G} \subset \mathcal{U}^n$ of X. If $U \in \mathcal{V}$ and $V \in \mathcal{G}$ then

$$S_{m+n}\phi(U \cap T^{-m}(V)) \leq S_m\phi(U) + S_n\phi(V)$$

and thus

$$\exp\bigl(S_{m+n}\phi(U \cap T^{-m}(V))\bigr) \leq \exp S_m\phi(U) \exp S_n\phi(V)$$

Since $U \cap T^{-m}(V) \in \mathcal{V} \vee T^{-m}(\mathcal{G}) \subset \mathcal{U}^m \vee T^{-m}(\mathcal{U}^n) = \mathcal{U}^{m+n}$, we therefore obtain

$$\begin{aligned} Z_{m+n}(\phi, \mathcal{U}) &\leq \sum_{U\in\mathcal{V}} \sum_{V\in\mathcal{G}} \exp\bigl(S_{m+n}f(U \cap T^{-m}(V))\bigr) \leq \sum_{U\in\mathcal{V}} \sum_{V\in\mathcal{G}} \exp S_m\phi(U) \exp S_n\phi(V) \\ &= \sum_{U\in\mathcal{V}} \exp S_m\phi(U) \times \sum_{V\in\mathcal{G}} \exp S_n\phi(V) \end{aligned} \tag{2.2.4}$$

Ranging now over all $\mathcal{V}$ and $\mathcal{G}$ as specified in definition (2.2.3) we get $Z_{m+n}(\phi,\mathcal{U}) \leq Z_m(\phi,\mathcal{U}) \cdot Z_n(\phi,\mathcal{U})$ which implies that

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

Moreover, $Z_n(\phi,\mathcal{U}) \geq \exp(-n||\phi||_\infty)$. So, $\log Z_n(\phi,\mathcal{U}) \geq -n||\phi||_\infty$ and applying now Lemma 1.4.3 finishes the proof. ♣

Notice that, although in the notation $\mathrm{P}(\phi,\mathcal{U})$, the transformation T does not directly appear, however this quantity depends obviously also on T. If we want to indicate this dependence we write $\mathrm{P}(T,\phi,\mathcal{U})$ and similarly $Z_n(T,\phi,\mathcal{U})$ for $Z_n(\phi,\mathcal{U})$. Given an open cover $\mathcal{V}$ of X let

$$\mathrm{osc}(\phi,\mathcal{V}) = \sup_{V\in\mathcal{V}} \big(\sup\{|\phi(x)-\phi(y)| : x,y \in V\}\big).$$

Lemma 2.2.2. If $\mathcal{U}$ and $\mathcal{V}$ are finite open covers of X such that $\mathcal{U} \succ \mathcal{V}$, then $\mathrm{P}(\phi,\mathcal{U}) \geq \mathrm{P}(\phi,\mathcal{V}) - \mathrm{osc}(\phi,\mathcal{V})$.

Proof. Take $U \in \mathcal{U}^n$. Then there exists $V = i(U) \in \mathcal{V}^n$ such that $U \subset V$. For every $x, y \in V$ we have $|S_n\phi(x) - S_n\phi(y)| \leq \mathrm{osc}(\phi,\mathcal{V})n$ and therefore

$$S_n\phi(U) \geq S_n\phi(V) - \mathrm{osc}(\phi,\mathcal{V})n \tag{2.2.5}$$

Let now $\mathcal{G} \subset \mathcal{U}^n$ be a cover of X and let $\tilde{\mathcal{G}} = \{i(U) : U \in \mathcal{U}^n\}$. The family $\tilde{\mathcal{G}}$ is also an open finite cover of X and $\tilde{\mathcal{G}} \subset \mathcal{V}^n$. In view of (2.2.5) and (2.2.3) we get

$$\sum_{U\in\mathcal{G}} \exp S_n\phi(U) \geq \sum_{V\in\tilde{\mathcal{G}}} \exp S_n\phi(V) e^{-\mathrm{osc}(\phi,\mathcal{V})n} \geq e^{-\mathrm{osc}(\phi,\mathcal{V})n} Z_n(\phi,\mathcal{V})$$

Therefore applying (2.2.3) again, we get $Z_n(\phi,\mathcal{U}) \geq \exp(-\mathrm{osc}(\phi,\mathcal{V})n) Z_n(\phi,\mathcal{V})$. Hence $\mathrm{P}(\phi,\mathcal{U}) \geq \mathrm{P}(\phi,\mathcal{V}) - \mathrm{osc}(\phi,\mathcal{V})$.

Definition 2.2.3. Consider now the family of all sequences $\{\mathcal{V}_n : n = 1,2,\ldots\}$ of open finite covers of X such that

$$\lim_{n\to\infty} \mathrm{diam}(\mathcal{V}_n) = 0 \tag{2.2.6}$$

and define the topological pressure $\mathrm{P}(T,\phi)$ as the supremum of upper limits

$$\limsup_{n\to\infty} \mathrm{P}(\phi,\mathcal{V}_n)$$

taken over all such sequences. Notice that by Lemma 2.2.1, $\mathrm{P}(T,\phi) \geq -||\phi||_\infty$.

The following lemma gives us a simpler way to calculate topological pressure showing that in fact in its definition we do not have to take the supremum.

Lemma 2.2.4. If $\{\mathcal{U}_n : n = 1,2,\ldots\}$ is a sequence of open finite covers of X such that $\lim_{n\to\infty} \mathrm{diam}(\mathcal{U}_n) = 0$, then the limit $\lim_{n\to\infty} \mathrm{P}(\phi,\mathcal{U}_n)$ exists and equals $\mathrm{P}(T,\phi)$.

Proof. Assume first that $\mathrm{P}(T,\phi)$ is finite and fix $\varepsilon > 0$. By the definition of pressure and uniform continuity of ϕ there exists $\mathcal{W}$, an open cover of X, such that

$$\operatorname{osc}(\phi,\mathcal{W}) \le \frac{\varepsilon}{2} \quad\text{and}\quad \mathrm{P}(\phi,\mathcal{W}) \ge \mathrm{P}(T,\phi) - \frac{\varepsilon}{2}. \tag{2.2.7}$$

Fix now $q \ge 1$ so large that for all $n \ge q$, $\operatorname{diam}(\mathcal{U}_n)$ does not exceed a Lebesgue number of the cover $\mathcal{W}$. Take $n \ge q$. Then $\mathcal{U}_n \succ \mathcal{W}$ and applying (2.2.7) and Lemma 2.2.2 we get

$$\mathrm{P}(\phi,\mathcal{U}_n) \ge \mathrm{P}(\phi,\mathcal{W}) - \frac{\varepsilon}{2} \ge \mathrm{P}(T,\phi) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = \mathrm{P}(T,\phi) - \varepsilon. \tag{2.2.8}$$

Hence, letting $\varepsilon \to 0$, $\liminf_{n\to\infty} \mathrm{P}(\phi,\mathcal{U}_n) \ge \mathrm{P}(T,\phi)$. This finishes the proof in the case of finite pressure $\mathrm{P}(T,\phi)$. Notice also that actually the same proof goes through in the infinite case. ♣

Since in the definition of numbers $\mathrm{P}(\phi,\mathcal{U})$ no metric is involved, they do not depend on a compatible metric under consideration. And since also the convergence to zero of diameters of a sequence of subsets of X does not depend on a compatible metric, we come to the conclusion that the topological pressure $\mathrm{P}(T,\phi)$ is independent of any compatible metric (depends of course on topology).

The reader familiar with directed sets will notice easily that the family of all finite open covers $\mathcal{U}$ of X equipped with the relation "$\prec$" is a directed set and topological pressure $\mathrm{P}(T,\phi)$ is the limit of the generalized sequence $\mathrm{P}(\phi,\mathcal{U})$. However we can assure him/her that this remark will not be used anywhere in this book.

If the funcion ϕ is identically zero, the pressure $\mathrm{P}(T,\phi)$ is usually called *topological entropy* of the map T and is denoted by $\mathrm{h}_{\mathrm{top}}(T)$.

In the rest of this section we establish basic elementary properties of pressure and provide its more effective characterizations. Applying Lemma 2.2.2 we obtain

Corollary 2.2.5. If $\mathcal{U}$ is a finite, open cover of X, then $\mathrm{P}(T,\phi) \ge \mathrm{P}(\phi,\mathcal{U}) - \operatorname{osc}(\phi,\mathcal{U})$.

Lemma 2.2.6. $\mathrm{P}(T^n, S_n\phi) = n\mathrm{P}(T,\phi)$ for every $n \ge 1$. In particular $\mathrm{h}_{\mathrm{top}}(T^n) = n\mathrm{h}_{\mathrm{top}}(T)$.

Proof. Put $g = S_n\phi$. Take $\mathcal{U}$, a finite open cover of X. Let $\overline{\mathcal{U}} = \mathcal{U} \vee T^{-1}(\mathcal{U}) \vee \ldots \vee T^{-(n-1)}(\mathcal{U})$. Since now we actually deal with two separate transformations T and T^n, we do not use the symbol $\mathcal{U}^n$ just to avoid possible misunderstandings. For any $m \ge 1$ consider an open set $U \in \mathcal{U}\vee T^{-1}(\mathcal{U})\vee\ldots\vee T^{-(nm-1)}(\mathcal{U}) = \overline{\mathcal{U}}\vee T^{-n}(\overline{\mathcal{U}})\vee\ldots\vee T^{-n(m-1)}(\overline{\mathcal{U}})$. Then for every $x \in U$ we have

$$\sum_{k=0}^{mn-1} \phi\circ T^k(x) = \sum_{k=0}^{m-1} g\circ T^{nk}(x)$$

and therefore $S_{mn}\phi(U) = S_m g(U)$, where the symbol S_m is considered with respect to the map T^n. Hence $Z_{mn}(T,\phi,\mathcal{U}) = Z_m(T^n, g, \overline{\mathcal{U}})$ and this implies that $\mathrm{P}(T^n, g, \overline{\mathcal{U}}) =$

$n\mathrm{P}(T,\phi,\mathcal{U})$. Since given a sequence $\mathcal{U}_k$ of open covers of X whose diameters converge to zero, the diameters of the sequence of its refinements $\overline{\mathcal{U}_k}$ also converge to zero, applying now Lemma 2.2.4 finishes the proof. ♣

Lemma 2.2.7. If $T: X \to X$ and $S: Y \to Y$ are continuous mappings of compact metric spaces and $\pi: X \to Y$ is a continuous surjection such that $S \circ \pi = \pi \circ T$, then for every continuous function $\phi: Y \to I\!R$ we have $\mathrm{P}(S,\phi) \le \mathrm{P}(T,\phi\circ\pi)$.

Proof. For every finite, open cover $\mathcal{U}$ of Y we get

$$\mathrm{P}(S,\phi,\mathcal{U}) = \mathrm{P}(T,\phi\circ\pi,\pi^{-1}(\mathcal{U})). \tag{2.2.9}$$

In view of Corollary 2.2.5 we have

$$\mathrm{P}(T,\phi\circ\pi) \ge \mathrm{P}(T,\phi\circ\pi,\pi^{-1}(\mathcal{U})) - \mathrm{osc}(\phi\circ\pi,\pi^{-1}(\mathcal{U})) = \mathrm{P}(T,\phi\circ\pi,\pi^{-1}(\mathcal{U})) - \mathrm{osc}(\phi,\mathcal{U}) \tag{2.2.10}$$

Let $\{\mathcal{U}_n : n = 1,2,\ldots\}$, be a sequence of open finite covers of Y whose diameters converge to 0. Then also $\lim_{n\to\infty}\mathrm{osc}(\phi,\mathcal{U}_n)) = 0$ and therefore, using Lemma 2.2.4, (2.2.9) and (2.2.10) we obtain

$$\mathrm{P}(S,\phi) = \lim_{n\to\infty}\mathrm{P}(S,\phi,\mathcal{U}_n) = \lim_{n\to\infty}\mathrm{P}(T,\phi\circ\pi,\pi^{-1}(\mathcal{U}_n)) \le \mathrm{P}(T,\phi\circ\pi)$$

The proof is finished. ♣

In the sequel we will need the following technical result.

Lemma 2.2.8. If $\mathcal{U}$ is a finite open cover of X then $\mathrm{P}(\phi,\mathcal{U}^k) = \mathrm{P}(\phi,\mathcal{U})$ for every $k \ge 1$.

Proof. Fix $k \ge 1$ and let $\gamma = \sup\{|S_{k-1}\phi(x)| : x \in X\}$. Since $S_{k+n-1}\phi(x) = S_n\phi(x) + S_{k-1}\phi(T^n(x))$ for every $n \ge 1$ and $x \in X$ we get

$$S_n\phi(x) - \gamma \le S_{k+n-1}\phi(x) \le S_n\phi(x) + \gamma$$

and therefore for every $n \ge 1$ and every $U \in \mathcal{U}^{k+n-1}$

$$S_n\phi(U) - \gamma \le S_{k+n-1}\phi(U) \le S_n\phi(U) + \gamma$$

Since $(\mathcal{U}^k)^n = \mathcal{U}^{k+n-1}$, these inequalities imply that

$$e^{-\gamma}Z_n(\phi,\mathcal{U}^k) \le Z_{n+k-1}(\phi,\mathcal{U}) \le e^{\gamma}Z_n(\phi,\mathcal{U}^k).$$

Letting now $n \to \infty$, we get the result required. ♣

§2.2a PRESSURE ON COMPACT METRIC SPACES

Let ρ is a metric on X. For every $n \ge 1$ we define on X the new metric ρ_n by putting

$$\rho_n(x,y) = \max\{\rho(T^j(x),T^j(y)) : j = 0,1,\ldots,n-1\}$$

Given $r > 0$ and $x \in X$ by $B_n(x,r)$ we denote the open ball in the metric ρ_n centered at x and of radius r. Let $\varepsilon > 0$ and let $n \geq 1$ be an integer. A set $F \subset X$ is said to be (n,ε)-*spanning* if and only if the family of balls $\{B_n(x,\varepsilon) : x \in F\}$ covers the space X. A set $S \subset X$ is said to be (n,ε)-*separated* if and only if $\rho_n(x,y) \geq \varepsilon$ for any pair x, y of different points in S. The following fact is obvious.

Lemma 2.2.9. Every maximal in the sense of inclusion (n,ε)-separated set forms an (n,ε)-spanning set.

We would like to emphasize here that the world maximal refering to separated sets will be in this book always understood in the sense of inclusion and not in the sense of cardinality. We finish this section with the following characterization of pressure.

Theorem 2.2.10. For every $\varepsilon > 0$ and every $n \geq 1$ let $F_n(\varepsilon)$ be a maximal (n,ε)-separated set in X. Then

$$\mathrm{P}(T,\phi) = \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\sum_{x\in F_n(\varepsilon)}\exp S_n\phi(x) = \lim_{\varepsilon\to 0}\liminf_{n\to\infty}\frac{1}{n}\log\sum_{x\in F_n(\varepsilon)}\exp S_n\phi(x)$$

Proof. Fix $\varepsilon > 0$ and let $\mathcal{U}(\varepsilon)$ be a finite cover of X by open balls of radii $\varepsilon/2$. For any $n \geq 1$ consider $\underline{\mathcal{U}}$, a subcover of $\mathcal{U}(\varepsilon)^n$ such that

$$Z_n(\phi,\mathcal{U}(\varepsilon)) = \sum_{U\in\underline{\mathcal{U}}}\exp S_n\phi(U),$$

where $Z_n(\phi,\mathcal{U}(\varepsilon))$ was defined by formula (2.2.3). For every $x \in F_n(\varepsilon)$ let $U(x)$ be an element of $\underline{\mathcal{U}}$ containing x. Since $F_n(\varepsilon)$ is an (n,ε)-separated set, we deduce that the function $x \mapsto U(x)$ is injective. Therefore

$$Z_n(\phi,\mathcal{U}(\varepsilon)) = \sum_{U\in\underline{\mathcal{U}}}\exp S_n\phi(U) \geq \sum_{x\in F_n(\varepsilon)}\exp S_n\phi(U(x)) \geq \sum_{x\in F_n(\varepsilon)}\exp S_n\phi(x).$$

Thus by Lemma 2.2.1

$$\mathrm{P}(\phi,\mathcal{U}(\varepsilon)) \geq \limsup_{n\to\infty}\frac{1}{n}\log\sum_{x\in F_n(\varepsilon)}\exp S_n\phi(x).$$

Hence, letting $\varepsilon \to 0$ and applying Corollary 2.2.4 we get

$$\mathrm{P}(T,\phi) \geq \limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\sum_{x\in F_n(\varepsilon)}\exp S_n\phi(x). \tag{2.2.11}$$

Let now $\mathcal{V}$ be an arbitrary finite open cover of X and let $\delta > 0$ be a Lebesgue number of $\mathcal{V}$. Take $\varepsilon < \delta/2$. Since for any $k = 0, 1, \ldots, n-1$ and for every $x \in F_n(\varepsilon)$

$$\mathrm{diam}\big(T^k(B_n(x,\varepsilon))\big) \leq 2\varepsilon < \delta,$$

we conclude that for some $U_k(x) \in \mathcal{V}$

$$T^k(B_n(x,\varepsilon)) \subset U_k(x)$$

Since the family $\{B_n(x,\varepsilon) : x \in F_n(\varepsilon)\}$ covers X (by Lemma 2.2.9), it implies that the family $\{U(x) : x \in F_n(\varepsilon)\} \subset \mathcal{V}^n$ also covers X, where $U(x) = U_0(x) \cap T^{-1}(U_1(x)) \cap \ldots \cap T^{-(n-1)}(U_{n-1}(x))$. Therefore

$$Z_n(\phi,\mathcal{V}) \le \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(U(x)) \le \exp\bigl(\mathrm{osc}(\phi,\mathcal{V})n\bigr) \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(x)$$

Hence

$$\mathrm{P}(\phi,\mathcal{V}) \le \mathrm{osc}(\phi,\mathcal{V}) + \liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(x)$$

and consequently

$$\mathrm{P}(\phi,\mathcal{V}) - \mathrm{osc}(\phi,\mathcal{V}) \le \liminf_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(x).$$

Letting $\mathrm{diam}(\mathcal{V}) \to 0$ and applying Corollary 2.2.4 we get

$$\mathrm{P}(T,\phi) \le \liminf_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(x).$$

Combining this and (2.2.11) finishes the proof. ♣

Frequently the limits

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(x)$$

and

$$\liminf_{n\to\infty} \frac{1}{n} \log \sum_{x\in F_n(\varepsilon)} \exp S_n\phi(x)$$

will be denoted respectively by $\overline{\mathrm{P}}(T,\phi,\varepsilon)$ and $\underline{\mathrm{P}}(T,\phi,\varepsilon)$. Actually these depend also on the sequence $\{F_n(\varepsilon) : n = 1,2,\ldots\}$ of maximal (n,ε)-separated sets under consideration. However it will be always clear from the context which such sequence is meant.

§2.3 VARIATIONAL PRINCIPLE

In this section we shall prove the following theorem, called *variational principle*, which has a long history and which establishes an interesting relationship between measure-theoretic dynamics and topological dynamics.

Theorem 2.3.1. (Variational principle) If $T : X \to X$ is a continuous transformation of a compact metric space X and $\phi : X \to I\!R$ is a continuous function then

$$\mathrm{P}(T,\phi) = \sup\left\{\mathrm{h}_\mu(T) + \int \phi\, d\mu : \mu \in M(T)\right\},$$

where $M(T)$ denotes the set of all Borel probability T-invariant measures on X.

The proof of this theorem consists of two parts. In the Part I we show that $\mathrm{h}_\mu(T)+\int \phi\, d\mu \le \mathrm{P}(T,\phi)$ for every measure $\mu \in M(T)$ and the Part II is devoted to proving inequality $\sup\{\mathrm{h}_\mu(T) + \int \phi\, d\mu : \mu \in M(T)\} \ge \mathrm{P}(T,\phi)$.

Proof of Part I. Let $\mu \in M(T)$. Fix $\varepsilon > 0$ and consider a finite partition $\mathcal{U} = \{A_1, \dots, A_s\}$ of X into Borel sets. One can find compact sets $B_i \subset A_i$, $i = 1, 2, \dots, s$, such that for the partition $\mathcal{V} = \{B_1, \dots, B_s, X \setminus (B_1 \cup \ldots \cup B_s)\}$ we have

$$\mathrm{H}_\mu(\mathcal{U}|\mathcal{V}) \le \varepsilon,$$

where the conditional entropy $\mathrm{H}_\mu(\mathcal{U}|\mathcal{V})$ has been defined in (1.3.3).

Therefore, as in the proof of Theorem 1.4.4 (d) we get for every $n \ge 1$

$$\mathrm{H}_\mu(\mathcal{U}^n) \le \mathrm{H}_\mu(\mathcal{V}^n) + n\varepsilon. \tag{2.3.1}$$

Our first aim is to estimate from above the number $\mathrm{H}_\mu(\mathcal{V}^n) + \int S_n\phi\, d\mu$. Putting $b_n = \sum_{B\in\mathcal{V}^n} \exp S_n\phi(B)$, keeping notation $k(x) = -x\log x$ and using concavity of the function logarithm we obtain by Jensen inequality

$$\begin{aligned}
\mathrm{H}_\mu(\mathcal{V}^n) + \int S_n\phi\, d\mu &\le \sum_{B\in\mathcal{V}^n} \mu(B)\big(S_n\phi(B) - \log\mu(B)\big) \\
&= \sum_{B\in\mathcal{V}^n} \mu(B)\log\big(e^{S_n\phi(B)}/\mu(B)\big) \\
&\le \log\Big(\sum_{B\in\mathcal{V}^n} e^{S_n\phi(B)}\Big) \qquad (2.3.2)
\end{aligned}$$

(compare the Lemma in Introduction).

Take now $0 < \delta < \frac{1}{2}\inf\{\rho(B_i, B_j) : 1 \le i \ne j \le s\} > 0$ so small that

$$|\phi(x) - \phi(y)| < \varepsilon \tag{2.3.3}$$

if $\rho(x,y) < \delta$. Consider an arbitrary maximal (n,δ)-separated set $E_n(\delta)$. Fix $B \in \mathcal{V}^n$. Then, by Lemma 2.2.9, for every $x \in B$ there exists $y \in E_n(\delta)$ such that $x \in B_n(y,\delta)$, whence $|S_n\phi(x) - S_n\phi(y)| \le \varepsilon n$ by (2.3.3). Therefore, using finiteness of the set $E_n(\delta)$, we see that there exists $y(B) \in E_n(\delta)$ such that

$$S_n\phi(B) \le S_n\phi(y(B)) + \varepsilon n \tag{2.3.4}$$

and

$$B \cap B_n(y(B), \delta) \neq \emptyset.$$

The definitions of δ and partition $\mathcal{V}$ imply that for every $z \in X$

$$\#\{B \in \mathcal{V} : B \cap B(z, \delta) \neq \emptyset\} \leq 2$$

Thus

$$\#\{B \in \mathcal{V}^n : B \cap B_n(z, \delta) \neq \emptyset\} \leq 2^n$$

Therefore the function $\mathcal{V}^n \ni B \mapsto y(B) \in E_n(\delta)$ is at most 2^n to 1. Hence, using (2.3.4),

$$2^n \sum_{y \in E_n(\delta)} \exp S_n\phi(y) \geq \sum_{B \in \mathcal{V}^n} \exp\big(S_n\phi(B) - \varepsilon n\big) = e^{-\varepsilon n} \sum_{B \in \mathcal{V}^n} \exp S_n\phi(B)$$

Taking now the logarithms of both sides of this inequality, dividing them by n and applying (2.3.2), we get

$$\begin{aligned}
\log 2 + \frac{1}{n} \log\big(\sum_{y \in E_n(\delta)} \exp S_n\phi(y)\big) &\geq -\varepsilon + \frac{1}{n} \log\big(\sum_{B \in \mathcal{V}^n} \exp S_n\phi(B)\big) \\
&\geq \frac{1}{n} \mathrm{H}_\mu(\mathcal{V}^n) + \frac{1}{n} \int S_n\phi \, d\mu - \varepsilon.
\end{aligned}$$

So, by (2.3.1),

$$\frac{1}{n} \log\big(\sum_{y \in E_n(\delta)} \exp S_n\phi(y)\big) \geq \frac{1}{n} \mathrm{H}_\mu(\mathcal{U}^n) + \int \phi \, d\mu - (2\varepsilon + \log 2).$$

In view of the definition of entropy $\mathrm{h}_\mu(T, \mathcal{U})$ presented just after Lemma 1.4.2, by letting $n \to \infty$, we get

$$\underline{\mathrm{P}}(T, \phi, \delta) \geq \mathrm{h}_\mu(T, \mathcal{U}) + \int \phi \, d\mu - (2\varepsilon + \log 2).$$

Applying now Theorem 2.2.10 with $\delta \to 0$ and next letting $\varepsilon \to 0$ and taking supremum over all Borel partitions $\mathcal{U}$ lead us to the following

$$\mathrm{P}(T, \phi) \geq \mathrm{h}_\mu(T) + \int \phi \, d\mu - \log 2.$$

And applying with every $n \geq 1$ this estimate to the transformation T^n and the function $S_n\phi$ we obtain

$$\mathrm{P}(T^n, S_n\phi) \geq \mathrm{h}_\mu(T^n) + \int S_n\phi \, d\mu - \log 2$$

or equivalently, by Lemma 2.2.6 and Theorem 1.4.6(a)

$$n\mathrm{P}(T, \phi) \geq n\mathrm{h}_\mu(T) + n \int \phi \, d\mu - \log 2$$

Dividing both sides of this inequality by n and letting then $n \to \infty$, the proof of Part I follows. ♣

In the proof of part II we will need the following two lemmas.

Lemma 2.3.2. If μ is a Borel probability measure on X, then for every $\varepsilon > 0$ there exists a finite partition $\mathcal{A}$ such that $\mathrm{diam}(\mathcal{A}) \le \varepsilon$ and $\mu(\partial A) = 0$ for every $A \in \mathcal{A}$.

Proof. Let $E = \{x_1, \ldots, x_s\}$ be an $\varepsilon/4$-spanning set (that is with respect to the metric $\rho = \rho_0$) of X. Since for every $i \in \{1, \ldots, s\}$ the sets $\{x : \rho(x, x_i) = r\}$, $\varepsilon/4 < r < \varepsilon/2$, are closed and mutually disjoint, only countably many of them can have positive measure μ. Hence, there exists $\varepsilon/4 < t < \varepsilon/2$ such that for every $i \in \{1, \ldots, s\}$

$$\mu(\{x : \rho(x, x_i) = t\}) = 0 \tag{2.3.5}$$

Define inductively the sets $A_1, A_2, \ldots, A_s$ putting $A_1 = \{x : \rho(x, x_1) \le t\}$ and for every $i = 2, 3, \ldots, s$

$$A_i = \{x : \rho(x, x_i) \le t\} \setminus (A_1 \cup A_2 \cup \ldots \cup A_{i-1})$$

The family $\mathcal{U} = \{A_1, \ldots, A_s\}$ is a partition of X with diameter not exceeding ε. Using (2.3.5) and noting that generally $\partial(A \setminus B) \subset \partial A \cup \partial B$, we conclude by induction that $\mu(\partial A_i) = 0$ for every $i = 1, 2, \ldots, s$. ♣

Proof of Part II. Fix $\varepsilon > 0$ and let $E_n(\varepsilon)$, $n = 1, 2, \ldots$, be a sequence of maximal (n, ε)-separated set in X. For every $n \ge 1$ define measures

$$\mu_n = \frac{\sum_{x \in E_n(\varepsilon)} \delta_x \exp S_n\phi(x)}{\sum_{x \in E_n(\varepsilon)} \exp S_n\phi(x)}, \quad m_n = \frac{1}{n} \sum_{k=0}^{n-1} \mu_n \circ T^{-k}$$

where δ_x denotes the Dirac measure concentrated at the point x (see (2.1.2)). Let $\{n_i, i = 1, 2, \ldots\}$ be an increasing sequence such that m_{n_i} converges weakly, say to m and

$$\lim_{i \to \infty} \frac{1}{n_i} \log \sum_{x \in E_{n_i}(\varepsilon)} \exp S_n\phi(x) = \limsup_{n \to \infty} \frac{1}{n} \log \sum_{x \in E_n(\varepsilon)} \exp S_n\phi(x) \tag{2.3.9}$$

Clearly $m \in M(T)$. In view of Lemma 2.3.2 there exists a finite partition γ such that $\mathrm{diam}(\gamma) \le \varepsilon$ and $\mu(\partial G) = 0$ for every $G \in \gamma$. For any $n \ge 1$ put $g_n = \sum_{x \in E_n(\varepsilon)} \exp S_n\phi(x)$. Since $\#(G \cap E_n(\varepsilon)) \le 1$ for every $G \in \gamma^n$, we obtain

$$\begin{aligned} \mathrm{H}_{\mu_n}(\gamma^n) + \int S_n\phi \, d\mu_n &= \sum_{x \in E_n(\varepsilon)} \big(-\log \mu_n(x) + S_n\phi(x)\big)\mu_n(x) \\ &= \sum_{x \in E_n(\varepsilon)} \frac{\exp S_n\phi(x)}{g_n} \big(S_n\phi(x) - \log\big(\frac{\exp S_n\phi(x)}{g_n}\big)\big) \\ &= g_n^{-1} \sum_{x \in E_n(\varepsilon)} \exp S_n\phi(x)\big(S_n\phi(x) - S_n\phi(x) + \log g_n\big) = \log g_n \end{aligned} \tag{2.3.10}$$

Fix now $M \in I\!N$ and $n \geq 2M$. For $j = 0, 1, \dots, M-1$ let $s(j) = \mathrm{E}(\frac{n-j}{M}) - 1$, where $\mathrm{E}(x)$ denotes the integer part of x. Note that

$$\bigvee_{k=0}^{s(j)} T^{-(kM+j)}\gamma^M = T^{-j}\gamma \vee \ldots \vee T^{-(s(j)M+j)-(M-1)}\gamma = T^{-j}\gamma \vee \ldots \vee T^{-((s(j)+1)M+j-1)}\gamma$$

and

$$(s(j)+1)M + j - 1 \leq n - j + j - 1 = n - 1$$

Therefore, setting $R_j = \{0, 1, \dots, j-1, (s(j)+1)M + j, \dots, n-1\}$, we can write

$$\gamma^n = \bigvee_{k=0}^{s(j)} T^{-(kM+j)}\gamma^M \vee \bigvee_{i \in R_j} T^{-i}\gamma.$$

Hence

$$\begin{aligned}
\mathrm{H}_{\mu_n}(\gamma^n) &\leq \sum_{k=0}^{s(j)} \mathrm{H}_{\mu_n}\big(T^{-(kM+j)}\gamma^M\big) + \mathrm{H}_{\mu_n}\big(\bigvee_{i \in R_j} T^{-i}\gamma\big) \\
&\leq \sum_{k=0}^{s(j)} \mathrm{H}_{\mu_n \circ T^{-(kM+j)}}(\gamma^M) + \log\big(\#(\bigvee_{i \in R_j} T^{-i}\gamma)\big).
\end{aligned}$$

Summing now over all $j = 0, 1, \dots, M-1$ we then get

$$\begin{aligned}
M\mathrm{H}_{\mu_n}(\gamma^n) &\leq \sum_{j=0}^{M-1}\sum_{k=0}^{s(j)} \mathrm{H}_{\mu_n \circ T^{-(kM+j)}}(\gamma^M) + \sum_{j=0}^{M-1} \log\big(\#\gamma^{\# R_j}\big) \\
&\leq \sum_{l=0}^{n-1} \mathrm{H}_{\mu_n \circ T^{-l}}(\gamma^M) + 2M^2 \log\#\gamma \leq n\mathrm{H}_{\frac{1}{n}\sum_{l=0}^{n-1} \mu_n \circ T^{-l}}(\gamma^M) + 2M^2 \log\#\gamma.
\end{aligned}$$

And applying (2.3.10) we obtain

$$M \log\big(\sum_{x \in E_n(\varepsilon)} \exp S_n\phi(x)\big) \leq n\mathrm{H}_{m_n}(\gamma^M) + M\int S_n\phi \, d\mu_n + 2M^2 \log\#\gamma$$

Dividing both sides of this inequality by Mn, we get

$$\frac{1}{n}\log\big(\sum_{x \in E_n(\varepsilon)} \exp S_n\phi(x)\big) \leq \frac{1}{M}\mathrm{H}_{m_n}(\gamma^M) + \int \phi \, dm_n + 2\frac{M}{n}\log\#\gamma.$$

Since $\partial T^{-1}(A) \subset T^{-1}(\partial A)$ for every set $A| \subset X$, the measure m of the boundaries of the partition γ^M is equal to 0. Letting therefore $n \to \infty$ along the subsequence $\{n_i\}$ we conclude from this inequality, (2.3.7) and Lemma 2.1.15 that

$$\overline{\mathrm{P}}(T, \phi, \varepsilon) \leq \frac{1}{M}\mathrm{H}_m(\gamma^M) + \int \phi \, dm.$$

Now letting $M \to \infty$ we get

$$\overline{\mathrm{P}}(T,\phi,\varepsilon) \le \mathrm{h}_m(T,\gamma) + \int \phi\, dm \le \sup\left\{h_\mu(T) + \int \phi\, d\mu : \mu \in M(T)\right\}.$$

Applying finally Theorem 2.2.10 and letting $\varepsilon \searrow 0$, we get the desired inequality. ♣

Corollary 2.3.4. Under assumptions of Theorem 2.3.1

$$\mathrm{P}(T,\phi) = \sup\{\mathrm{h}_\mu(T) + \int \phi\, d\mu : \mu \in M_e(T)\},$$

where $M_e(T)$ denotes the set of all Borel ergodic probability T-invariant measures on X.

Proof. Let $\mu \in M(T)$ and let $\{\mu_x : x \in X\}$ be the ergodic decomposition of μ. Then $\mathrm{h}_\mu = \int \mathrm{h}_{\mu_x}\, d\mu(x)$ and $\int \phi\, d\mu = \int(\int \phi\, d\mu_x)\, d\mu(x)$. Therefore

$$\mathrm{h}_\mu + \int \phi\, d\mu = \int \left(\mathrm{h}_{\mu_x} + \int \phi\, d\mu_x\right) d\mu(x)$$

and consequently there exists $x \in X$ such that $\mathrm{h}_{\mu_x} + \int \phi\, d\mu_x \ge \mathrm{h}_\mu + \int \phi\, d\mu$ which finishes the proof. ♣

Corollary 2.3.5. If $T : X \to X$ is a continuous transformation of a compact metric space X, $\phi : X \to I\!\!R$ is a continuous function and Y is a forward invariant subset of X (i.e. $T(Y) \subset Y$), then $\mathrm{P}(T|_Y, \phi|_Y) \le \mathrm{P}(T,\phi)$.

Proof. The proof follows immediatly from Theorem 2.3.1 by the remark that each $T|_Y$-invariant measure on Y can be treated as a measure on X and is T-invariant. ♣

§2.4 EQUILIBRIUM STATES AND EXPANSIVE MAPS

We keep in this section the notation of the previous one. A measure $\mu \in M(T)$ is called an *equilibrium state* for the transformation T and function ϕ if $\mathrm{P}(T,\phi) = \mathrm{h}_\mu(T) + \int \phi\, d\mu$. The set of all those measures will be denoted by $E(\phi)$. In the case $\phi = 0$ the equilibrium states are also called as *maximal measures*. Similarly (in fact even easier) as Corollary 2.3.5 one can prove the following.

Proposition 2.4.1 If $E(\phi) \ne \emptyset$ then $E(\phi)$ contains ergodic measures.

As the following example shows there exist transformations and functions which admit no equilibrium states.

Example 2.4.2. Let $\{T_n : X_n \to X_n\}_{n\ge 1}$ be a sequence of continuous mappings of compact metric spaces X_n such that for every $n \ge 1$

(2.4.1) $$\mathrm{h}_{\mathrm{top}}(T_n) < \mathrm{h}_{\mathrm{top}}(T_{n+1}) \text{ and } \sup_n \mathrm{h}_{\mathrm{top}}(T_n) < \infty$$

The disjoint union $\oplus_{n=1}^{\infty} X_n$ of the spaces X_n is a locally compact space and let $X = \{\omega\} \cup \oplus_{n=1}^{\infty} X_n$ be its one-point (Alexandroff) compactification. Define the map $T : X \to X$ by $T|_{X_n} = T_n$ and $T(\omega) = \omega$. The reader will check easily that T is continuous. Suppose that μ is an ergodic maximal measure for T. Then $\mu(X_n) = 1$ for some $n \geq 1$ and therefore $\mathrm{h}_{\mathrm{top}}(T) = \mathrm{h}_\mu(T_n) \leq \mathrm{h}_{\mathrm{top}}(T_n)$ which contradicts formula (2.4.1) and Corollary 2.3.5. In view of Proposition 2.4.1 this shows that T has no maximal measure.

A more difficult problem is to find a transitive and smooth example (see for instance [Mi, 1973]).

The remaining part of this section is devoted to provide sufficient conditions for the existence of equilibrium states and we start with the following simple general criterion which will be the base to obtain all others.

Proposition 2.4.3. If the function $M(T) \ni \mu \to \mathrm{h}_\mu(T)$ is upper semi-continuous then each continuous function $\phi : X \to I\!\!R$ has an equilibrium state.

Proof. By the definition of weak topology the function $M(T) \ni \mu \to \int \phi \, d\mu$ is continuous. Therefore the lemma follows from the assumption, the sequential compactness of the space $M(T)$ and Theorem 2.3.1 (variational principle). ♣

As an immediate consequence of this lemma and Theorem 2.3.1 we obtain the following.

Corollary 2.4.4. If $\mathrm{h}_{\mathrm{top}}(T) = 0$ then each continuous function on X has an equilibrium state.

A continuous transformation $T : X \to X$ of a compact metric space X equipped with a metric ρ is said to be *(positively) expansive* if and only if

$$\exists \delta > 0 \; [\forall n \geq 0 \; \rho(T^n(x), T^n(y)) \leq \delta] \Longrightarrow \; x = y$$

and the number δ which has appeared in this definition is called an *expansive constant.*

Although at the end of this section we will introduce a related but different notion of expansiveness of homeomorphisms we will frequently omit the word "positively". Note that the property of being expansive does not depend on the choice of a metric compatible with the topology. From now on in this chapter the transformation T will be assumed to be positively expansive, unless stated otherwise. The following lemma is an immediate consequence of expansiveness.

Lemma 2.4.5. If $\mathcal{A}$ is a finite Borel partition of X with diameter not exceeding an expansive constant then $\mathcal{A}$ is a generator for every Borel probability T-measure μ on X.

The main result concerning expansive maps is the following.

Theorem 2.4.6. If $T : X \to X$ is positively expansive then the function $M(T) \ni \mu \to \mathrm{h}_\mu(T)$ is upper semi-continuous and consequently (by Lemma 2.4.3) each continuous function on X has an equilibrium state.

Proof. Let $\delta > 0$ be an expansive constant of T and let $\mu \in M(T)$. By Lemma 2.3.2 there exists a finite partition $\mathcal{A}$ of X such that $\mathrm{diam}(\mathcal{A}) \le \delta$ and $\mu(\partial A) = 0$ for every $A \in \mathcal{A}$. Thus in view of Lemma 2.4.5 and Theorem 1.8.7(b) $\mathrm{h}_\mu(T) = \mathrm{h}_\mu(T, \mathcal{A})$ whence by the definition of the entropy $\mathrm{h}_\mu(T, \mathcal{A})$ (cf. Lemma 1.4.2) there exists $m \ge 1$ such that

$$\frac{1}{m+1}\mathrm{H}_\mu(\mathcal{A}^m) \le \mathrm{h}_\mu(T) + \frac{\varepsilon}{2} \tag{2.4.2}$$

Consider now a sequence $\{\mu_n : n = 1, 2, \dots\}$ of invariant measures converging weakly to μ. By the definition of the entropy of partition, by Theorem 2.1.15 and by the choice of partition $\mathcal{A}$, $\lim_{n\to\infty} \mathrm{H}_{\mu_n}(\mathcal{A}^m) = \mathrm{H}_\mu(\mathcal{A}^m)$. Therefore there exists $n_0 \ge 1$ such that for every $n \ge n_0$

$$\frac{1}{m+1}\left|\mathrm{H}_{\mu_n}(\mathcal{A}^m) - \mathrm{H}_\mu(\mathcal{A}^m)\right| \le \frac{\varepsilon}{2}$$

Combining this and 2.4.2, and using Lemma 1.4.3 we get for every $n \ge n_0$

$$\mathrm{h}_{\mu_n}(T) = \mathrm{h}_{\mu_n}(T, \mathcal{A}) \le \frac{1}{m+1}\mathrm{H}_{\mu_n}(\mathcal{A}^m) \le \frac{1}{m+1}\mathrm{H}_\mu(\mathcal{A}^m) + \frac{\varepsilon}{2} \le \mathrm{h}_\mu(T) + \varepsilon$$

The proof is finished. ♣

Below we prove three additional interesting results about expansive maps.

Lemma 2.4.7. If $\mathcal{U}$ is a finite open cover of X with diameter not exceeding an expansive constant of an expansive map $T : X \to X$, then $\lim_{n\to\infty} \mathrm{diam}(\mathcal{U}^n) = 0$.

Proof. Let $\mathcal{U} = \{U_1, U_2, \dots, U_s\}$. By expansiveness for every sequence $\{a_n : n = 0, 1, 2, \dots\}$ of elements of the set $\{1, 2, \dots, s\}$

$$\#\Big(\bigcap_{n=0}^{\infty} T^{-n}(\overline{U_{a_n}})\Big) \le 1$$

and hence

$$\lim_{k\to\infty} \mathrm{diam}\Big(\bigcap_{n=0}^{k} T^{-n}(\overline{U_{a_n}})\Big) = 0$$

Therefore, given a fixed $\varepsilon > 0$ there exists a minimal finite $k = k(\{a_n\})$ such that

$$\mathrm{diam}\Big(\bigcap_{n=0}^{k} T^{-n}(\overline{U_{a_n}})\Big) < \varepsilon$$

Note now that the function $\{1, 2, \dots, s\}^{\mathbb{N}} \ni \{a_n\} \mapsto k(\{a_n\})$ is continuous, even more, it is locally constant. Thus, by compactness of the space $\{1, 2, \dots, s\}^{\mathbb{N}}$, it is bounded, say by t, and therefore

$$\mathrm{diam}(\mathcal{U}^n) < \varepsilon$$

for every $n \geq t$. The proof is finished. ♣

Combining now Lemma 2.2.4, Lemma 2.4.7 and Lemma 2.2.8 we get the following.

Proposition 2.4.8. If $\mathcal{U}$ is a finite open cover of X with diameter not exceeding an expansive constant then $\mathrm{P}(T, \phi) = \mathrm{P}(T, \phi, \mathcal{U})$.

As the last result of this section we shall prove the following.

Proposition 2.4.9. There exists a constant $\eta > 0$ such that $\forall\, 0 < \varepsilon < \eta \; \exists\, n(\varepsilon) \geq 1$

$$\rho(x, y) \geq \varepsilon \implies \rho_{n(\varepsilon)}(x, y) > \eta$$

Proof. Let $\mathcal{U} = \{U_1, U_2, \ldots, U_s\}$ be a finite open cover of X with diameter not exceeding an expansive constant δ and let η be a Lebesgue number of $\mathcal{U}$. Fix $\varepsilon > 0$. In view of Lemma 2.4.7 there exists an $n(\varepsilon) \geq 1$ such that

$$\operatorname{diam}(\mathcal{U}^{n(\varepsilon)}) < \varepsilon. \tag{2.4.3}$$

Let $\rho(x, y) \geq \varepsilon$ and suppose that $\rho_{n(\varepsilon)}(x, y) \leq \eta$. Then

$$\forall\, (0 \leq j \leq n(\varepsilon) - 1) \; \exists\, (U_{i_j} \in \mathcal{U}) \;\; T^j(x), T^j(y) \in U_{i_j}$$

and therefore

$$x, y \in \bigcap_{j=0}^{n(\varepsilon)-1} T^{-j}(U_{i_j}) \in \mathcal{U}^{n(\varepsilon)}$$

Hence $\operatorname{diam}(\mathcal{U}^{n(\varepsilon)}) \geq \rho(x, y) \geq \varepsilon$ which contradicts (2.4.3). The proof is finished. ♣

As we mentioned at the begining of this section there is a notion related to positive expansiveness which makes sense only for homeomorphisms. Namely we say that a homeomorphism $T : X \to X$ is *expansive* if and only if

$$\exists \delta > 0 \; [\forall n \in \boldsymbol{Z} \; \rho(T^n(x), T^n(y)) \leq \delta] \implies x = y$$

We will not explore this notion in our book – we only want to emphasize that for expansive homeomorphisms analogous results (with obvious modifications) can be proved (in the same way) as for positively expansive mappings. Of course each positively expansive homeomorphism is expansive.

§2.5 TOPOLOGICAL PRESSURE AS A FUNCTION ON THE BANACH SPACE OF CONTINUOUS FUNCTIONS. THE ISSUE OF UNIQUENESS OF EQUILIBRIUM STATES

Let $T : X \to X$ be a continuous mapping of a compact topological space X. We shall discuss here the topological pressure function $\mathrm{P} : C(X) \to I\!R$, $\mathrm{P}(\phi) = \mathrm{P}(T, \phi)$. Assume that the topological entropy is finite, $\mathrm{h_{top}}(T) < \infty$. Hence the pressure P is also finite, for example

$$\mathrm{P}(\phi) \leq \mathrm{h_{top}}(T) + ||\phi|||_\infty. \tag{2.5.1}$$

This estimate follows directly from the definitions, see Section 2. It is also an immediate consequence of Theorem 2.3.1 (Variational Principle) in case X is metrizable.

Let us start with the following easy

Theorem 2.5.1. The pressure function P is Lipschitz continuous with the Lipschitz constant 1.

Proof. Let $\phi \in C(X)$. Recall from Section 2.2 that in the definition of pressure we have considered the following partition function

$$Z_n(\phi, \mathcal{U}) = \inf_{\mathcal{V}} \{ \sum_{U \in \mathcal{V}} \exp S_n\phi(U) \},$$

where $\mathcal{V}$ ranges over all covers of X contained in $\mathcal{U}^n$. Now if also $\psi \in C(X)$, then we obtain for every open cover $\mathcal{U}$ and positive integer n

$$Z_n(\psi, \mathcal{U}) e^{-||\phi-\psi||_\infty n} \leq Z_n(\phi, \mathcal{U}) \leq Z_n(\psi, \mathcal{U}) e^{||\phi-\psi||_\infty n}$$

Taking limits if $n \nearrow \infty$ we get $\mathrm{P}(\psi) - ||\phi - \psi||_\infty \leq \mathrm{P}(\phi) \leq \mathrm{P}(\psi) + ||\phi - \psi||_\infty$, hence $|P(\psi) - P(\phi)| \leq ||\psi - \phi||_\infty$. ♣

Theorem 2.5.2. If X is metrizable, then the topological pressure function $\mathrm{P} : C(X) \to I\!R$ is convex.

we provide two different proofs of this important theorem. One elementary, the second relying on the variational principle (Theorem 2.3.1).

Proof 1. By Hölder inequality applied with the exponents $a = 1/\alpha, b = 1/(1-\alpha)$, so that $1/a + 1/b = \alpha + 1 - \alpha = 1$ we obtain for an arbitrary finite set $E \subset X$

$$\frac{1}{n} \log \sum_E e^{S_n(\alpha\phi) + S_n(1-\alpha)\psi)} = \frac{1}{n} \log \sum_E e^{\alpha S_n(\phi)} e^{(1-\alpha) S_n(\psi)} \leq$$

$$\frac{1}{n} \log \big(\sum_E e^{S_n(\phi)}\big)^\alpha \big(\sum_E e^{S_n(\psi)}\big)^{1-\alpha} \leq \alpha \frac{1}{n} \log \big(\sum_E e^{S_n(\phi)}\big) + (1-\alpha) \frac{1}{n} \log \big(\sum_E e^{S_n(\psi)}\big).$$

To conclude the proof use the definition of pressure *via* $E = F_n(\varepsilon)$ that is (n, ε)-separated sets, Theorem 2.2.10.

Proof 2. It is sufficient to prove that the function

$$\hat{P} := \sup_{\mu \in M(X,T)} L_\mu \phi = \mathrm{h}_\mu(T) + \mu\phi$$

is convex, because by variational principle $\hat{P}(\phi) = P(\phi)$.

We need to prove that the set

$$A := \{(\phi, y) \in C(X) \times I\!R\} : y \geq \hat{P}(\phi))$$

is convex. Observe however that that by its definition $A = \bigcap_{\mu \in M(X,T)} L_\mu^+$, where by L_μ^+ we denote the upper half space $\{(f, y) : y \geq L_\mu \phi\}$. All the halfspaces L_μ^+ are convex, hence A is convex as their intersection. ♣

Remark 2.5.3. We can write $L_\mu \phi = \mu\phi - (-h_\mu(T))$. The function $\hat{\mathrm{P}}(\phi) = \sup_{\mu \in M(T)} L_\mu \phi$ defined on the space $C(X)$ is called the Legendre-Fenchel transform of the convex function $\mu \mapsto -h_\mu(T)$ on the convex set $M(T)$. We shall abbreviate the name Legendre-Fenchel transform to LF-transform. Observe that this transform generalizes the standard Legendre transformation of a strictly convex function h on a finite dimensional linear space, say $I\!R^n$, $y \mapsto \sup_{x \in I\!R^n} < x, y > -h(x)$, where $< x, y >$ is the scalar (inner) product of x and y.

Note that $-h_\mu(T)$ is not strictly convex (unless $M(X,T)$ is a one element space) because it is affine, see Th.1.4.7.

Proof 2 just repeats the standard proof that Legendre transform is convex.

In the sequel we will need so called geometric form of the Hahn-Banach theorem (see [Bourbaki, Th.1, Ch.2.5] or Ch. 1.7 of [Edwards, 1995].

Theorem 2.5.4 (Hahn-Banach). Let A be an open convex non-empty subset of a real topological vector space V and let M be a non-empty affine subset of V (linear subspace moved by a vector) which does not meet A. Then there exists a codimension 1 closed affine subset H which contains M and does not meet A.

Suppose now that $P : V \to I\!R$ is an arbitrary convex continuous function on a real topological vector space V. We call a continuous linear functional $F : V \to I\!R$ *tangent* to P at $x \in V$ if

$$F(y) \leq P(x + y) - P(x) \tag{2.5.2}$$

for every $y \in V$. We denote the set of all such functionals by $V^*_{x,P}$. Sometimes the term *supporting functional* is being used in the literature.

Applying Theorem 2.5.4 we easily prove that for every x the set $V^*_{x,P}$ is non-empty. Indeed, we can consider the open convex set $A = \{(\phi, y) \in V \times I\!R\} : y > P(x)\}$ in the vector space $V \times I\!R$ with the product topology and the one-point set $M = \{x, P(x)\}$, and define a supporting functional we look for, as having the graph $H - \{x, P(x)\}$ in $V \times I\!R$.

We would also like to bring reader's attention to the following another general fact from functional analysis.

Theorem 2.5.5. Let V be a Banach space and $P : V \to I\!R$ be a convex continuous function. Then for every $x \in V$ the function P is differentiable at x in every direction (Gateaux differentiable), or in a dense in the weak topology set of directions, if and only if $V^*_{x,P}$ is a singleton.

Proof. Suppose first that P is not differentiable at some point x and direction y. Choose an arbitrary $F \in V^*_{x,P}$. Non-differentiability in the direction $y \in V$ implies that there exist $\varepsilon > 0$ and a sequence $\{t_n\}_{n\geq 1}$ converging to 0 such that

$$P(x + t_n y) - P(x) \geq t_n F(y) + \varepsilon |t_n|. \tag{2.5.3}$$

In fact we can assume that all t_n, $n \geq 1$, are positive by passing to a subsequence and replacing y by $-y$ if necessary. We shall prove that (2.5.3) implies the existence of $\hat{F} \in V^*_{x,P} \setminus \{F\}$. Indeed, take $F_n \in V^*_{x+t_n y,P}$. Then, by (2.5.1), we have

$$P(x) - P(x + t_n y) \geq F_n(-t_n y) \tag{2.5.4}$$

The inequalities (2.5.3) and (2.5.4) give

$$t_n F(y) + \varepsilon t_n \leq t_n F_n(y).$$

Hence

$$(F_n - F)(y) \geq \varepsilon. \tag{2.5.5}$$

In the case when P is Lipschitz continuous, and this is the case of topological pressure see (Theorem 2.5.1) which we are mostly interested in, all F_n's, $n \geq 1$, are uniformly bounded. Indeed, let L be a Lipschitz constant of P. Then for every $z \in V$ and every $n \geq 1$

$$F_n(z) \leq P(x + t_n y + z) - P(x + t_n y) \leq L||z||$$

So, $||F_n|| \leq L$ for every $n \geq 1$. Thus, there exists $\hat{F} = \lim_{n\to\infty} F_n$, a weak*-limit of a sequence $\{F_n\}_{n\geq 1}$ (subsequence of the previous sequence). By (2.5.5) $(\hat{F} - F)(y) \geq \varepsilon$. Hence $\hat{F} \neq F$. Since

$$P(x + t_n y + v) - P(x + t_n y) \geq F_n(v) \quad \text{for all } n \text{ and } v \in V$$

passing with n to ∞ and using continuity of P, we conclude that $\hat{F} \in V^*_{x,P}$.

If we do not assume that P is Lipschitz continuous, we restrict F_n to the 1-dimensional space spanned by y i.e. we consider $F_n|_{I\!Ry}$. In view of (2.5.5) for every $n \geq 1$ there exists $0 \leq s_n \leq 1$ such that $F_n(s_n y) - F(s_n y) = \varepsilon$. Passing to a subsequence, we may assume that $\lim_{n\to\infty} s_n = s$ for some $s \in [0, 1]$. Define

$$f_n = s_n F_n|_{I\!Ry} + (1 - s_n) F|_{I\!Ry}$$

Then $f_n(y) = F(y) = \varepsilon$ hence $||f_n|| = ||F|| = \frac{\varepsilon}{||y||}$ for every $n \geq 1$. Thus the sequence $\{f_n\}_{n\geq 1}$ is uniformly bounded and, consequently, it has a weak-* limit $\hat{f} : I\!Ry \to I\!R$. Now we use Theorem 2.5.4 (Hahn-Banach) for the affine set M being the graph of $\hat{f}$ translated by $(x, P(x))$ in $V \times I\!R$. for every $\alpha \in I\!R$ and every $n \geq 1$. We extend M to H and find the linear functional $\hat{F} \in V^*_{x,P}$ whose graph is H. Since $\hat{F}(y) - F(y) = \hat{f}(y) - F(y) = \varepsilon$, $\hat{F} \neq F$.

Suppose now that Proposition 2.5.4, $V^*_{x,P}$ contains at least two distinct linear functionals, say F and $\hat{F}$. So, $F(y) - \hat{F}(y) > 0$ for some $y \in V$. Suppose on the contrary that P is differentiable in every direction at the point x. In particular P is differentiable in the direction y. Hence

$$\lim_{t\to 0} \frac{P(x+ty) - P(x)}{t} = \lim_{t\to 0} \frac{P(x-ty) - P(x)}{-t}$$

and consequently

$$\lim_{t\to 0} \frac{P(x+ty) + P(x-ty) - 2P(x)}{t} = 0.$$

On the other hand, for every $t > 0$, we have $P(x+ty) - P(x) \geq F(t) = tF(y)$ and $P(x-ty) - P(x) \geq \hat{F}(-ty) = -t\hat{F}(y)$, hence

$$\liminf_{t\to 0} \frac{P(x+ty) + P(x-ty) - 2P(x)}{t} \geq F(y) - \hat{F}(y) > 0.$$

, a contradiction.

In fact $F(y) - \hat{F}(y) = \varepsilon > 0$ implies $F(y') - \hat{F}(y') \geq \varepsilon/2 > 0$ for all y' in the neighbourhood of y in the weak topology defined just by $\{y' : (F - \hat{F})(y - y') < \varepsilon/2\}$. Hence P is not differentiable in a weak*-open set of directions. ♣

Let us go back now to our special situation:

Proposition 2.5.6. If $\mu \in M(T)$ is an equilibrium state for $\phi \in C(X)$, then the linear functional represented by μ is tangent to P at ϕ.

Proof. We have

$$\mu(\phi) + h_\mu = P(\phi)$$

and for every $\psi \in C(X)$

$$\mu(\phi + \psi) + h_\mu \leq P(\phi + \psi).$$

Subtracting the sides of the equality from the respective sides of the latter inequality we obtain $\mu(\psi) \leq P(\phi + \psi) - P(\phi)$ which is just the inequality defining tangent functionals. ♣

As an immediate consequence of Proposition 2.5.6. and Theorem 2.5.5 we get the following.

Corollary 2.5.7. If the pressure function P is differentiable at ϕ in every direction, or at least in a dense in the weak topology set of directions, then there is at most one equilibrium state for ϕ.

Due to this Corollary, in future (see Chapter 4) to prove uniqueness it will be sufficient to prove differentiability of the pressure function in a weak*-dense set of directions.

The next part of this section will be devoted to kind of reversing Proposition 2.5.6 and Corollary 2.5.7. and better understanding of the mutual Legendre-Fenchel transforms $-h$ and P. This is a beautiful topic but will not have applications in the rest of this book. Let us start with a characterization of T invariant measures in the space of all signed measures $C(X)^*$ provided by the pressure function P.

Theorem 2.5.8. For every $F \in C(X)^*$ the following three conditions are equivalent:

(i) For every $\phi \in C(X)$ it holds $F(\phi) \le \mathrm{P}(\phi)$.

(ii) There exists $C \in I\!R$ such that for every $\phi \in C(X)$ it holds $F(\phi) \le \mathrm{P}(\phi) + C$.

(iii) F is represented by a probability invariant measure $\mu \in M(X,T)$.

Proof. (iii) $\Rightarrow$ (i) follows immediately from the variational principle:

$$F(\phi) \le F(\phi) + h_\mu(T) \le \mathrm{P}(\phi) \quad \text{for every} \quad \phi \in C(X).$$

(i) $\Rightarrow$ (ii) is obvious. Let us prove that (ii) $\Rightarrow$ (iii). Take an arbitrary positive $\phi \in C(X)$, i.e. such that for every $x \in X$, $\phi(x) \ge 0$. For every real $t < 0$ we have

$$F(t\phi) \le \mathrm{P}(t\phi) + C$$

Since $t\phi \le 0$ it immediately follows from (2.5.1) that $\mathrm{P}(t\phi) \le \mathrm{P}(0)$. Hence $F(t\phi) \le P(0) + C$. So

$$|t|F(\phi) \ge -(C + \mathrm{P}(0)) \quad \text{hence} \quad F(\phi) \ge \frac{-(C + \mathrm{P}(0))}{|t|}.$$

Letting $t \to -\infty$ we obtain $F(\phi) \ge 0$. We estimate the value of F on constant functions t. For every $t > 0$ we have $F(t) \le \mathrm{P}(t) + C \le \mathrm{P}(0) + t + C$. Hence $F(1) \le 1 + \frac{\mathrm{P}(0)+C}{t}$. Similarly $F(-t) \le \mathrm{P}(-t) + C = \mathrm{P}(0) - t + C$ and therefore $F(1) \ge 1 - \frac{\mathrm{P}(0)+C}{t}$. Letting $t \to \infty$ we thus obtain $F(1) = 1$. Therefore by Theorem 2.1.1 (Riesz Representation Theorem) the functional F is represented by a probability measure $\mu \in M(X)$. Let us finally prove that μ is T-invariant. For every $\phi \in C(X)$ and every $t \in I\!R$ we have by (i)

$$F(t(\phi \circ T - \phi)) \le \mathrm{P}(t(\phi \circ T - \phi)) + C$$

It immediately follows from Theorem 2.3.1 (Variational principle) that $\mathrm{P}(t(\phi \circ T - \phi)) = \mathrm{P}(0)$. Hence

$$|F(\phi \circ T) - F(\phi)| \le \left| \frac{\mathrm{P}(0) + C}{t} \right|.$$

Thus, letting $|t| \to \infty$, we obtain $F(\phi \circ T) = F(\phi)$, i.e T-invariance of μ. ♣

We shall prove the following.

Corollary 2.5.9. Every functional F tangent to P at $\phi \in C(X)$, i.e. $F \in C(X)^*_{\phi,P}$, is represented by a probability T-invariant measure $\mu \in M(X,T)$.

Proof. Using Theorem 2.5.1, we get for every $\psi \in C(X)$ that

$$F(\psi) \le P(\phi+\psi) - P(\phi) \le P(\psi) + |P(\phi+\psi) - P(\psi)| - P(\phi) \le P(\psi) + ||\phi||_\infty - P(\phi).$$

So condition (ii) of Theorem 2.5.8 holds hence (iii) holds, F is represented by $\mu \in M(X,T)$. ♣

We can now almost reverse Proposition 2.5.6. Namely being a functional tangent to P at ϕ implies being an "almost" equilibrium state for ϕ.

Theorem 2.5.10. $F \in C(X)^*_{\phi,P}$ if and only if F, in other words the measure $\mu = \mu_F \in M(X,T)$ representing F, is a weak*-limit of measures $\mu_n \in M(X,T)$ such that

$$\mu_n\phi + \mathrm{h}_{\mu_n}(T) \to \mathrm{P}(\phi).$$

Proof. In one way the proof is simple. Assume that $\mu = \lim_{n\to\infty} \mu_n$ in the weak* topology and $\mu_n\phi + \mathrm{h}_{\mu_n}(T) \to P(\phi)$. We proceed as in Proof of Theorem 2.5.6. In view of Theorem 2.3.1 (Variational principle) $\mu_n(\psi+\phi) + \mathrm{h}_{\mu_n}(T) \le \mathrm{P}(\phi+\psi)$ which means that $\mu_n(\psi) \le \mathrm{P}(\phi+\psi) - (\mu_n\phi + \mathrm{h}_{\mu_n}(T))$. Thus, letting $n\to\infty$, we get $\mu(\psi) \le \mathrm{P}(\phi+\psi) - \mathrm{P}(\psi)$. This means that $\mu \in C(X)^*_{\phi,\mathrm{P}}$.

Now, let us prove our Theorem in the other direction. Recall again that the function $\mu \mapsto \mathrm{h}_\mu(T)$ on $M(T)$ is affine, hence concave. Denote $\overline{\mathrm{h}}_\mu = \limsup_{\nu\to\mu} \mathrm{h}_\nu(T)$, with $\nu \to \mu$ in weak*-topology. It is also concave and upper semicontinuous on $M(T) := M(X,T)$. In the sequel we shall prefer to consider the function $\mu \mapsto -\overline{\mathrm{h}}_\mu(T)$ which is lower semicontinuous and convex.

Step 1. For every $\vartheta \in C(X)$ we have

$$\mu\vartheta - \sup_{\nu\in M(T)} (\nu\vartheta - -\mathrm{h}_\nu(T)) \le \mu\vartheta - (\mu\vartheta - -\overline{\mathrm{h}}_\mu(T)) = -\overline{\mathrm{h}}_\mu(T)).$$

We obtained here $-\overline{\mathrm{h}}_\nu(T))$ rather than merely $-\mathrm{h}_\nu(T))$ by taking every sequence $\mu_n \to \mu$ writing the right hand side of the above inequality : $\mu\vartheta - (\mu_n\vartheta - -\mathrm{h}_{\mu_n}\vartheta(T))$ and letting $n \to \infty$. So

$$\sup_{\vartheta\in C(X)} \Big(\mu\vartheta - \sup_{\nu\in M(T)} (\nu\vartheta - -\mathrm{h}_\nu(T))\Big) \le -\overline{\mathrm{h}}_\mu(T).$$

This says that the LF-transform of the LF-transform of $-\mathrm{h}_\mu(T)$ is less or equal to $-\overline{\mathrm{h}}_\mu(T)$.

Let us prove now the opposite inequality. We refer to the following corollary of the geometric form of Hahn-Banach Theorem [Bourbaki, Ch.II.§5. Prop.5]:

Let M be a closed convex set in a locally convex vector space V. Then every lower semicontinuous convex function f defined in M is supremum of a family of functions bounded from above by f, which are restrictions to M of continuous affine functions in V.

We shall apply this theorem to $C^*(X)$ endowed with weak*-topology and use the fact that every linear functional continuous in this topology is represented by an element belonging to $C(X)$. (This is a general fact concerning pairs of vector spaces in duality, [Bourbaki, Ch.II.§6. Prop.3.].). Thus, for every $\varepsilon > 0$ there exists $\psi \in C(X)$ such that for every $\nu \in M(T)$

$$(\nu - \mu)(\psi) \leq -\mathrm{h}_\nu(T) - -\overline{\mathrm{h}}_\mu + \varepsilon \tag{2.5.6}$$

So

$$\mu\psi - \sup_{\nu\in M(T)} (\nu\psi - -\mathrm{h}_\nu(T)) \geq -\overline{\mathrm{h}}_\mu(T) - \varepsilon.$$

Letting $\varepsilon \to 0$ we obtain

$$\sup_{\vartheta\in C(X)} \Big(\mu\vartheta - \sup_{\nu\in M(T)} (\nu\vartheta - -\mathrm{h}_\nu(T))\Big) \geq -\overline{\mathrm{h}}_\mu(T).$$

Thus we proved the standard fact that the LF-transform of the LF-transform of $-h_\mu(T)$ is back $-\overline{\mathrm{h}}_\mu(T)$. Remind now that by variational principle the LF-transform of $-\mathrm{h}_\nu(T)$, i.e. the supremum $\sup_{\nu\in M(T)}(\nu\vartheta - -\mathrm{h}_\nu(T))$ is pressure $\mathrm{P}(\vartheta)$. We conclude that

$$\overline{\mathrm{h}}_\mu(T) = \inf_{\vartheta\in C(X)} \{\mathrm{P}(\vartheta) - \mu\vartheta\} \tag{2.5.7}$$

Step 2. Fix $\mu \in C(X)^*_{\phi,\mathrm{P}}$. From $\mu\psi \leq P(\phi+\psi) - P(\phi)$ we obtain

$$P(\phi+\psi) - \mu(\phi+\psi) \geq \mathrm{P}(\phi) - \mu\phi \quad \text{for all} \quad \psi \in C(X)$$

or

$$\inf_{\psi\in C(X)} \{\mathrm{P}(\psi) - \mu\psi\} \geq \mathrm{P}(\phi) - \mu\phi \tag{2.5.8}$$

(This expresses the fact that the supremum ($-$ infimum above) in the definition of the LF-transform of P at F is attained at ϕ at which F is tangent to P.) By (2.5.7) and (2.5.8) we obtain

$$\overline{\mathrm{h}}_\mu \geq \mathrm{P}(\phi) - \mu\phi \tag{2.5.9}$$

so by the definition of $\overline{\mathrm{h}}_\mu$ there exists a sequence of measures $\mu_n \in M(T)$ such that $\lim_{n\to\infty} \mu_n = \mu$ and $\lim_{n\to\infty} h_{\mu_n} \geq P(\phi) - \mu\phi$. The proof is finished. ♣

Remark. In Step 1 of the above proof it did not matter whether we considered μ tangent to P or an arbitrary $\mu \in M(T)$. In Step 2, where we started with all $\mu \in M(T)$, considering $\varepsilon > 0$ in (2.5.6) is necessary; without $\varepsilon > 0$ this formula may happen to be false, see Example 2.5.13. For $\mu \in C(X)^*_{\phi,\mathrm{P}}$ we obtain from (2.5.9) and inequality $\mathrm{h}_\nu(T) \leq \mathrm{P}(\phi) - \nu\phi$ for every $\nu \in M(T)$ that $h_\nu(T) - \overline{\mathrm{h}}_\mu(T) \leq (\mu-\nu)\phi$ which is just (2.5.6) with $\varepsilon = 0$. So *a posteriori* we know that ε in (2.5.6) can be omitted.

The meaning of this, is that if μ is tangent to P at ϕ then ϕ is tangent to $-\overline{\mathrm{h}}$, the LF-transform of P, at μ.

Conversely, if ψ satisfies (2.5.6) with $\varepsilon = 0$ i.e. ψ is tangent to $-\overline{\mathrm{h}}$ at $\mu \in M(T)$ then in the same way as in Step 2. we can prove the inequality analogous to (2.5.8), namely that

$$\sup_{\nu \in M(T)} \nu\psi - -\overline{\mathrm{h}}_\mu(T) = \mathrm{P}(\psi) \leq \mu\psi - -\overline{\mathrm{h}}_\mu(T).$$

Hence μ is tangent to P at ψ (by the "if" part of Theorem 2.5.10).

Assume now the upper semicontinuity of the entropy $\mathrm{h}_\mu(T)$ as a function of μ. Then we obtain.

Corollary 2.5.11. If the entropy is upper semicontinuous, then a functional $F \in C(X)^*$ is tangent to P at $\phi \in C(X)$ if and only if it is represented by a measure which is an equilibrium state for ϕ.

Proof. This is just the previous Theorem with the observation that $\lim_{n\to\infty} \mathrm{h}_{\mu_n}(T) \leq \mathrm{h}_{\mu_T}(T)$. (Remark that this uses only the upper semicontinuity of the entropy at the measure μ.) ♣

Recall that already the upper semicontinuity above implies the existence of at least one equilibrium state (Lemma 2.4.3)

Now we can complete Corollary 2.5.7.

Corollary 2.5.12. If the entropy is upper semicontinuous then the pressure function P is differentiable at $\phi \in C(X)$ in every direction, or in a set of directions dense in the weak topology, if and only if there is at most one equilibrium state for ϕ.

Proof. This Corollary follows directly from Corollary 2.5.11 and Theorem 2.5.5. ♣

After discussing functionals tangent to P and proving that they coincide with the set of equilibrium states for maps for which the entropy is upper semicontinuous as the function on $M(T)$ the question arises of whether all measures in $M(T)$ are equilibrium states of some continuous functions. The answer given below is no.

Example 2.5.13. We shall construct a measure $m \in M(T)$ which is not an equilibrium state for any $\phi \in C(X)$. Here X is the one sided shift space Σ^2 with the left side shift map σ. Since this map is obviously expansive, it follows from Theorem 2.4.6 that the entropy function is upper semicontinues. Let $m_n \in M(\sigma)$ be the measure equidistributed on the set Per_n of points of period n, i.e.

$$m_n = \sum_{x \in \mathrm{Per}_n} \frac{1}{\mathrm{Card\ Per}_n} \delta_x$$

where δ_x is the Dirac measure supported by x. m_n converge weak* to $\mu_{\max}$, the measure of maximal entropy: $\log 2$. (Check that this follows for example from the proof of variational principle, part II.) Let $t_n, n = 0, 1, 2, ...$ be a sequence of positive real numbers such that $\sum_{n=0}^{\infty} t_n = 1$. Finally define

$$m = \sum_{n=0}^{\infty} t_n m_n$$

Let us prove that there is no $\phi \in C(X)$ tangent to h at m. Let $\mu_n = R_n \mu_{\max} + \sum_{j=0}^{n-1} t_j m_j$, where $R_n = \sum_{j=n}^{\infty} t_j$. We have of course $\mathrm{h}_{m_n}(\sigma) = 0,\ n = 1, 2, ...$ Also $\mathrm{h}_m(\sigma) = 0$. This follows from the fact that h is affine on $M(\sigma)$, the function h is bounded by the topological entropy $\mathrm{h}_{\mathrm{top}}(\sigma) = \log 2$ and

$$m = \sum_{j=0}^{n-1} t_j m_j + R_n \sum_{j=n}^{\infty} \frac{t_j}{R_n} m_j .$$

Thus

$$\mathrm{h}_{\mu_n}(\sigma) - \mathrm{h}_m(\sigma) = R_n \mathrm{h}_{\mu_{\max}}(\sigma) = R_n \log 2$$

and for an arbitrary $\phi \in C(\Sigma^2)$

$$(\mu_n - m)\phi = (R_n \mu_{\max} - \sum_{j=n}^{\infty} t_j m_j)\phi \le R_n \varepsilon$$

where $\varepsilon \to 0$ as $n \to \infty$ because $m_j \to \mu_{\max}$. The inequality $\mathrm{h}_{\mu_n}(\sigma) - \mathrm{h}_m(\sigma) \ge R_n \log 2$ and the latter inequality prove that ϕ is not tangent to h at m. Indeed $\mathrm{h}_{\mu_n}(\sigma) - \mathrm{h}_m(\sigma) > (\mu_n - m)\phi$ for n large, contrary to (2.5.5) with $\varepsilon = 0$.

By Remark after Theorem 2.5.10 we know that m is not tangent to any ϕ for the pressure function P. In fact it is easy to see it directly: For an arbitrary $\phi \in C(\Sigma^2)$ we have $\mu_{\max}\phi < \mathrm{P}(\phi)$ because $\mathrm{h}_{\mu_{\max}}(\sigma) > 0$, so $m_n\phi < P(\phi)$ for all n large enough as $m_n \to \mu_{\max}$. Also $m_n\phi \le P(\phi)$ for all n's. So for the average of m_n's namely m we have $m\phi < \mathrm{P}(\phi)$. So ϕ is not an equilibrium state. ♣

The measure m in this example is very non-ergodic, this is necessary as will follow from Exercise 5.

EXERCISES

Exercise 1. Let $T : X \to X$ and $S : Y \to Y$ be two continuous maps of compact metric spaces respectively. Show that $\mathrm{h}_{\mathrm{top}}(T \times S) = \mathrm{h}_{\mathrm{top}}(T) + \mathrm{h}_{\mathrm{top}}(S)$.

Exercise 2. Prove that $T : X \to X$ is an isometry of a compact metric space X, then $\mathrm{h}_{\mathrm{top}}(T) = 0$

Exercise 3. Show that if $T : X \to X$ is a local homeomorphism of a compact metric space, then the number $d = \#T^{-1}(x)$ is finite and independent of $x \in X$.

Exercise 4. With the assumtions and notation of Exercise 3, demonstrate that $\mathrm{h_{top}}(T) \geq \log d$

Exercise 5. Prove that if $f : M \to M$ is a C^1 endomorphism of a compact differentiable manifold M, then $\mathrm{h_{top}}(T) \geq \log \deg(f)$. (Hint: see [MP]).

Exercise 6. Let $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle and let $f_n : S^1 \to S^1$ be the map defined by the formula $f_n(z) = z^n$. Show that $\mathrm{h_{top}}(f_n) = \log n$.

Exercise 7. Let $\sigma_A : \Sigma_A \to \Sigma_A$ be the shift map generated by the incidence matrix A. Prove that $\mathrm{h_{top}}(\sigma_A)$ is equal to the logarithm of the spectral radius of A.

Exercise 8. Show that for every continuous potential ϕ, $\mathrm{P}(\phi) \leq \mathrm{h_{top}}(T) + \sup(\phi)$.

Exercise 9. Provide an example of a transitive diffeomorphism without measures of maximal entropy.

Exercise 10 Provide an example of a transitive diffeomorphism with at least measures of maximal entropy.

Exercise 11. Find a sequnce of continuous maps $T_n : X_n \to X_n$ such that $\mathrm{h_{top}}(T_{n+1}) > \mathrm{h_{top}}(T_n)$ and $\lim_{n\to\infty} \mathrm{h_{top}}(T_n) < \infty$.

Exercise 12. Prove that for an arbitrary convex continuous function $P : V \to \mathbb{R}$ on a real Banach space V the set of tangent functionals: $\bigcup_{x\in V} V^*_{x,P}$ is dense in the norm topology in

$$\{F \in V^* :\ \text{there exists } C \in \mathbb{R} \text{ such that for every } x \in V,\ F(x) \leq P(x) + C\}$$

(such functionals are called P-bounded)

Remark. The conclusion is that for P the pressure function on $C(X)$ tangent measures are dense in $M(X,T)$, see Theorem 2.4.6. Hint: This follows from Bishop – Phelps Theorem, see [BP] or Israel's book [I, pp.112-115], which can be stated as follows: For every P-bounded F_0 $x_0 \in V$ and $\varepsilon > 0$ there exists $x \in V$ and $F \in V^*$ tangent to P at X such that

$$\|F - F_0\| \leq \varepsilon \text{ and } \|x - x_0\| \leq \frac{1}{\varepsilon}\Big(P(x_0) - F_0(x_0) + s(x_0)$$

where $s(F_0) := \sup_{x'\in V} F_0 x' - P(x')$ (the LF-transform of P. The idea of the proof of this theorem is as follows: If we replace P by $Q(x0 := P(x) - F_0(x) + s(x_0)$ the theorem reduces to the case $F_0 \equiv 0, s(F_0) = 0$. For each $x \in V$ consider the cone

$$\mathcal{C}(x) = \{(x', y) : y - Q(x) < -\varepsilon\|x' - x\|\}.$$

There is $x \in V$ such that $\mathcal{C}(x) \cap \mathrm{graph} Q = \{x\}$ Now F can be defined as a functional which graph translated by a constant separates $\mathcal{C}(x)$ from $\{y \geq Q(x)$.

Exercise 13. Prove that in the situation from Exercise 1 for every $x \in V$ $V^*_{x,P}$ is convex and weak*-compact.

Exercise 14. Let E_ϕ denote the set of all equilibrium states for $\phi \in C(X)$.

(i) Prove that E_ϕ is convex.

(ii) Find an example that E_ϕ is not weak*-compact.

(iii) Prove that extremal points of E_ϕ are extremal points of $M(X,T)$.

(iv) Prove that almost all measures in the ergodic decomposition of an arbitrary $\mu \in E_\phi$ belong also to E_ϕ. (One says that every equilibrium state has a unique decomposition into pure, i.e. ergodic, equilibrium states .)

Hints: In (ii) consider a sequence of Smale horseshoes of topological entropies $\log 2$ converging to a point fixed for T. To prove (iii) and (iv) use the fact that entropy is an affine function of measure.

Exercise 15. Find an example showing that the point (iii) of Exercise 3 is false if we consider $C(X)^*_{\phi,P}$ rather than E_ϕ.

Hint: An idea is to have two fixed points p, q and two trajectories $(x_n), (y_n)$ such that $x_n \to p, y_n \to q$ for $n \to \infty$ and $x_n \to q, y_n \to p$ for $n \to -\infty$. Now take a sequence of periodic orbits γ_k approaching $\{p,q\} \cup \{x_n\} \cup \{y_n\}$ with periods tending to ∞. Take their Cartesian products with corresponding invariant subsets A_k's of small horseshoes of topological entropies less than $\log 2$ but tending to $\log 2$, diameters of the horseshoes shrinking to 0 as $k \to \infty$. Then for $\phi \equiv 0$ $C(X)^*_{\phi,P}$ consists only of measure $\frac{1}{2}(\delta_p + \delta_q)$. One cannot repeat the proof in Exercise 3(iii) with the function $\overline{\mathrm{h}}_\mu$ instead of the entropy function h_μ, because $\overline{\mathrm{h}}_\mu$ is no more affine !

This is Peter Walters' example, for details see the preprint [W2].

Exercise 16. Suppose that the entropy function h_μ is upper semicontinuous (then for each $\phi \in C(x)$ $C(X)^*_{\phi,\mathrm{P}} = E_\phi$, see Corollary 2.5.11). Prove that

(i) every $\mu \in M(T)$ which is a finite combination of ergodic masures $\mu = \sum t_j m_j$, $m_j \in M(T)$, is tangent to P more precisely there exists $\phi \in C(X)$ such that $\mu, m_j \in C(X)^*_{\phi,\mathrm{P}}$ and moreover they are equilibrium states for ϕ.

(ii) if $\mu = \int_{M_e(T)} m d\alpha(m)$ where $M_e(X,T)$ consists of ergodic measures in $M(X,T)$ and α is a probability non-atomic measure on $M_e(X,T)$, then there exists $\phi \in C(X)$ which has uncountably many ergodic equilibria in the support of α.

(iii) the set of elements of $C(X)$ with uncountably many ergodic equilibria is dense in $C(X)$.

Hint: By Bishop – Phelps Theorem (Exercise 12) there exists $\nu \in E_\phi$ arbitrarily close to μ. Then in its ergodic decomposition there are all the measures μ_j because all ergodic measures are far apart from each other (in the norm in $C(X)^*$). These measures by Exercise 14 belong to the same E_ϕ what proves (i). For more details and proofs of (ii) and (iii) see [Israel, Theorem V.2.2] or [Ruelle, 1978, 3.17, 6.15].

Remark. In statistical physics the occurence of more then one equilibrium for $\phi \in C(X)$ is called "phase transition". (iii) says that the set of functions with "very rich" phase transition is dense. For the further discussion see also [Israel, V.2].

Exercise 17. Prove the following. Let $P : V \to \mathbb{R}$ be a continuous convex function on a real Banach space V with norm $\|\cdot\|_V$. Suppose P is differentiable at $x \in V$ in every direction. Let $W \subset V$ be an arbitrary linear subspace with norm $\|\cdot\|_W$ such that the embedding $W \subset V$ is continuous and the unit ball in $(W, \|\cdot\|_W)$ is compact in $(V, \|\cdot\|_V)$. Then $P|_W$ is differentiable in the sense that there exists a functional $F \in V^*$ such that for

$y \in W$ it holds

$$|P(x+y) - P(x) - F(y)| = o(\|y\|_W).$$

Remark. In Chapter 3 we shall discuss W being the space of Hölder continuous functions with an arbitrary exponent $\alpha < 1$ and the entropy function will be upper semicontinuous. So the conclusion will be that uniqueness of the equilibrium state at an arbitrary $\phi \in C(X)$ is equivalent to the differentiability in the direction of this spaece of Hö .lder functions.

Exercise 18. (Walters) Prove that the pressure function P is Frechet differentiable at $\phi \in C(X)$ if and only if P affine in a neighbourhood of ϕ. Prove also the conclusion: P is Frechet differentiable at every $\phi \in C(X)$ if and only if T is uniquely ergodic, namely if $M(X,T)$ consists of one element.

Exercise 19. Prove S. Mazur's Theorem: If $P : V \to I\!R$ is a continuous convex function on a real separable Banach space V then the set of points at which there exists a unique functional tangent to P is dense G_δ.

Remark. In the case of the pressure function on $C(X)$ this says that for a dense G_δ set of functions there exists at most one equilibrium state. Mazur's Theorem contrasts with the theorem from Exercise 16 (iii).

BIBLIOGRPHICAL NOTES

The concept of topological pressure in dynamical context was introduced by D. Ruelle in [Ruelle, 1973] and since then have been studied in many papers and books. Let us mention only [Bowen, 1975], [Wallters, 1976], [Wallters, 1982] and [Ruelle, 1978]. The topological entropy was introduced earlier in [AKM, 1965]. The variational principle (Theorem 2.3.1) has been proved for some maps in [Ruelle, 1973]. The first proofs of this principle in its full generality can be found in [Walters, 1976] and [Bowen, 1975]. The simplest proof presented in this chapter is taken from [Mi, 1976]. In the case of topological entropy (potential $\phi = 0$) the corresponding results have been obtained earlier: Goodwyn in [Goodwyn, 1969] proved the first part of the variational principle, Dinaburg in [Dinaburg, 1971] proved its full version assuming that the space X has finite covering topological dimension and finally Goodman proved in [Goodman, 1971] the variational principle for topological entropy without any additional assumptions. The concept of equilibrium states and expansive maps in mathematical setting was introduced in [Ruelle, 1973] where the first existence and uniqueness type results have appeared. Since then these concepts have been explored by many authors, in particular in [Bowen, 1975] and [Ruelle, 1978]. The material of Section 2.5 is mostly taken from [Ruelle, 1978], [Israel, 1979] and [Ellis, 1985].

References

N.Bourbaki: Espaces Vectoriels Topologiques. Masson,Paris, 1981.

M.Denker, Ch.Grillenberger, K.Sigmund: Ergodic Theory on Compact Spaces. Lecture Notes in Mathematics, 527, Springer-Verlag, Berlin 1976.

CHAPTER 3
DISTANCE EXPANDING MAPS

version Feb.1, 2000

We devote this Chapter to a closer topological study of distance expanding maps. Often however weaker assumptions will be sufficient. We always assume the maps are continuous on a compact metric space X and usually assume the maps are open, which means that open sets have open images. This is equivalent to saying that if $f(x) = y$ and $y_n \to y$ then there exist $x_n \to x$ such that $f(x_n) = y_n$ for n large enough.

In theorems with assertions of topological character the assumption that a map is only expansive gives in fact always the same as if we assumed that the map is expanding, in view of Sec.6. We shall prove in Sec.6 that for every expansive map there always exists a metric compatible with the topology on X given by an original metric, so that the map is distance expanding in it.

Recall that for (X, ρ) a compact metric space, a continuous mapping $T : X \to X$ is said to be *distance expanding* (with respect to the metric ρ) if there exist constants $\lambda > 1$ and $\eta > 0$ such that

$$\rho(x,y) \le 2\eta \implies \rho(T^n(x), T^n(y)) \ge \lambda\rho(x,y) \tag{3.1.1}$$

We say that T is *distance expanding at a set* $Y \subset X$ if the above holds for every $x, y \in B(z, \eta)$ for $z \in Y$.

In future we shall usually be able to assume that $n = 1$ i.e. that

$$\rho(x,y) \le 2\eta \implies \rho(T(x), T(y)) \ge \lambda\rho(x,y) \tag{3.1.2}$$

One can achieve this in two ways:

(1) If T is Lipschitz continuous (say with constant $L > 1$) replace the metric $\rho(x, y)$ by $\sum_{j=0}^{n-1} \rho(T^j(x), T^j(y))$. Of course then λ and η change. As an exercise you can check that the number $1 + (\lambda - 1)\left(\frac{L-1}{L^n-1}\right)$ can play the role of λ in (3.1.2).

(2) Consider T^n instead of T.

Sometimes we shall write for short *expanding*, instead of *distance expanding.*

§3.1 DISTANCE EXPANDING OPEN MAPS, BASIC PROPERTIES

Let us start with a lemma where we assume $T : X \to X$ is a continuous open map of a compact metric space X. We do not need to assume in this lemma that T is distance expanding.

Lemma 3.1.2. If $T : X \to X$ is a continuous open map, then for every $\eta > 0$ there exists $\xi > 0$ such that $T(B(x, \eta)) \supset B(T(x), \xi)$ for every $x \in X$.

Proof. For every $x \in X$ let

$$\xi(x) = \sup\{r > 0 : T(B(x, \eta)) \supset B(T(x), r)\}.$$

Since T is open, $\xi(x) > 0$. Since $T(B(x, \eta)) \supset B(T(x), \xi(x))$, it suffices to show that $\xi = \inf\{\xi(x) : x \in X\} > 0$. Suppose conversely that $\xi = 0$. Then there exists a sequence of points $x_n \in X$ such that

$$\xi(x_n) \to 0 \quad \text{as} \quad n \to \infty \tag{3.1.3}$$

and, as X is compact, we can assume that $x_n \to y$ for some $y \in X$. Hence $B(x_n, \eta) \supset B(y, \frac{1}{2}\eta)$ for all n large enough. Therefore

$$T(B(x_n, \eta)) \supset T\left(B(y, \frac{1}{2}\eta)\right) \supset B(T(y), \varepsilon) \supset B\left(T(x_n), \frac{1}{2}\varepsilon\right)$$

for some $\varepsilon > 0$ and again for every n large enough. The existence of ε such that the second inclusion holds follows from the openness of T. Consequently $\xi(x_n) \geq \frac{1}{2}\varepsilon$ for these n, which contradicts (3.1.3). ♣

If $T : X \to X$ is an open, expanding map, then by (3.1.1), for all $x \in X$, the restriction $T|_{B(x,\eta)}$ is injective and therefore it has an inverse map. The same holds for expanding at Y for all $x \in Y$. In view of Lemma 3.1.2 we can introduce the following definition.

Notation 3.1.3. If $T : X \to X$ is expanding then for all $x \in X$ the inverse of the map $T|_{B(x,\eta)}$ restricted to the ball $B(T(x), \xi)$ will be denoted by T_x^{-1}.

Observe that for every $y \in X$

$$T^{-1}(B(y, \xi)) = \bigcup_{x \in T^{-1}(y)} T_x^{-1}(B(y, \xi)) \tag{3.1.4}$$

Indeed, suppose that $y' = T(x') \in B(y, \xi)$. Then $y \in B(y', \xi)$. Let $x = T_{x'}^{-1}(y)$. As T_x^{-1} and $T_{x'}^{-1}$ coincide on y, they coincide on y' because they map y' into $B(x, \eta)$ and T is injective on $B(x, \eta)$. Thus $x' = T_x^{-1}(y')$.

A map T with the property that there exists ξ such that for each $B(x, \xi)$ (3.1.4) holds with the sets in the union disjoint from each other and T restricted to each of them being a homeomorphism, is called a *covering map*. So we proved that a continuous open locally injective map of a compact metric space is a covering map. This is well known but we gave the proof for the completeness of the exposition.

Immediately from Definition 3.1.3 we have

$$T_x^{-1}(B(T(x), \xi)) \subset B(x, \eta) \tag{3.1.5}$$

¿From now on throughout this section we assume also the expanding property, i.e. (3.1.2). We then get the following.

Lemma 3.1.4. If $x \in X$ and $y, z \in B(T(x), \xi)$ then

$$\rho(T_x^{-1}(y), T_x^{-1}(z)) \leq \lambda^{-1}\rho(y, z)$$

In particular $T_x^{-1}(B(T(x), \xi)) \subset B(x, \lambda^{-1}\xi) \subset B(x, \xi)$.

Definition 3.1.5. For every $x \in X$, every $n \geq 1$ and every $j = 0, 1, \ldots, n-1$ write $x_j = T^j(x)$. In view of Lemma 3.1.4 the composition $T_{x_0}^{-1} \circ T_{x_1}^{-1} \circ \ldots \circ T_{x_{n-1}}^{-1} : B(T^n(x), \xi) \to X$ is well-defined and will be denoted by T_x^{-n}.

Below we collect the basic elementary properties of maps T_x^{-n} following immediately from the above. For every $y \in X$

$$T^{-n}(B(y, \xi)) = \bigcup_{x \in T^{-n}(y)} T_x^{-n}(B(y, \xi)); \tag{3.1.6}$$

$$\rho(T_x^{-n}(y), T_x^{-n}(z)) \leq \lambda^{-n}\rho(y, z) \text{ for all } y, z \in B(T^n(x), \xi); \tag{3.1.7}$$

$$T_x^{-n}(B(T^n(x), r)) \subset B(x, \min\{\eta, \lambda^{-n}r\}) \text{ for every } r \leq \xi. \tag{3.1.8}$$

Remark. All these properties hold, and notation makes sense, also for open maps $T : X \to X$ expanding at $Y \subset X$, provided $x, T(x), ..., T^n(x) \in Y$.

§3.2 SHADOWING OF PSEUDOORBITS

We keep the notation of Section 3.1. We consider an open distance expanding map $T : X \to X$ with the constants η, λ, ξ.

Let n be a non-negative integer or ∞. Given $\alpha \geq 0$ a sequence $(x_i : i = 0, \dots, n)$ is said to be an *α-pseudo-orbit* for $T : X \to X$ if and only if for every $i = 0, \dots, n-1$

$$\rho(T(x_i), x_{i+1}) \leq \alpha \tag{3.2.1}$$

Of course every (real) orbit $(x, T(x), \dots, T^n(x))$, $x \in X$, is an α-pseudo-orbit for every $\alpha \geq 0$. We shall prove a kind of a converse fact, that in case of open, distance expanding maps, each "sufficiently good" pseudo-orbit can be approximated (shadowed) by a real orbit. To make this precise we proceed as follows. Let $\beta > 0$. We say that an orbit of $x \in X$, *β-shadows* the pseudo-orbit $(x_i : i = 0, ..., n)$ if and only if for every $i = 0, \dots, n$

$$\rho(T^i(x), x_i) \leq \beta \tag{3.2.2}$$

Definition 3.2.0. We say that a continuous map $T : X \to X$ has *shadowing property* if for every $\beta > 0$ there exists $\alpha > 0$ such that for every finite n every α-pseudo-orbit can be β-shadowed by an orbit.

Note that due to the compactness of X this property implies the same with $n = \infty$ included.

Below is a simple observation on the uniqueness of the shadowing. Assume only that T is expansive (cf. Section 2.2.).

Proposition 3.2.1. If 2β is less than an expansiveness constant of T (we do not need to assume here that T is expanding with respect to the metric ρ) and $n = \infty$ then there exists at most one point x whose orbit β-shadows the pseudo-orbit $(x_i)_{i=1}^{\infty}$.

Proof. Suppose the forward orbits of x and y shadow (x_i). Then for every $n \geq 0$ we have $\rho(T^n(x), T^n(y)) \leq 2\beta$. Then by the definition of the expansiveness $x = y$. ♣

We shall now prove some less trivial results, concerning the existence of β-shadowing orbits.

Lemma 3.2.2. Let $T : X \to X$ be an open distance expanding map. Let $0 < \beta < \xi$, $0 < \alpha \leq \min\{(\lambda - 1)\beta, \xi\}$. If $(x_i : i = 0, 1, \dots, n)$, $0 \leq n \leq \infty$, is an α-pseudo-orbit and $x_i' = T_{x_i}^{-1}(x_{i+1})$, then

(a) For all $i = 0, 1, 2, \dots, n-1$

$$T_{x_i'}^{-1}(\overline{B(x_{i+1}, \beta)}) \subset \overline{B(x_i, \beta)}$$

and consequently for all $i = 0, 1, \dots, n$ the compositions

$$T_i = T_{x_0'}^{-1} \circ T_{x_1'}^{-1} \circ \dots \circ T_{x_{i-1}'}^{-1} : \overline{B(x_i, \beta)} \to X$$

are well-defined.

(b) The sequence of closed sets $T_i(\overline{B(x_i,\beta)})$, $i=0,1,\ldots,n$, is decreasing in the sense of inclusion.

(c) The intersection

$$\bigcap_{i=0}^{n} T_i(\overline{B(x_i,\beta)}$$

is non-empty and the forward orbits (for times $0,1,...,n$) of all the points of this intersection β-shadow the pseudo-orbit $(x_i : i=0,1,\ldots,n)$.

Proof. In order to prove (a) observe that by (3.1.8) and (3.1.7) we have

$$T_{x_i'}^{-1}(\overline{B(T(x_{i+1}),\beta)}) \subset \overline{B(x_i',\lambda^{-1}\beta)} \subset \overline{B(x_i,\lambda^{-1}\beta+\lambda^{-1}\alpha)}$$

and $\lambda^{-1}\beta+\lambda^{-1}\alpha \le \beta$. The statement (b) follows immediately from (a). The first part of (c) follows immediately from (b) and the compactness of the space X. To prove the second part call the intersection which appears in (c) by A. By the definition of A we have $T^i(A) \subset \overline{B(x_i,\beta)}$ for all $i=0,1,\ldots,n$. Thus the forward orbit for the times $0,1,...,n$ of every point in A, β shadows $(x_i : i=0,1,\ldots,n)$. The proof is finished. ♣

As an immediate consequence of Lemma 3.2.2 we get the following.

Corollary 3.2.3. (Shadowing lemma) Every open, distance expanding map satisfies the shadowing property. More precisely, for all $\beta>0$ and $\alpha>0$ as in Lemma 3.2.2 every α-pseudo-orbit $(x_i : i=0,\ldots,n)$ can be β-shadowed by an orbit in X.

As a consequence of Corollary 3.2.3 we shall prove the following.

Corollary 3.2.4. (Closing lemma) Let $T : X \to X$ be an expansive map, satisfying the shadowing property. Then for every $\beta>0$ there exists $\alpha>0$ such that if $x\in X$ and $\rho(x,T^l(x)) \le \alpha$ for some $l\ge 1$, then there exists a periodic point of period l whose orbit β-shadows the pseudo-orbit $(x,T(x),\ldots,T^{l-1}(x))$. The choices of α to β are the same as in the definition of shadowing, for 2β not exceeding the expansivness constant.

In particular the above holds for $T : X \to X$ open, expanding.

Proof. Since $\rho(x,T^l(x)) \le \alpha$, the sequence made up as the infinite concatenation of the sequence $(x,T(x),\ldots,T^{l-1}(x))$ is an α-pseudo-orbit. Hence, by shadowing with $n=\infty$, there is a point $y\in X$ whose orbit β-shadows this pseudo-orbit. But note that then the orbit of the point $T^l(y)$ also does it and therefore, by Lemma 3.2.1, $T^l(y)=y$. The proof is finished. ♣

Note that the assumption T is expansive is substantial. The adding machine map, see Ch.0.3' ???, satisfies the shadowing property, whereas it has no periodic orbits at all. In fact the same proof yields the following *periodic shadowing.*

Definition 3.2.5. We say that a continuous map $T : X \to X$ satisfies *periodic shadowing* property if for every $\beta > 0$ there exists $\alpha > 0$ such that for every finite n and every *periodic α-pseudo-orbit* $x_0, ..., x_{n-1}$, that is a sequence of points $x_0, ..., x_{n-1}$ such that $\rho(T(x_i), x_{(i+1)(\mathrm{mod} n)}) \le \alpha$, there exists a point $y \in X$ of period n such that for all $0 \le i < n \;\; \rho(T^i(y), x_i) \le \beta$.

Note that shadowing and periodic shadowing can hold for the maps that are not expansive. One can just add artificially the missing periodic orbits, of periods 2^n to the adding machine space. This example appears in fact as the nonwandering set for any Feigenbaum-like map of the interval, see Ch ??? (dawny 4.6)

§3.3 SPECTRAL DECOMPOSITION. MIXING PROPERTIES.

Let us start with general observations concerning iterations of continuous mappings

Definition 3.3.1 We call a continuous mapping $T : X \to X$ for a compact metric space X *topologically transitive* if for all non-empty open sets $U, V \subset X$ there exists $n \geq 0$ such that $T^n(U) \cap V \neq \emptyset$. By the compactness of X topological transitivity implies that T maps X onto X.

Example 3.3.2 Consider a topological Markov chain Σ_A, or $\tilde{\Sigma}_A$ in a one-sided or two-sided shift space of d states, see Example 0.3. Observe that the left shift map s on the topological Markov chain is topologically transitive iff the matrix A is *irreducible* that is for each i, j there exists an $n > 0$ such that the i, j-th entry $A^n_{i,j}$ of the n-th composition matrix A^n is non-zero.
One can consider a directed graph consisting of d vertices such that there is an edge from a vertex v_i to v_j iff $A_{i,j} \neq 0$; then one can identify elements of the topological Markov chain with infinite paths in the graph (that is sequences of edges indexed by all integers or nonnegative integers depending as we consider the two-sided or one-sided case, such that each edge begins at the vertex where the preceding edge ends). Then it is easy to see that A is irreducible iff for every two vertices v_1, v_2 there exists a finite path from v_i to v_j.

A notion stronger than the topological transitivity, which makes a non-trivial sense only for f non-invertible, is the following

Definition 3.3.2 A continuous mapping $T : X \to X$ for a compact metric space X is called *topologically exact* (or *locally eventually onto*) if for every open set $U \subset X$ there exists $n > 0$ such that $T^n(U) = X$.

In Example 3.3.2 in the one-sided shift space case topological exactness is equivalent to the property that there exists $n > 0$ such that the matrix A^n has all entries positive. Such a matrix is called *aperiodic*.

In the two-sided case aperiodicity of the matrix is equivalent to *topological mixing* of the shift map. We say a continuous map is *topologically mixing* if for every non-empty open sets $U, V \subset X$ there exists $N > 0$ such that for every $n \geq N$ we have $T^n(U) \cap V \neq \emptyset$.

Proposition 3.3.3 The following 3 conditions are equivalent:

(1) $T : X \to X$ is topologically transitive.
(2) For every non-empty open sets $U, V \subset X$ and every $N \geq 0$ there exists $n \geq N$ such that $T^n(U) \cap V \neq \emptyset$.
(3) There exists a T-trajectory $(x_n, n = 0, 1, ...)$, such that every $x \in X$ is its ω-limit point, that is for every $N \geq 0$ the set $\{x_n : n > N\}$ is dense in X.

Proof. Let us prove first the implication (1)$\Rightarrow$(3). So, suppose $T : X \to X$ is topologically transitive. Then for every open non-empty set $V \subset X$, the set

$$K(V) := \{x \in X : \text{there exists } n \geq 0 \text{ such that } \; T^n(x) \in V\} = \bigcup_{n \geq 0} T^{-n}(V)$$

is open and dense in X. Let $\{V_k\}_{k \geq 1}$ be a countable basis of topology of X. By Baire's category theorem, the intersection

$$K := \bigcap_{k \geq 1} \bigcap_{N \geq 0} K(T^{-N}(V_k))$$

is a dense G_δ subset of X. In particular K is non-empty and by its definition the trajectory $\{T^n(x) : n \geq N\}$ is dense in X for every $x \in K$. Thus (1) implies (3).

Let us now prove that (3)$\Rightarrow$(2). Indeed, if x_n is a trajectory satisfying the condition (3), then for all non-empty open sets $U, V \subset X$ and $N \geq 0$, there exist $n \geq m > 0, n - m \geq N$ such that $x_m \in U$ and $x_n \in V$. Hence $T^{n-m}(U) \cap V \neq \emptyset$. Thus (3) implies (2). Since (2) implies (1) trivially the proof is complete. ♣

Definition 3.3.a. A point $x \in X$ is called wandering if there exists an open neighhbourhood V of x such that $V \cap T^n(V) = \emptyset$ for all $n \geq 1$. Otherwise x is called non-wandering. We denote the set of all non-wandering points for T by Ω or $\Omega(T)$.

Proposition 3.3.b For $T : X \to X$ satisfying the periodic shadowing property, the set of periodic points is dense in the set Ω of non-wandering points.

Proof. Take any $x \in \Omega(T)$ and given $\beta > 0$ its neighborhood V in X of diameter α chosen for β in the definition of periodic shadowing. Then by the definition of $\Omega(T)$ there exists $y \in V$ and $n > 0$ such that $T^n(y) \in V$. So $rho(y, T^n(y)) \leq \operatorname{diam} V$ hence $(y, T(y), ..., T^n(y))$ can be β-shadowed by a periodic orbit. We can take β arbitrarily small hence we obtain the density of periodic points in $\Omega(T)$. ♣

Remark 3.3.c. It is not true that for every open, distance expanding map $T : X \to X$ we have $\overline{\text{Per}} = X$. Here is an example: Let $X = \{(1/2)^n : n = 0, 1, 2, ...\} \cup \{0\}$. Let $T((1/2)^n) = (1/2)^{(n-1)}$ for $n > 0$, $T(0) = 0, T(1) = 1$. Let the metric be the restriction to X of the standard metric on the real line. This T is distance expanding on X but $\Omega(T) = \text{Per}(T) = \{0\} \cup \{1\}$. See also Exercise 3.3.1.

Here is the main theorem of this section. Its assertion holds under the assumption that $T : X \to X$ is open, distance expanding and even under weaker assumptions below.

Theorem 3.3.4 (on the existence of Spectral Decomposition) Suppose that $T : X \to X$ is an open map which satisfies also the periodic shadowing property and is expanding at the set $\overline{\text{Per}(T)}$, the closure of the set of periodic points.

Then $\overline{\text{Per}(T)}$ is the union of finitely many disjoint compact sets $\Omega_j, j = 1, ..., J$ with

$$(T|_{\overline{\text{Per}(T)}})^{-1}(\Omega_j) = \Omega_j$$

and $T|_{\Omega_j}$ topologically transitive.

Each Ω_j is the union of $k(j)$ disjoint compact sets Ω_j^k which are cyclically permuted by T and such that $T^{k(j)}|_{\Omega_j^k}$ is topologically exact.

Proof of Theorem 3.3.4 Let us start with defining an equivalence relation $\sim$ on $\mathrm{Per}(T)$. For $x, y \in \mathrm{Per}(T)$ we write $x \sim_{\rightarrow} y$ if for every $\varepsilon > 0$ there exist $x' \in X$ and positive integer m such that $\rho(x, x') < \varepsilon$ and $T^m(x') = T^m(y)$. We write $x \sim y$ if $x \sim y$ and $y \sim x$. Of course for every $x \in \mathrm{Per}(T)$, $x \sim x$. Suppose that $x \sim y$ and $y \sim z$. Let k_y, k_z denote periods of y, z respectively.

Let x' be close to x and $T^n(x') = T^n(y) = y$; an integer n satisfiying the latter equality exists since we can take an integer so that the first equality holds and then take any larger integer divisible by k_y. Choose n divisible by $k_y k_z$. Next, since T is open, for y' close enough to y, with $T^m(y') = T^m(z) = z$ for m divisible by k_z, there exists x'' close to x' such that $T^m(x'') = y'$. Hence $T^{n+m}(x'') = T^m(y') = y' = T^{n+m}(y')$, since both m and n are divisible by k_z. Thus $x \sim z$. This proof is illustrated at Fig 3.1.a.

Fig.3.1.a

Fig.3.1.b

Fig.3.1.b illustrates the transitivity for hyperbolic sets $\overline{\mathrm{Per}(T)}$ (see Exercises or [KH] ???), where $x \sim y$ if the unstable manifold of x intersects transversally the stable manifold of y. In our expanding case the role of transversality is played by the openness of T.

Till this point we did not use the expanding assumption.

Observe now that for every $x, y \in \mathrm{Per}(T)$, $\rho(x, y) \leq \xi$ implies $x \sim y$. Indeed, we can take $x' = T_x^{-nk_x k_y}(y)$ for n arbitrarily large. Then x' is arbitrarily close to x and $T^{nk_x k_y}(x') = y = T^{nk_x k_y}(y)$. Hence the number of equivalence classes of $\sim$, denote them $P_1, ..., P_N$, is finite. Moreover the sets $\overline{P}_1, ..., \overline{P}_N$ are pairwise disjoint and the distances between them are at least ξ. We have $T(\mathrm{Per}(T)) = \mathrm{Per}(T)$, and if $x \sim y$ then $T(x) \sim T(y)$. The latter follows straight from the definition of $\sim$. So T permutes the sets P_i. This permutation decomposes into cyclic permutations we were looking for. More precisely: consider the partition of $\overline{\mathrm{Per}(T)}$ into the sets of the form

$$\bigcup_{n=0}^{\infty} T^n(\overline{P}_i), \ \ i = 1, ..., N.$$

The unions are in fact over finite families. It does not matter in which place the closure is placed because X is compact so for every $A \subset X$ we have $T(\overline{A}) = \overline{T(A)}$. We consider this partition as a partition into Ω_j's we were looking for. Ω_j^k's are the summands $T^n(\overline{P}_i)$ in the unions.

Observe now that T is topologically transitive on each Ω_j.

Indeed, if x, y belong to the same Ω_j there exist $x' \in B(x, \xi)$ and $y' \in B(y, \xi)$ such that $T^n(x') = T^{n_0}(y)$ and $T^m(y') = T^{m_0}(x)$ for some natural numbers n, m and $n_0 \leq k_y, m_0 \leq k_x$. For an arbitrary $\beta > 0$ choose $\alpha > 0$ from the definition of periodic shadowing and consider x'', y'' such that $\rho(x'', x) \leq \alpha$, $\rho(y'', y) \leq \alpha$ and $T^{n_1}(x'') = x'$, $T^{m_1}(y'') = y'$ for some natural numbers n_1, m_1, existing by the expanding property at $\overline{\mathrm{Per}(T)}$. Then the sequence of points $T(x''), ..., T^{n_1+n+k_y-n_0}(x''), T(y''), ..., T^{m_1+m+k_x-m_0}(y'')$ is a periodic α-pseudo-orbit, of period $n_1+n+k_y-n_0+m_1+m+k_x-m_0$, so it can be β-shadowed by a periodic orbit. Thus, there exists $z \in \mathrm{Per}(T)$ such that $\rho(z, x) \leq \beta$ and $\rho(T^N(z), y) \leq \beta$ for an integer $N > 0$. Now take arbitrary neighbourhoods $U \ni x$ and $V \ni y$ and take β such that $B(x, \beta) \subset U$ and $B(y, \beta) \subset V$. We find a periodic point z as above. Note that, provided $\beta \leq \xi$, $z \sim x$ and $T^N(z) \sim y$. We obtain $T^N(z) \in T^N(U \cap \Omega_j) \cap (V \cap \Omega_j)$ so this set is nonempty. This proves the topological transitivity.

Note that by the way we proved that the orbits $x'', ..., T^{n_1}(x'') = x', ..., T^n(x')$ with n_1, n arbitrarily large, can be arbitrarily well shadowed by parts of periodic orbits. This corresponds to the approximation of transversal cycles of heteroclinic orbits by periodic ones, in the hyperbolic theory (see also Exercise 3.3.3).

This analogy justifies the name *heteroclinic cycle points* for the points x' and y', or *heteroclinic cycle orbits* for their orbits discussed above. Thus we proved

Lemma 3.3.7'. Under the assumptions of Theorem 3.3.4 every heteroclinic cycle point is a limit of periodic points.

Now we can prove another fact interesting in itself:

Lemma 3.3.8 $T|_{\overline{\mathrm{Per}(T)}}$ is an open map.

Proof. Fix $x, y \in \overline{\mathrm{Per}(T)}$ and $\rho(T(x), y) \leq \varepsilon \leq \xi/3$. Since T is open, by Lemma 3.1.2, and due to the expanding property at $\overline{\mathrm{Per}(T)}$ there exists $\hat{y} = T_x^{-1}(y) \in B(x, \lambda^{-1}\xi/3)$ We want to prove that $\hat{y} \in \overline{\mathrm{Per}(T)}$.

There exist $z_1, z_2 \in \mathrm{Per}(T)$ such that $\rho(z_1, x) \leq \lambda^{-1}\xi/3$ and $\rho(z_2, y) \leq \xi/3$. Hence $\rho(T(z_1), z_2) \leq \xi$, hence $T(z_1) \sim z_2$. Then $T_x^{-1}(z_2)$ is a heteroclinic cycle point, so by Lemma 3.3.7' it is a limit of periodic points. ♣

We go back to Proof of Theorem 3.3.4. We can prove now the topological exactness of $T^{k(j)}|_{\Omega_j^k}$. So fix $\Omega_j^k = P_i$ with $T^{k(j)}(P_i) = P_i$. Let $\{x_s\}, s = 1, ...S$ be a $\xi'/2$-spanning set in P_i, where ξ' is a constant having the properties of ξ for the map $T|_{\overline{\mathrm{Per}}}$, existing by the openness of $T|_{\overline{\mathrm{Per}(T)}}$ (Lemma 3.1.2). Write $k(P_i) = \prod_{s=1}^S k_{x_s}$. Take an arbitrary open set $U \subset \overline{P}_i$. It contains a periodic point x.

Note that for every ball $B = B(y, r)$ in $\overline{\mathrm{Per}(T)}$ with the origin at $y \in \mathrm{Per}(T)$ and radius r less than η and $\lambda^{-k_y}\xi'$, we have $T^{k_y}(B) \supset B(y, \lambda^{k_y} r)$. Repeating this step by step we obtain $T^{nk(y)}(B) \supset B(y, \xi')$, (see (3.1.8).

Let us go back to U and consider $B_x = B(x, r) \subset U$ with $r \leq \lambda^{-k(P_i)}\xi'$. Then $T^{nk(P_i)}(B_x)$ is an increasing family of sets for $n = 0, 1, 2, ...$

By the definition of $\sim$, the set $\bigcup_{n \geq 0} T^{nk(P_i)}(B_x)$ contains $\{x_s :, s = 1, ...S\}$, because the points x_s are in the relation $\sim$ with x. This uses the fact proved above, see Lemma 3.3.7', that x' in the definition of $\sim$, such that $T^m(x') = T^m(x_s)$, belongs to $\overline{\mathrm{Per}(T)}$. It belongs even to P_i, since for $z \in \mathrm{Per}(T)$ close to x' we have $z \sim x_s$, with the use of the same

x' as one of a heteroclinic cycle points. Hence, by the observation above $\bigcup_{n\ge 0} T^{nk(P_i)}(B_x)$ contains the ball $B(x_s, \xi')$ for each s. So it contains $\overline{P}_i$. Since $T^{nk(P_i)}(B_x)$ is an increasing family of open sets in $\overline{\mathrm{Per}(T)}$ that is compact, just one of these sets covers $\overline{\mathrm{Per}(T)}$. The topological exactness is proved. ♣

Remark. It is easy to see that if T is a covering map than the assumption of periodic shadowing can be skipped. We used it only to approximate heteroclinic cycle points by periodic ones. See also Exercise 3.3.2.

As a corollary we obtain the following two theorems.

Theorem 3.3.9. Let $T : X \to X$ be an open distance expanding map, or expanding at the set $\overline{\mathrm{Per}(T)}$ satisfying the periodic shadowing property. Then, if T is topologically transitive, or is surjective and its spectral decomposition consists of just one set $\Omega_1 = \bigcup_{k=1}^{k(1)} \Omega_1^k$, the following properties hold:

1. The set of periodic points is dense in X, which is thus equal to Ω_1.
2. For every open $U \subset X$ there exists $N = N(U)$ such that $\bigcup_{j=0}^{N} T^j(U) = X$.
3. $(\forall r > 0)(\exists N)(\forall x \in X) \bigcup_{j=0}^{N} T^j(B(x, r)) = X$.
4. The following *specification property* holds: For every $\beta > 0$ there exists a positive integer N such that for every n and every T-orbit $(x_0, ...x_n)$ there exists a periodic point y of period not larger than $n + N$ whose orbit for the times $0, ..., n$ β-shadows $(x_0, ...x_n)$.

Proof. By the topological transitivity for all open U there exist $n \ge 1$ such that $T^n(U) \cap U \ne \emptyset$, (use the condition (2) in Proposition 3.3.3 for $N = 1$). Hence for the set Ω of the non-wandering points we have $\Omega = X$. This gives the density of $\mathrm{Per}(T)$ by Proposition 3.3.b.

If we assume only that there is one $\Omega_1 (= \Omega = \overline{\mathrm{Per}(T)})$ in the Spectral Decomposition, then for an arbitrary $z \in X$ we find by the surjectivity a backward orbit z_{-n} of z and notice that $z_{-n} \to \Omega$ and $T^n(z) \to \Omega$, that follows easily from the definition of Ω. So for every $\alpha > 0$ there exist $w_1, w_2 \in \mathrm{Per}(T)$ and natural numbers k, n such that $T^k(w_2) \sim w_1$, $\rho(w_1, z_{-n}) \le \alpha$ and $\rho(w_2, T^n(z)) \le \alpha$. This allows to find a periodic point in $B(z, \beta)$, where $\beta > 0$ is arbitrarily small and α chosen for β from the periodic shadowing property.

We conclude that $X = \bigcup_{j=1}^{J} \Omega_j$, each Ω_j is T-invariant, closed, and also open since Ω_j's are at least ξ-distant from each other. So $J = 1$. Otherwise, by the topological transitivity, for $j \ne i$ there existed n such that $T^n(\Omega_j) \cap \Omega_i \ne \emptyset$, what would contradict the T-invariance of Ω_j.

Thus $X = \bigcup_{k=1}^{k(1)} (\Omega_1^k)$ and the assertion 2. follows immediately from the exactness of $T^{k(1)}$ on each Ω_1^k, $k = 1, ..., k(1)$.

The property 3. follows from 2. where given r we choose $N = \max\{N(U)\}$ where we consider a finite covering of X by sets U of diameter not exceeding $r/2$. Indeed, then for every $B(x, r)$ the set U containing x is a subset of$B(x, r)$.

Now let us prove the specification property. By the property 3. for every $\alpha > 0$ there exists $N = N(\alpha)$ such that for every $v, w \in X$ there exists $m \le N$ and $z \in B(v, \alpha)$ such that $T^m(z) \in B(w, \alpha)$. Consider any T-orbit $x_0, ...x_n$. Then consider an α-pseudo-orbit $x_0, ...x_{n-1}, z, ..., T^{m-1}(z)$ with $m \leq N$ and $z \in B(x_n, \alpha, T^m(z) \in B(x_0, \alpha)$. By Corollary 3.2.4 we can β-shadow it by a periodic orbit of period $n + m \leq n + N$. ♣

The same proof yields

Theorem 3.3.10. Let T satisfies the assumptions of Theorem 3.3.9, and be also topologically mixing, i.e. $k(1) = 1$. Then

1. T is topologically exact, i.e. for every open $U \subset X$ there exists $N = N(U)$ such that $T^N(U) = X$.

2. $(\forall r > 0)(\exists N)(\forall x \in X)\ \ T^N(B(x, r)) = X$.

§3.4 HÖLDER CONTINUOUS FUNCTIONS

For distance expanding maps, Hölder continuous functions play a special role. Recall that a function $\phi : X \to \mathbb{C}$ (or $\mathbb{R}$) is said to be *Hölder continuous* with an exponent $0 < \alpha \le 1$ if and only if there exists $C > 0$ such that

$$|\phi(y) - \phi(x)| \le C\rho(y, x)^\alpha$$

for all $x, y \in X$. All Hölder continuous functions are continuous, if $\alpha = 1$ they are usually called Lipschitz continuous.

Let $C(X)$ denote as in the previous chapters the space of all continuous, real or complex-valued functions defined on a compact space X and for $\psi : X \to \mathbb{C}$ we write $\|\psi\|_\infty := \sup\{|\psi(x)| : x \in X\}$ for its supremum norm. For any $\alpha > 0$ let $\mathcal{H}_\alpha(X)$ denote the space of all Hölder continuous functions with exponent $\alpha > 0$. If $\psi \in \mathcal{H}_\alpha(X)$ let

$$\vartheta_{\alpha,\xi}(\psi) = \sup\left\{\frac{|\psi(y) - \psi(x)|}{\rho(y, x)^\alpha} : x, y \in X,\ x \ne y \text{ and } \rho(x, y) \le \xi\right\}$$

and

$$\vartheta_\alpha(\psi) = \sup\left\{\frac{|\psi(y) - \psi(x)|}{\rho(y, x)^\alpha} : x, y \in X, x \ne y\right\}.$$

Note that

$$\vartheta_\alpha(\psi) \le \max\left\{\frac{2||\psi||_\infty}{\xi^\alpha}, \vartheta_{\alpha,\xi}(\psi)\right\}.$$

The reader will check easily that $\mathcal{H}_\alpha(X)$ becomes a Banach space when equipped with the norm

$$\|\psi\|_{\mathcal{H}_\alpha} = \vartheta_\alpha(\psi) + \|\psi\|_\infty.$$

Thus, to estimate in future $||\psi||_{\mathcal{H}_\alpha}$ it is enough to estimate $\vartheta_{\alpha,\xi}(\psi)$ and $||\psi||_\infty$. The following result is a straightforward consequence of Arzela-Ascoli theorem.

Theorem 3.4.1. Any bounded subset of the Banach space $\mathcal{H}_\alpha(X)$ with the norm $\|\cdot\|_{\mathcal{H}_\alpha}$ is relatively compact as a subset of the Banach space $C(X)$ with the supremum norm $\|\cdot\|_\infty$. Moreover if $\{\psi_n : n = 1, 2, \ldots\}$ is a sequence of continuous functions in $\mathcal{H}_\alpha(X)$ such that $\|x_n\|_{\mathcal{H}_\alpha} \le C$ for all $n \ge 1$ and some constant C and if $\lim_{n\to\infty} \|\psi_n - \psi\|_\infty = 0$ for some $\psi \in C(X)$, then $\psi \in \mathcal{H}_\alpha(X)$ and $\|\psi\|_{\mathcal{H}_\alpha} \le C$.

Now let us formulate a simple but very basic lemma in which you will see a coherence of the expanding property of T and the Hölder continuity property of a function.

Lemma 3.4.2 (pre-Bounded Distortion Lemma for Iteration). Let $T : X \to X$ be a distance expanding map and $\phi : X \to \mathbb{C}$ be a Hölder continuous function with the exponent α. Then for every positive integer n and $x, y \in X$ such that

$$\rho(T^j(x), T^j(y)) < 2\eta \quad \text{for every} \quad j = 0, 1, ..., n-1 \tag{3.4.1}$$

we have

$$(3.4.2) \qquad |S_n\phi(x) - S_n\phi(y)| \le \rho(T^n(x), T^n(y))^\alpha \Big(\frac{\vartheta_\alpha(\phi)}{1-\lambda^{-\alpha}}\Big).$$

If T is open we can assume $x, y \in T_z^{-n}(B(T^n(z), \xi)$ for a point $z \in X$, instead of (3.4.1). Then in (3.4.2) we can replace ϑ_α by $\vartheta_{\alpha,\xi}$.

The sense of (3.4.2) is that the coefficient $\frac{\vartheta_\alpha(\phi)}{1-\lambda^{-\alpha}}$ does not depend on x, y, n).

Proof. By (3.1.2) we have $\rho(T^j(x), T^j(y)) \le \lambda^{-(n-j)}\rho(T^n(y), T^n(z))$ for every $0 \le j \le n$. Hence

$$|\phi(T^j(y)) - \phi(T^j(z))| \le \vartheta_\alpha(\phi)\lambda^{-(n-j)\alpha}\rho(T^n(y), T^n(z))^\alpha$$

Thus

$$\begin{aligned}|S_n\phi(y) - S_n\phi(z)| &\le \vartheta_\alpha(\phi)\rho(T^n(y), T^n(z))^\alpha \sum_{j=0}^{n-1} \lambda^{-(n-j)\alpha} \\ &\le \vartheta_\alpha(\phi)\rho(T^n(y), T^n(z))^\alpha \sum_{j=0}^{\infty} \lambda^{-j\alpha} = \frac{\vartheta_\alpha(\phi)}{1-\lambda^{-\alpha}}\rho(T^n(y), T^n(z))^\alpha\end{aligned}$$

The proof is finished. ♣

For an open distance expanding topologically transitive map we can replace topological pressure defined in Chapter 2 by a corresponding notion related with a "tree" of pre-images of an arbitrary point (compare this with Exercise 4 ??? in Chapter 2).

Proposition 3.4.3. If $T : X \to X$ is a topologically transitive distance expanding map, then for every Hölder continuous potential $\phi : X \to I\!R$ and for every $x \in X$ there exists the limit

$$\mathrm{P}_x(T, \phi) := \lim_{n\to\infty} \frac{1}{n} \log \sum_{\overline{x}\in T^{-n}(x)} \exp S_n\phi(\overline{x})$$

and it is equal to the topological pressure $\mathrm{P}(T, \phi)$. In addition, there exists a constant C such that for every $x, y \in X$ and every positive integer n

$$(3.4.3) \qquad \frac{\sum_{\overline{x}\in T^{-n}(x)} \exp S_n\phi(\overline{x})}{\sum_{\overline{y}\in T^{-n}(y)} \exp S_n\phi(\overline{y})} < C$$

Proof. If $\rho(x, y) < \xi$ then (3.4.3) follows immediately from Lemma 3.4.2 with some constant, say C_1. Now observe that by the topological transitivity of T there exists N (depending on ξ) such that for all $x, y \in X$ there exists $0 \le m < N$ such that $T^m(B(x, \xi)) \cap B(y, \xi) \ne \emptyset$. Indeed, for example by the condition 3) in Proposition 3.3.3 we can find two blocks of a trajectory of z with dense ω-limit set, say $T^k(z), ..., T^{k'}(z)$ and $T^l(z), ..., T^{l'}(z)$

with $l > k'$, each ξ-dense in X. Then we set $N = l' - k$. We can find t between k and k' and s between l and l' so that $T^t(z) \in B(x, \xi)$ and $T^s(z) \in B(y, \xi$. We have $m := s - t \leq N$.

Now fix arbitrary $x, y \in X$. So, there exists a point $y' \in T^{-m}(B(y,\xi)) \cap B(x,\xi)$. We then have

$$\begin{aligned}
\sum_{\overline{x}\in T^{-n}(x)} \exp S_n\phi(\overline{x}) &\leq C_1 \sum_{\overline{y'}\in T^{-n}(y')} \exp S_n\phi(\overline{y'}) \\
&= C_1 \exp(-S_m\phi(T^m(y'))) \sum_{\overline{y'}\in T^{-n}(y')} \exp S_{n+m}\phi(\overline{y'}) \\
&\leq C_1 \exp(-m \inf \phi) \sum_{\overline{y'}\in T^{-(n+m)}(T^m(y'))} \exp S_{n+m}\phi(\overline{y'}) \\
&\leq C_1 \exp(-m \inf \phi) \sum_{\overline{y'}\in T^{-(n+m)}(T^m(y'))} \exp S_n\phi(T^m(\overline{y'})) \exp S_m\phi(\overline{y'}) \\
&\leq C_1 \exp(m \sup \phi - m \inf \phi) \sum_{\overline{y'}\in T^{-(n+m)}(T^m(y'))} \exp S_n\phi(T^m(\overline{y'})) \\
&\leq C_1 \exp(2N||\phi||_\infty) D^N \sum_{\overline{y'}\in T^{-n}(T^m(y'))} \exp S_n\phi(\overline{y'}) \\
&\leq C_1^2 \exp(2N||\phi||) \sum_{\overline{y}\in T^{-n}(y)} \exp S_n\phi(\overline{y}),
\end{aligned}$$

where $D = \sup\{\#(T^{-1}(z)) : z \in X\} < \infty$. This proves (3.4.3).

Observe that each set $T^{-n}(x)$ is $(n, 2\eta)$-separated, whence

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{\overline{x}\in T^{-n}(x)} \exp S_n\phi(\overline{x}) \leq \mathrm{P}(T,\phi),$$

by the characterization of pressure given in Theorem 2.2.10.

In order to prove the opposite inequality fix $\varepsilon < 2\xi$ and for every $n \geq 1$, an (n,ε)-separated set F_n. Cover X by finitely many balls $B(z_1, \varepsilon/2), B(z_2, \varepsilon/2), \ldots, B(z_k, \varepsilon/2)$. Then $F_n = F_n \cap \left(\bigcup_{j=1}^k T^{-n}\big(B(z_j, \varepsilon/2)\big)\right)$ and therefore

$$\sum_{z\in F_n} \exp(S_n\phi(z)) = \sum_{j=1}^{k} \sum_{F_n\cap T^{-n}(B(z_j,\varepsilon/2))} \exp(S_n\phi(z)).$$

Given $y \in X$ choose as $j(y)$ an arbitrary j such that $y \in T^{-n}(B(z_{j(y)}, \varepsilon/2))$. Let $\overline{z_{j(y)}} \in T^{-n}(z)$ be defined by $y \in T^{-n}_{\overline{z_{j(y)}}}(B(z_{j(y)}, \varepsilon/2)$. We shall show that the function $y \mapsto \overline{z_{j(y)}}$ is injective. Indeed, suppose that $\overline{z_j} = \overline{z_{j(a)}} = \overline{z_{j(b)}}$ for some $a, b \in F_n \cap T^{-n}(B(z_j, \varepsilon/2))$. Then

$$\rho(T^l(a), T^l(b)) \leq \rho(T^l(a), T^l(\overline{z_j})) + \rho(T^l(\overline{z_j}), T^l(b)) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for every $0 \le l \le n$. So, $a = b$ since F_n is (n, ε)-separated.

Hence, using (3.4.3), we obtain

$$\sum_{z \in F_n} \exp(S_n\phi(z)) \le \sum_{j=1}^{k} C \sum_{\overline{z_j}} \exp(S_n\phi(\overline{z_j})) \le kC^2 \sum_{\overline{x} \in T^{-n}(x)} \exp(S_n\phi(\overline{x}))$$

Letting $n \nearrow \infty$ and next $\varepsilon \to 0$, applying Theorem 2.2.10, we therefore get

$$\mathrm{P}(T, \phi) \le \liminf_{n\to\infty} \frac{1}{n} \log \sum_{\overline{x} \in T^{-n}(x)} \exp S_n\phi(\overline{x}).$$

Thus

$$\liminf_{n\to\infty} \frac{1}{n} \log \sum_{\overline{x} \in T^{-n}(x)} \exp S_n\phi(\overline{x}) \ge \mathrm{P}(T, \phi) \ge \limsup_{n\to\infty} \frac{1}{n} \log \sum_{\overline{x} \in T^{-n}(x)} \exp S_n\phi(\overline{x}).$$

So liminf=limsup above, the limit exists and is equal to $\mathrm{P}(T, \phi)$. ♣

Remark 3.4.4. It follows from Proposition 3.4.3, the proof of the Variational Principle Part II (see Section 2.3) and the expansiveness of T that for every $x \in X$ every weak limit of the measures $\frac{1}{n}\sum_{k=0}^{n-1} \mu_n \circ T^{-k}$ where

$$\mu_n = \frac{\sum_{\overline{x} \in T^{-n}(x)} \delta_x \exp S_n\phi(\overline{x})}{\sum_{\overline{x} \in T^{-n}(x)} \exp S_n\phi(\vee x)}$$

and δ_x denotes the Dirac measure concentrated at the point x, is an equilibrium state. In fact our very special situation allows to say a lot more about the measures involved. Chapter 3 will be devoted to this end.

Let us finish this section with one more very useful fact (compare Theorem 1.11.3.)

Proposition 3.4.5. Let $T : X \to X$ be an open, distance expanding, topologically transitive map. If $\phi, \psi \in \mathcal{H}_\alpha(X)$, then the following conditions are equivalent.

(1) If $x \in X$ is a periodic point of T and if n denotes its period, then $S_n\phi(x) - S_n\psi(x) = 0$.
(2) There exists a constant $C > 0$ such that for every $x \in X$ and integer $n \ge 0$, we have $|S_n\phi(x) - S_n\psi(x)| \le C$.
(3) There exists a function $u \in \mathcal{H}_a$ such that $\phi - \psi = u \circ T - u$.

Proof. The implications (3) $\Longrightarrow$ (2) $\Longrightarrow$ (1) are very easy. The first one is obtained by summing up the equation in (3) along the orbit $x, T(x), ..., T^{n-1}(x)$ which gives $C = 2\sup|\phi - \psi|$. The second one holds because otherwise, if $S_n\phi(x) - S_n\psi(x) = K \ne 0$ for x of period n, then we have $S_{jn}\phi(x) - S_{jn}\psi(x) = jK$ which contradicts (2) for j large enough. Now let us prove (1) $\Longrightarrow$ (3). Let $x \in X$ be a point such that for every $N \ge 0$ the orbit $(x_n : n = N, N+1, ...)$ is dense in X. Such x exists by topological transitivity of T, see Proposition 3.3.3. Write $\eta = \phi - \psi$. Define u on the forward orbit of x, the set $A = \{x_n : n = 0, 1, ...\}$ by $u(x_n) = S_n\eta(x)$. If x is periodic then X is just the orbit of x and the function u is well defined due to the equality in (1). So, suppose that x is not periodic. Then $x_n \neq x_m$ for $m \neq n$ hence u is well defined on A. We will show that it extends in a Hölder continuous manner to $\overline{A} = X$. Indeed, if we take points $x_m, x_n \in A$ such that $m < n$ and $\rho(x_m, x_n) < \varepsilon$ for ε small enough, then $x_m, ..., x_{n-1}$ can be β-shadowed by a periodic orbit $y, ..., T^{n-m-1}(y)$ of period $n - m$ by Corollary 3.2.4, where ε is related to β in the same way as α related to β in that Corollary. Then by the Lemma 3.4.2

$$\begin{aligned} |u(x_n) - u(x_m)| &= |S_n\eta(x) - S_m\eta(x)| = |S_{n-m}\eta(x_m)| \\ &= |S_{n-m}\eta(x_m) - S_{n-m}\eta(y)| \leq \vartheta(\phi)_\alpha \varepsilon^\alpha. \end{aligned}$$

In particular we proved that u is uniformly continuous on A which allows to extend u continuously to $\overline{A}$. By taking limits we see that this extension satisfies the same Hölder estimate on $\overline{A}$ as on A. Also the equality in (3) true on A, extends to $\overline{A}$ by the definition of u and by the continuity of η and u . The proof is finished. ♣

The equality in (3) is called *cohomology equation*, u is a solution of the equation, see Ch.1.11.2. Here the cohomology equation is solvable in the space $\mathcal{K} = \mathcal{H}_\alpha$. Note that proving 3) $\Longrightarrow$ 2) we used only the assumption that u is bounded. So, going through 2) $\Longrightarrow$ 1) $\Longrightarrow$ 3) we prove that if the cohomology equation is solvable with u bounded, then automatically $u \in \mathcal{H}_\alpha$. Later on ??? you will see that an assumption that u is finite measurable, for some probability T-invariant measure with support X, would be sufficient, even under assumptions on T weaker than expanding. Often u is forced to be as good as ϕ and ψ. This type of theorem is called *Livśic type theorem.*

§3.5 MARKOV PARTITIONS AND SYMBOLIC REPRESENTATION

We shall prove in this section that the topological Markov chains (Ch.0.3) describe quite precisely dynamics of general open expanding maps.

This can be done through so called Markov partitions of X. The sets of a partition will play the role of "cylinders" $\{i_0 = \text{ Const }\}$ in Σ_A.

Definition 3.5.1. A finite cover $\Re = \{R_1, \ldots, R_n\}$ of X is said to be a *Markov partition* of the space X for the mapping T if $\operatorname{diam}(\Re) < \min\{\eta, \xi\}$ and the following conditions are satisfied.

(a) $R = \overline{\operatorname{Int} R_i}$ for all $i = 1, 2, \ldots, d$

(b) $\operatorname{Int} R_i \cap \operatorname{Int} R_j = \emptyset$ for all $i \neq j$

(c) $\operatorname{Int} R_j \cap T(\operatorname{Int} R_i) \neq \emptyset \implies R_j \subset T(R_i)$ for all $i, j = 1, 2, \ldots, d$

Theorem 3.5.2. For the open, distance expanding mapping T there exist Markov partitions of arbitrarily small diameters.

Proof. Fix $\beta < \min\{\eta/4, \xi\}$ and let α be the number associated to β as in Lemma 3.2.2. Choose $0 < \gamma \leq \min\{\beta/2, \alpha/2\}$ so small that

$$\rho(x,y) \leq \gamma \implies \rho(T(x), T(y)) \leq \alpha/2 \tag{3.5.1}$$

and let $E = \{z_1, \ldots, z_r\}$ be a γ-spanning set of X. Define the space Ω putting

$$\Omega = \{q = (q_i) \in E^{\boldsymbol{Z}^+} : \rho(T(q_i), q_{i+1}) \leq \alpha \text{ for all } i \geq 0\}$$

By definition all elements of the space Ω are α-pseudo-orbits and therefore in view of Corollary 3.2.3 and Lemma 3.2.1 for every sequence $q \in \Omega$ there exists a unique point whose orbit for $n = 0, 1, \ldots$ β-shadows q. Denote this point by $\Theta(q)$. In this way we have defined a map $\Theta : \Omega \to X$. We will need some of its properties.

Let us show first that Θ is surjective. Indeed, since E is a γ spanning set, for every $x \in X$ and every $i \geq 0$ there exists $q_i \in E$ such that

$$\rho(T^i(x), q_i) < \gamma$$

and therefore, using also (3.5.1),

$$\rho(T(q_i), q_{i+1}) \leq \rho(T(q_i), T(T^i(x))) + \rho(T^{i+1}(x), q_{i+1}) < \alpha/2 + \gamma \leq \alpha/2 + \alpha/2 = \alpha$$

for all $i \geq 0$. Thus $q = (q_i : i = 0, 1, \ldots) \in \Omega$ and (as $\gamma < \beta$) $x = \Theta(q)$. The surjectivity of Θ is proved.

Now we shall show that Θ is continuous. For this aim we will need the following notation. If $q \in \Omega$ then we put

$$q(n) = \{p \in \Omega : p_i = q_i \text{ for every } i = 0, 1, \ldots, n\} \tag{3.5.2}$$

To prove continuity suppose now that $p, q \in \Omega$, $p(n) = q(n)$ with some $n \geq 0$ and denote $x = \Theta(q)$, $y = \Theta(p)$. Then for all $i = 0, 1, \ldots, n$

$$\rho(T^i(x), T^i(y)) \leq \rho(T^i(x), q_i) + \rho(p_i, T^i(y)) \leq \beta + \beta = 2\beta$$

As $\beta < \xi$, we therefore obtain by (3.1.2) that $\rho(T^{i+1}(x), T^{i+1}(y)) \geq \lambda \rho(T^i(x), T^i(y))$ for $i = 0, 1, \ldots, n-1$, (see (3.1.7)), and consequently $\rho(x, y) \leq \lambda^{-n} 2\beta$. The continuity of Θ is proved.

Now for every $k = 1, \ldots, r$ define the sets

$$P_k = \Theta(\{q \in \Omega : q_0 = z_k\})$$

Since Θ is continuous, Ω is a compact space, and the sets $\{q \in \Omega : q_0 = z_k\}$ are closed in Ω, all sets P_k are closed in X.

Denote

$$W(k) = \{l : \rho(T(z_k), z_l) \leq \alpha\}$$

We have the following basic property satisfied:

$$T(P_k) = \bigcup_{l \in W(k)} P_l \tag{3.5.3}$$

Indeed, if $x \in P_k$ then $x = \Theta(q)$ for $q \in \Omega$ with $q_0 = z_k$. By the definition of Ω we have $q_1 = z_l$ for some $l \in W(k)$. We obtain $T(x) \in P_l$.

Conversely, let $x \in P_l$ for $l \in W(k)$. It means that $x = \Theta(q)$ for some $q \in \Omega$ with $q_0 = z_l$. By the definition of $W(k)$ the concatenation $z_k q$ belongs to Ω and therefore the point $T(\Theta(z_k q))$ β-shadows q. Thus $T(\Theta(z_k q)) = \Theta(q) = x$ hence $x \in T(P_k)$.

Let now

$$Z = X \setminus \bigcup_{n=0}^{\infty} T^{-n}\big(\bigcup_{k=1}^{r} \partial P_k\big)$$

and for any $x \in Z$ denote

$$P(x) = \{k \in \{1, \ldots, r\} : x \in P_k\},$$

$$Q(x) = \{l \notin P(x) : P_l \cap \big(\bigcup_{k \in P(x)} P_k\big) \neq \emptyset\},$$

and

$$S(x) = \bigcap_{k \in P(x)} \mathrm{Int} P_k \setminus \Big(\bigcup_{k \in Q(x)} P_k\Big) = \bigcap_{k \in P(x)} \mathrm{Int} P_k \setminus \Big(\bigcup_{k \notin P(x)} P_k)\Big)$$

We shall show that the family $\{S(x) : x \in Z\}$ is in fact finite and moreover, that the family $\{\overline{S(x)} : x \in Z\}$ is a Markov partition of diameter not exceeding 2β.

Indeed, since $\mathrm{diam}(P_k) \leq 2\beta$ for every $k = 1, \dots, r$ we have

$$\mathrm{diam}(S(x)) \leq 2\beta \tag{3.5.4}$$

As the sets $S(x)$ are open, we have

$$\overline{\mathrm{Int}\overline{S(x)}} = \overline{S(x)} \tag{3.5.5}$$

for all $x \in Z$. This proves the property (a) of the Theorem.

We shall now show that for every $x \in Z$

$$T(S(x)) \supset S(Tx) \tag{3.5.6}$$

Note first that for $K(x) := \bigcup_{k \in P(x)} P_k \cup \bigcup_{l \in Q(x)} P_l$ we have $\mathrm{diam} K \leq 8\beta$ and therefore by the assumption $\beta < \eta/4$, the map T restricted to K is injective.

Consider $k \in P(x)$. Then there exists $l \in W(k)$ such that $T(x) \in P_l$ cf. (3.5.3), and using the definition of Z we get $T(x) \in \mathrm{Int}(P_l)$. Using the assumption that T is open and next (3.5.3) we obtain

$$T(\mathrm{Int}P_k) = \mathrm{Int}(T(P_k)) \supset \mathrm{Int}P_l \supset S(T(x))$$

and therefore

$$T\Big(\bigcap_{k \in P(x)} \mathrm{Int}P_k\Big) \supset S(T(x)) \tag{3.5.7)}$$

Now consider $k \in Q(x)$. We remind (3.5.3) and observe that by the injectivity of $T|_K$ the assumption $x \notin P_k$ implies $T(x) \notin P_l$, $l \in W(k)$.

Thus

$$T(P_k) \subset \bigcup_{l \notin P(T(x))} P_l$$

hence

$$T\Big(\bigcup_{l \in Q(x)} P_l\Big) \cap S(T(x)) = \emptyset$$

Combining this and (3.5.7) gives

$$T\Big(\bigcap_{k \in P(x)} \mathrm{Int}P_k \setminus \Big(\bigcup_{k \in Q(x)} P_k\Big)\Big) \supset S(T(x))$$

which exactly means that formula (3.5.6) is satisfied and therefore

$$T(\overline{S(x)}) \supset \overline{S(Tx)} \tag{3.5.8}$$

We shall now prove the following claim.

Claim. If $x, y \in Z$ then either $S(x) = S(y)$ or $S(x) \cap S(y) = \emptyset$.

Indeed, if $P(x) = P(y)$ then also $Q(x) = Q(y)$ and consequently $S(x) = S(y)$. If $P(x) \neq P(y)$ then there exists $k \in P(x) \div P(y)$, say $k \in P(x) \setminus P(y)$. Hence $S(x) \subset \mathrm{Int} P_k$ and $S(y) \subset X \setminus P_k$. Therefore $S(x) \cap S(y) = \emptyset$ and the Claim is proved.

(One can write the family $S(x)$ as $\bigvee_{k=1,\dots,r}\{\mathrm{Int}P_k, X \setminus P_k\}$, compare notation in Ch.1. Then the assertion of the Claim is immediate.)

Since the family $\{P(x) : x \in Z\}$ is finite so is the family $\{S(x) : x \in Z\}$. Note that $S(x) \cap S(y) = \emptyset$ implies $\mathrm{Int}\overline{S(x)} \cap \mathrm{Int}\overline{S(y)} = \emptyset$. This is a general property of pairs of open sets, $U \cap V = \emptyset$ implies $\overline{U} \cap V = \emptyset$ implies $\mathrm{Int}\overline{U} \cap V = \emptyset$ implies $\mathrm{Int}\overline{U} \cap \overline{V} = \emptyset$ implies $\mathrm{Int}\overline{U} \cap \mathrm{Int}\overline{V} = \emptyset$.

In view of Baire's theorem the set Z is dense in X. Since $\bigcup_{x \in Z} S(x) \supset Z$, we thus have $\bigcup_{x\in Z} \overline{S(x)} = X$. That the family $\{\overline{S(x)} : x \in Z\}$ is a Markov partition for T of diameter not exceeding 2β follows now from (3.5.5), (3.5.6), (3.5.4) and from the claim. The proof is finished. ♣

Each Markov partition allows to introduce a coding (symbolic representation) of $T : X \to X$ as follows.

Theorem 3.5.3. Let $T : X \to X$ be an open, distance expanding map. Let $\{R_1, ..., R_d\}$ be a Markov partition. Let $A = (a_{i,j})$ be a $d \times d$ matrix with $a_{i,j} = 0$ or 1, $a_{i,j} = 1$ iff $T(\mathrm{Int}R_i) \cap \mathrm{Int}R_j \neq \emptyset$. Then consider the one-sided topological Markov chain Σ_A with the left shift σ, see Ch.0.3. Define a mapping $\pi : \Sigma_A \to X$ by

$$\pi((i_0, i_1, ...)) = \bigcap_{n=0}^{\infty} T^{-n}(R_{i_n}).$$

Then π is well defined Hölder continuous mapping onto X and $T\pi = \pi\sigma$. Moreover $\pi|_{\pi^{-1}(X \setminus \bigcup_{n=0}^{\infty} T^{-n}(\bigcup_i \partial R_i))}$ is injective.

Proof. For an arbitrary sequence $(i_0, i_1, ...) \in \Sigma_A$, $a_{i,j} = 1$ implies $T(R_{i_n}) \supset R_{i_{n+1}}$. Since $\mathrm{diam} R_{i_n} < 2\eta$, T is injective on R_{i_n}, hence there exists an inverse branch $T^{-1}_{R_{i_n}}$ on $R_{i_{n+1}}$ The subscript R_{i_n} indicates that we take the branch leading to R_{i_n}, compare notation from Ch.3.1. Thus, $T^{-1}_{R_{i_n}}(R_{i_{n+1}}) \subset R_{i_n}$. Hence

$$T^{-1}_{R_{i_0}} T^{-1}_{R_{i_1}} ... T^{-1}_{R_{i_n}}(R_{i_{n+1}}) \subset T^{-1}_{R_{i_0}} T^{-1}_{R_{i_1}} ... T^{-1}_{R_{i_{n-1}}}(R_{i_n}).$$

So $\bigcap_{n\geq 0} T^{-n}(R_{i_n}) \neq \emptyset$, as the intersection of the above decreasing family of compact sets. We have used here

$$T^{-1}_{R_{i_0}} ... T^{-1}_{R_{i_{n-1}}}(R_{i_n}) = T^{-1}_{R_{i_0}} ... T^{-1}_{R_{i_{n-2}}}(T^{-1}(R_{i_n}) \cap R_{i_{n-1}})$$

$$= T^{-1}_{R_{i_0}} ... T^{-1}_{R_{i_{n-3}}}(T^{-2}(R_{i_n}) \cap T^{-1}R_{i_{n-1}} \cap R_{i_{n-2}}) = ... = \bigcap_{k=0}^{n} T^{-k}(R_{i_k})$$

following from $T_{R_{i_k}}^{-1}(A) = T^{-1}(A) \cap R_{i_k}$ for every $A \subset R_{i_{k+1}}$, $k = 0, ..., n-1$.

Our infinite intersection consists of only one point, since $\mathrm{diam} R_i$ are less than the expansivness constant.

Let us prove now that π is Hölder continuous. Indeed, $\mathrm{dist}((i_n),(i'_n)) \le \lambda_1^{-N}$ implies $i_n = i'_n$ for all $n = 0, ..., N-1$, where we consider distance in the metric ρ' in Ch.0.3, with the factor λ'. Then, for $x = \pi((i_n))$, $y = \pi((i'_n))$ and every $n : 0 \le n < N$ we have $T^n(x), T^n(y) \in R_{i_n}$, hence $\mathrm{dist}(T^n(x), T^n(y)) \le \mathrm{diam} R_{i_n} \le \xi$, hence $\mathrm{dist}(x,y) \le \lambda^{-(N-1)}\xi$. Therefore π is Hölder with exponent $\min\{1, \log\lambda'/\log\lambda\}$.

Finally let us deal with the injectivity. If $x = \pi((i_n))$ and $T^n(x) \in \mathrm{Int} R_{i_n}$ for all $n = 0, 1, ...$, then $T^n(x) \notin R_j$ for all $j \ne i_n$. So, if $x \in \bigcap_n T^{-n}(R_{i'_n})$, then all $i'_n = i_n$. ♣

Remark. One would not think that π is always injective on the whole Σ_A. Consider for example the mapping of the unit interval $T(x) = 2x (\mathrm{mod}\ 1)$, compare Ch.0.3. Then dyadic expansion of x is not unique for $x \in \bigcup_{n=0}^{\infty} T^{-n}(\{\frac{1}{2}\})$. Dyadic expansion is the inverse, π^{-1}, of the coding obtained from the Markov partition $[0,1] = \{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$.

Remind finally that $\sigma : \Sigma_A \to \Sigma_A$ is an open, distance expanding map. The partition into the cylinders $C_i := \{(i_n) : i_0 = i\}$ for $i = 1, ..., d$, is a Markov partition into closed-open sets. The corresponding coding π is just the identity.

Another fact concerning a similarity between (Σ_A, σ) and (X, T) is the following

Theorem 3.5.4. For every Hölder continuous function $\phi : X \to \mathbb{R}$ the function $\phi \circ \pi$ is Hölder continuous on Σ_A and the pressures coincide, $\mathrm{P}(T, \phi) = \mathrm{P}(\sigma, \phi \circ \pi)$.

Proof. The functio $\pi \circ \Phi$ is Hölder as a composition of Hölder continuous functions. Consider next an arbitrary $x \in X \setminus \bigcup_{n=0}^{\infty} T^{-n}(\bigcup_i \partial R_i)$. Then, using Proposition 3.4.3 for T and σ we obtain

$$\mathrm{P}(T, \phi) = \mathrm{P}_x(T, \phi) = \mathrm{P}_{\pi^{-1}(x)}(\sigma, \phi \circ \pi) = \mathrm{P}(\sigma, \phi \circ \pi).$$

The middle equality follows directly from the definitions. ♣

Finally we shall prove that π is injective in a measure-theoretic sense.

Theorem 3.5.5. For every ergodic, invariant under the shift σ, probability Borel measure μ on Σ_A, positive on open sets, the mapping π yields an isomorphism between μ and the measure $\mu \circ \pi^{-1}$ on the Borel sets in X.

Proof. The set $\partial = \bigcup_{i=1}^{d} \partial(R_i)$, and hence $\pi^{-1}(\partial)$, have non-empty open complements in Σ_A. We have also $T(\partial) \subset \partial$ hence $\sigma(\pi^{-1}(\partial)) \subset \pi^{-1}(\partial)$. Hence, by the σ-invariance of μ we get $\mu(\pi^{-1}(\partial)) = \mu(\sigma(\pi^{-1}(\partial)))$, equal to 0 or 1 by the ergodicity. But the complement of $\pi^{-1}(\partial)$, as a non-empty open set, has positive measure μ. Hence $\mu(\pi^{-1}(\partial)) = 0$. Hence $\mu(E) = 0$ for $E := \bigcup_{n=0}^{\infty} T^{-n}(\pi^{-1}(\partial))$ and by Theorem 3.5.3 π is injective on $\Sigma_A \setminus E$. This proves that π is the required isomorphism. ♣

§3.6 EXPANSIVE MAPS ARE EXPANDING IN SOME METRIC

Theorem 3.1.1 says that distance expanding maps are expansive. In this section we prove the following much more difficult result which can be considered as a sort of the converse statement and which provides an additional strong justification to explore expanding maps.

Theorem 3.6.1. If a continuous map $T : X \to X$ of a compact metric space X is (positively) expansive then there exists a metric on X, compatible with the topology, such that the mapping T is distance expanding with respect to this metric.

The proof of Theorem 3.6.1 given here relies heavily on the old topological result of Frink (see [Frn], comp.[K, p.185]) which we state below without proof.

Lemma 3.6.2. (The Metrization Lemma of Frink) Let $\{U_n : n \geq 0\}$ be a sequence of open neighborhoods of the diagonal $\Delta \subset X \times X$ such that $U_0 = X \times X$,

$$\bigcup_{n=1}^{\infty} U_n = \Delta, \tag{3.6.1}$$

and for every $n \geq 1$

$$U_n \circ U_n \circ U_n \subset U_{n-1} \tag{3.6.2}$$

Then there exists a metric ρ, compatible with the topology on X, such that for every $n \geq 1$

$$U_n \subset \{(x, y) : \rho(x, y) < 2^{-n}\} \subset U_{n-1}. \tag{3.6.3}$$

We will also need the following almost obvious result.

Lemma 3.6.3. If $T : X \to X$ is a continuous map of a compact metric space X and T^n is distance expanding for some $n \geq 1$ then T is distance expanding with respect to some some metric compatible with the topology on X.

Proof. Let ρ be a compatible metric with respect to which T is distance expanding and let $\lambda > 1$ and $\eta > 0$ be constants such that

$$\rho(T^n(x), T^n(y)) \geq \lambda\rho(x, y)$$

whenever $\rho(x, y) < \eta$. Put $\xi = \lambda^{\frac{1}{n}}$ and define the new metric ρ' setting

$$\rho'(x, y) = \rho(x, y) + \frac{1}{\xi}\rho(T(x), T(y)) + \ldots + \frac{1}{\xi^{n-1}}\rho(T^{n-1}(x), T^{n-1}(y))$$

Then ρ' is a metric on X compatible with the topology and $\rho'(T^(x), T^(y)) \geq \xi\rho'(x, y)$ whenever $l\rho'(x, y) < \eta$. ♣

Now we can pass to the proof of Theorem 3.6.1.

Proof of Theorem 3.6.1. Let d be a metric on X compatible with the topology, and let $3\theta > 0$ be an expansive constant associated to T which does not exceed the constant η claimed in Proposition 2.4.9. For any $n \geq 1$ and $\gamma > 0$ let

$$V_n(\gamma) = \{(x, y) \in (X \times X) : d(T^j(x), T^j(y)) < \gamma \quad \text{for every } j = 0, \ldots, n\}.$$

Then in view of Proposition 2.4.9 there exists $M \geq 1$ such that

$$V_M(3\theta) \subset \{(x, y) : d(x, y) < \theta\}. \tag{3.6.4}$$

Define $U_0 = X \times X$ and $U_n = V_{Mn}(\theta)$ for every $n \geq 1$. We will check that the sequence $\{U_n : n \geq 0\}$ satisfies the assumptions of Lemma 3.6.2. Indeed, (3.6.1) follows immediately from expasiveness of T and condition (3.6.2) will be proved by induction. For $n = 1$ nothing has to be proved. Suppose that (3.6.2) holds for some $n \geq 1$ and let $(x, u),\ (u, v),\ (v, y) \in U_{n+1}$. Then by the triangle inequality

$$d(T^j(y), T^j(x)) < 3\theta \qquad \text{for every } j = 0, \ldots, (n + 1)M.$$

Therefore, using (3.6.4), we conclude that

$$d(T^j(y), T^j(x)) < \theta \qquad \text{for every } j = 0, \ldots, Mn$$

Equivalently $(x, y) \in V_{Mn}(\theta) = U_n$ which finishes the proof of (3.6.2).

So we have shown that the assumptions of Lemma 3.6.2 are satisfied, and therefore we obtain a compatible metric ρ on X satisfying (3.6.3). In view of Lemma 3.6.3 it sufficies to show that T^{3M} is expanding with respect to the metric ρ. So suppose that $0 < \rho(x, y) < \frac{1}{16}$. Then by (3.6.1) there exists an $n \geq 0$ such that

$$(x, y) \in U_n \setminus U_{n+1}. \tag{3.6.5}$$

As $0 < \rho(x, y) < \frac{1}{16}$, this and (3.6.3) imply that $n \geq 3$. It follows from (3.6.5) and the definitions of U_n and $V_{Mn}(\theta)$, that there exists $Mn < j \leq (n + 1)M$ such that $d(T^j(y), T^j(x)) \geq \theta$. Since $3 \leq n$ we conclude that $d(T^i(T^{3M}(x)), T^i(T^{3M}(y))) \geq \theta$ for some $0 \leq i \leq (n - 2)M$ and therefore $(T^{3M}(x), T^{3M}(y)) \notin U_{n-2}$. Consequently, by (3.6.3) and (3.6.5) we obtain that

$$\rho(T^{3M}(x), T^{3M}(y)) \geq 2^{-(n-1)} = 2 \cdot 2^{-n} > 2\rho(x, y).$$

The proof is finished. ♣

Exercises.

Exercise 3.2.1. Prove the following *Shadowing Theorem* generalizing Corollary 3.2.3 (Shadowing lemma) and Corollary 3.2.4 (Closing lemma):

Let $T : X \to X$ be an open map, expanding at a compact $Y \subset X$. Then, for every $\beta > 0$ there exists $\alpha > 0$ such that for every map $\Gamma : Z \to Z$ for a set Z and a map $\Phi : Z \to B(Y, \alpha)$ satisfying $\rho(T\Phi(z), \Phi\Gamma(z)) \leq \alpha$ for every $z \in Z$, there exists a map $\Psi : Z \to X$ satisfying $T\Phi = \Phi\Gamma$, hence $T(Y') \subset Y'$ for $Y' = \Psi(Z)$, and such that for every $z \in Z$, $\rho(\Psi(z), \Phi(z)) \leq \beta$. If Z is a metric space and Γ, Φ are continuous, then Ψ is continuous. If $T(Y) \subset Y$ and the map $T|_Y : Y \to Y$ be open, then $Y' \subset Y$.

(Hint: see Ch.5.1.)

Exercise 3.2.2. Prove the following *structural stability theorem.*

Let $T : X \to X$ be an open map with a compact $Y \subset X$ such that $T(Y) \subset Y$. Then for every $\lambda > 1$ and $\beta > 0$ there exists $\alpha > 0$ such that if $S : X \to X$ is distance expanding at Y with the expansion factor λ and for all $y \in Y$ $\rho(S(y), T(y)) \leq \alpha$ then there exists a continuous mapping $h : Y \to X$ such that $Sh|_Y = hT|_Y$, in particular $S(Y') \subset Y'$ for $Y' = h(Y)$, and $\rho(h(z), z) \leq \beta$.

(Hint: apply the previous exercise for $Z = Y, \Gamma = T|_Y, \Phi = \text{id}$, $T = S$ and $Y = Y$. Compare also Ch.5.1.)

Exercise 3.3.1. Prove that every $T : X \to X$ open, distance expanding, for X compact connected, is topologically exact.

Exercise 3.3.2. Prove Lemma 3.3.7' and hence Theorem 3.3.4 (Spectral Decomposition) without the assumption of periodic shadowing, assuming that T is a branched covering of the Riemann sphere.

Exercise 3.3.3. Prove the existence of stable and unstable manifolds for hyperbolic sets and *Smale's Spectral Decomposition Theorem for Axiom A diffeomorphisms.*

An invariant set Λ for a diffeomorphism T is called *hyperbolic* if there exist constants $\lambda > 1$ and $C > 0$ such that the tangent bundle on X restricted to tangent spaces over points in Λ, $T_\Lambda X$ decomposes into DT-invariant subbundles $T_\Lambda X = T^u_\Lambda X \oplus T^s_\Lambda X$ such that $||DT^n(v)|| \geq C\lambda^n$ for all $v \in T^u_\Lambda X$ and $n \geq 0$ and $||DT^n(v)|| \geq C\lambda^n$ for all $v \in T^s_\Lambda X$ and $n \leq 0$.

Prove that for every $x \in \Lambda$ the sets $W^u(x) = \{y \in X : \rho(T^n(x), T^n(y)) \to 0 \text{ as} n \to -\infty\}$, and $W^s(x) = \{y \in X : \rho(T^n(x), T^n(y)) \to 0 \text{ as} n \to \infty\}$ are immersed manifolds. (They are called *unstable* and *stable manifolds.*)

Assume next that a diffeomorphism $T : X \to X$ satisfies *Smale's Axiom A condition*, that is the set of non-wandering points Ω is hyperbolic and $\Omega = \overline{\text{Per}}$.

Then the relation between periodic points is as follows. $x \sim y$ if there are points $z \in W^u(x) \cap W^s(y)$ and $z' \in W^u(y) \cap W^s(x)$ where $W^u(x)$a and $W^s(y)$, and $W^u(y)$a and $W^s(x)$ respectively, intersect transversally, that is the tangent spaces to these manifolds at z and z'

span the whole tangent spaces. Prove that this relation yields Spectral Decomposition, as in Theorem 3.3.4, with topological transitivity assertion rather than topological exactness of course.

As one of the steps prove a lemma corresponding to Lemma 3.3.7' about approximation of a transversal heteroclinic cycle points by periodic ones. That is assume that $x_1, x_2, ..., x_n$ are hyperbolic periodic points (i.e. their orbits are hyperbolic sets) for a diffeomorphism, and $W^u_{x_i}$ has a point p_i of transversal intersection with $W^s_{x_{(i+1) \bmod n}}$ for each $i = 1, ..., n$. Then $p_i \in \overline{\mathrm{Per}}$.

(For the theory of hyperbolic sets for diffeomorphisms see for example [KH].)

Exercise 3.4.1. Prove directly that 1) $\Longrightarrow$ 2) in Proposition 3.4.5, using the specification property, Theorem 3.2.9.

***Exercise 3.5.1.** Suppose $T : X \to X$ is a distance expanding map on a closed surface. Prove that there exist a Markov partition for an iterate T^N compatible with a cell complex structure. That is elements R_i of the partitions are topological discs, the 1-dimensional "skeleton" $\bigcup_i \partial R_i$ is a graph consisting of a finite number of continuous curves "edges" intersecting one another only at end points, called "vertices". Intersection of each two R_i is empty or one vertex or one edge, each vertex is contained in 2 or 3 edges.

(Hint: Start with any cellular partition, with R_i being nice topological discs and correct it by adding or subtracting components of $T^{-N}(R_i), T^{-2N}(R_i)$, etc. See [FJ1] for details.)

***Exercise 3.5.2.** Prove that if T is an expanding map of the 2-dimensional torus $I\!R^2/\boldsymbol{Z}^2$, a factor map of a linear map of $I\!R^2$ given by an integer matrix with two irrational eigenvalues of different moduli (for example $\left(\begin{smallmatrix} 0 & 11 \\ -1 & 7 \end{smallmatrix}\right)$ but not $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right)$), then ∂R_i cannot be differentiable.

(Hint: Smooth curves $T^n(\partial R_i)$ become more and more dense in $I\!R^2/\boldsymbol{Z}^2$ as $n \to \infty$, stretching in the direction of the eigenspace corresponding to the eigenvalue with a larger modulus. So they cannot omit $\mathrm{Int} R_i$.

The same argument, looking backward, says that the components of $T^{-n}(\mathrm{Int} R_i)$ are dense and very distorted, since the eigenvalues have different moduli. The curve ∂R_i must manouver between them, so it is "fractal". See [PU] for more details.)

Historical and Bibliographical Notes.

For Shadowing Lemma in the hyperbolic setting see [Anosov], [Bowen] and [Kushnirenko] or [KH] (for the variant as in Exercise 3.2.1. For the expanding case see [Shub], where structural stability was proved for X a differentiable manifold, T being C^1. D. Sullivan introduced in [Sullivan] the notion *telescope* for the sequence $T^{-1}_{x'_i}(\overline{B(x_{i+1}, \beta)}) \subset$

$\overline{B(x_i,\beta)}$ to capture a shadowing orbit, hence to prove stability of expanding repellers, compare Ch.5.1. This stability was also proved in [Przytycki, 1977]. Recently a comprehensive monography on shadowing by S. Yu. Pilyugin [Pilyugin] appeared.

The existence of Spectral Decomposition in the sense of Theorem 3.3.4 (see Exercise 3.3.3.) was first proved by S. Smale [S] for diffeomorphisms which he called *Axiom A*, that is the set of non-wandering points Ω is hyperbolic and $\Omega = \overline{\mathrm{Per}}$, see also [KH] and further historical informations therein. In a topological setting this was considered by Bowen [B2], called *Axiom A** and for *Axiom A endomorphisms*, covering the diffeomorphisms and expanding (smooth) cases, in [Przytycki, 1977]. For open, distance expanding maps $\Omega = \overline{\mathrm{Per}}$ (Proposition 3.3.b.) corresponds to the analogous fact for Anosov diffeomorphisms. $\Omega = X$ is not known for Anosov diffeomorphisms. It is not true for some distance expanding endomorphisms (Remark 3.3.c), but true for X connected (Exercise 3.3.1), see [Shub] in the smooth case.

The construction of Markov partition in Sec. 5 is similar to the construction for basic sets of Axiom A diffeomorphisms in [Bowen, 1975]. For a general theory of cellular Markov partitions, including Exercise 3.5.1, see [FJ2]. The fact that Hausdorff dimension of the boundaries of 2-dimensional cells is greater than 1, in particular their non-differentiability, Exercise 3.5.2, follows from [PU].

References

[A] D. V. Anosov, On a class of invariant sets of smooth dynamical systems. Proceedings 5th Int. Conf. on Nonlin. Oscill. 2. Kiev 1970, 39-45 (in Russian).

[Bowen, 1975]

[B2] R. Bowen, Markov partitions for Axiom A diffeomorphisms. Amer. J. Math. 92 (1970), 725-749.

[FJ1] F. T. Farrell, L. E. Jones, Markov cell structures for expanding maps in dimension 2. Trans. Amer. Math. Soc. 255 (1979), 315-327.

[FJ2] F. T. Farrell, L. E. Jones, Markov Cell Structures near a Hyperbolic Sets. Memoirs of the AMS 491 (1993).

[KH] A. Katok, B. Hasselblatt, Introduction to the modern theory of dynamical systems. Cambrideg University Press 1995.

[Kushnirenko] A. G. Kushnirenko, Problems in dynamical systems theory on manifolds. IX Mathematical Summer School, Kiev 1972 (in Russian).

[Pilyugin] S.Yu. Pilyugin, Shadowing in Dynamical Systems. L.N.Math. 1706, Springer 1999.

[Przytycki, 1977] F. Przytycki, On Ω-stability and structural stability of endomorphisms satisfying Axiom A. Studia Math. 60 (1977), 61-77.

[PU] F. Przytycki, M. Urbański,

[Shub] M. Shub, Endomorphisms of compact differentiable manifolds. Amer. J. Math. 91 (1969), 175-199.

[Sullivan] D. Sullivan, Seminar on conformal and hyperbolic geometry. Preprint IHES, 1982.

CHAPTER 4. THERMODYNAMICAL FORMALISM

(version Nov.16, 2002)

In Chapter 2 (Th. 2.4.6) we proved that for every positively expansive map of a compact space $T : X \to X$ and an arbitrary continuous function $\phi : X \to I\!R$ there exists an equilibrium state. In Remark 3.4.4 we provided a specific construction for T open distance expanding topologically transitive and ϕ Hölder. Here we shall construct this equilibrium measure with a greater care and study its miraculous regularity with respect to the "potential" function ϕ, its "mixing" properties and uniqueness. So, for the whole chapter $T : X \to X$ we fix an open, distance expanding, topologically transitive map of a compact metric space (X, ρ), with constants η, λ, ξ introduced in Ch.3.

SECTION 4.1. GIBBS MEASURES: INTRODUCTORY REMARKS.

A probability measure μ on X and Borel σ-algebra of sets is said to be a *Gibbs state (measure)* for the potential ϕ if there exist $P \in I\!R$ and $C \geq 1$ such that for all $x \in X$ and all $n \geq 1$

$$C^{-1} \leq \frac{\mu\big(T_x^{-n}(B(T^n(x), \xi))\big)}{\exp(S_n\phi(x) - Pn)} \leq C \tag{4.1.1}$$

If additionally μ is T-invariant, we call μ *invariant Gibbs state* (or *measure*).

We denote the set of all Gibbs states of ϕ by G_ϕ. It is obvious that if μ is a Gibbs state of ϕ and ν is equivalent to μ with Radon-Nikodym derivatives uniformly bounded from above and below, then ν is also a Gibbs state. The following proposition shows that the converse is also true and it identifies the constant P appearing in the definition of Gibbs states as the topological pressure of ϕ.

Proposition 4.1.1. If μ and ν are Gibbs states associated to the map T and a Hölder continuous function ϕ and the corresponding constants are denoted respectively by P, C and Q, D then $P = Q = \mathrm{P}(T, \phi)$ and the measures μ and ν are equivalent with mutual Radon-Nikodym derivatives uniformly bounded.

Proof. Since X is a compact space, there exist finitely many points $x_1, \ldots, x_l \in X$ such that $B(x_1, \xi) \cup \ldots \cup B(x_l, \xi) = X$. We claim that for every compact set $A \subset X$, every $\delta > 0$ and for all $n \geq 1$ large enough

$$\mu(A) \leq CDl \exp((Q - P)n)(\nu(A) + \delta) \tag{4.1.2}$$

By the compactness of A and by the regularity of the measure ν there exists $\varepsilon > 0$ such that $\nu(B(A, \varepsilon)) \leq \nu(A) + \delta$. Fix an integer $n \geq 1$ so large that $\xi\lambda^{-n} < \frac{\varepsilon}{2}$ and for every $1 \leq i \leq l$ let

$$X(i) = \{x \in T^{-n}(x_i) : A \cap T_x^{-n}(B(x_i, \xi)) \neq \emptyset\}.$$

Then

$$A \subset \bigcup_{i=1}^{l} \bigcup_{x\in X(i)} T_x^{-n}(B(x_i,\xi)) \subset B(A,\varepsilon)$$

and since for any fixed $1 \le i \le l$ the sets $T_x^{-n}(B(x_i,\xi))$ for $x \in T^{-n}(x_i)$ are mutually disjoint, it follows from (4.1.1) that

$$\begin{aligned}
\mu(A) &\le \mu\Big(\bigcup_{i=1}^{l} \bigcup_{x\in X(i)} T_x^{-n}(B(x_i,\xi))\Big) \le \sum_{i=1}^{l} \sum_{x\in X(i)} \mu\big(T_x^{-n}(B(x_i,\xi))\big) \\
&\le C \sum_{i=1}^{l} \sum_{x\in X(i)} \exp(S_n\phi(x) - Pn) = C\exp((Q-P)n) \sum_{i=1}^{l} \sum_{x\in X(i)} \exp(S_n\phi(x) - Qn) \\
&\le CD\exp((Q-P)n) \sum_{i=1}^{l} \sum_{x\in X(i)} \nu\big(T_x^{-n}(B(x_i,\xi))\big) \le CD\exp((Q-P)n) l\nu(B(A,\varepsilon)) \\
&\le CDl\exp((Q-P)n)(\nu(A)+\delta)
\end{aligned}$$

Exchanging the roles of μ and ν we also obtain

$$\nu(A) \le CDl\exp((P-Q)n)(\mu(A)+\delta) \tag{4.1.3}$$

for all $n \ge 1$ large enough. So, if $P \ne Q$, say $P < Q$, then it follows from (4.1.3) applied to the compact set X that $\nu(X) = 0$. Hence $P = Q$, and as by regularity of μ and ν, (4.1.2) and (4.1.3) continue to be true for all Borel subsets of X, we conclude that μ and ν are equivalent with the Radon-Nikodym derivative $d\mu/d\nu$ bounded from above by CDl and from below by $(CDl)^{-1}$ (letting $\delta \to 0$).

It is left to show that $P = \mathrm{P}(T,\phi)$. Looking at the expression after the third inequality sign in our estimates of $\mu(A)$ with $A = X$ we get

$$0 = \log\mu(X) \le \log C + \log\Big(\sum_{i=1}^{l} \sum_{x\in X(i)} \exp(S_n\phi(x))\Big) - Pn.$$

Since for every i, $X(i)$ is an (η,n)-separated set, taking into account division by n in the definition of pressure, we can replace here $\sum_i$ by a largest summand for each n. We get $P \le P(T,\phi)$.

On the other hand for an arbitrary $x \in X$

$$\sum_{y\in T^{-n}(x)} \exp(S_n\phi(x) - Pn) \le C \sum_{y\in T^{-n}(x)} \mu\big(T_y^{-n}(B(x,\xi))\big) \le C\mu(X) = C$$

gives $\mathrm{P}(T,\phi) = \mathrm{P}_x(T,\phi) \le P$. The proof is finished. ♣

Remark 4.1.2 To prove Proposition 4.1.1 except the part identifying P as $\mathrm{P}(T,\phi)$ we used only the inequalities

$$C^{-1} \le \frac{\mu\big(T_x^{-n}(B(T^n(x),\xi)\big)\exp Pn}{\nu\big(T_x^{-n}(B(T^n(x),\xi)\big)\exp Qn} \le C.$$

We used the function ϕ in (4.1.1) and its Hölder continuity only to prove that $P = Q = P(T,\phi)$. Hölder continuity allows us also to replace x in $S_n\phi(x)$ by an arbitrary point contained in $T_x^{-n}(B(T^n(x),\xi))$.

Remark 4.1.3. For $\mathcal{R} = \{R_1, ..., R_d\}$, a Markov partition of diameter smaller than ξ, (4.1.1) produces a constant C depending on $\mathcal{R}$ (see Exercise 1) such that

$$C^{-1} \le \frac{\mu\big(R_{j_0,\dots,j_{n-1}}\big)}{\exp(S_n\phi(x) - Pn)} \le C \tag{4.1.4}$$

for every admissible sequence $j_0, j_1, \dots j_{n-1}$ and every $x \in R_{j_0,\dots,j_{n-1}}$. In particular (4.1.4) holds for the shift map of a one-sided topological Markov chain.

The following completes Proposition 4.1.1.

Proposition 4.1.4. If ϕ and ψ are two arbitrary Hölder continuous functions on X, then the following conditions are equivalent:

(1) $\phi - \psi$ is cohomologous to a constant in the space of bounded functions (see Def.1.11.2).

(2) $G_\phi = G_\psi$.

(3) $G_\phi \cap G_\psi \ne \emptyset$.

Proof. Of course (2) implies (3). That (1) implies (2) is also obvious. If (3) is satisfied, that is if there exists $\mu \in G_\phi \cap G_\psi$, then it follows from (4.1.1) that

$$D^{-1} \le \exp(S_n(\phi)(x) - S_n(\psi)(x) - n\mathrm{P}(\phi) + n\mathrm{P}(\psi)) \le D$$

for some constant D, all $x \in X$ and $n \in I\!N$. Applying logarithms we see that the condition (2) in Proposition 3.4.5 is satisfied with ϕ and ψ replaced by $\phi - \mathrm{P}(\phi)$ and $\psi - \mathrm{P}(\psi)$ respectively. Hence, by this Proposition $\phi - P(\phi)$ and $\psi - \mathrm{P}(\psi)$ are cohomologous which finishes the proof. ♣

We shall prove later that the class of Gibbs states associated to T and ϕ is not empty (Sec.3) and contains exactly one Gibbs state which is T-invariant (Corollary 4.2.9). Actually we shall prove a stronger uniqueness theorem. We shall prove that any invariant Gibbs state is an equilibrium state for T and ϕ and prove (Sec.6) uniqueness of the equilibrium state for open expanding T and Hölder continuous ϕ.

Proposition 4.1.5 A probability T-invariant Gibbs state μ is an equilibrium state for T and ϕ.

Proof. Consider an arbitrary finite partition $\mathcal{P}$ into Borel sets of diameter less than $\min(\eta,\xi)$. Then for every $x \in X$ we have $T_x^{-n}(B(T^n(x),\xi)) \supset \mathcal{P}^n(x)$, where $\mathcal{P}^n(x)$ is the element of the partition $\mathcal{P}^n = \bigvee_{j=0}^n \mathcal{P}$ that contains x. Hence $\mu\big(T_x^{-n}(B(T^n(x),\xi))\big) \geq \mu(\mathcal{P}^n(x))$. Therefore by the Shannon-McMillan-Breiman Theorem and (4.1.1) one obtains

$$\mathrm{h}_\mu(T) \geq \mathrm{h}_\mu(T,\mathcal{P}) \geq \int \Big(\limsup_{n\to\infty} \frac{1}{n}(n\,\mathrm{P}(T,\phi)) - S_n\phi(x)\Big)\,d\mu = \mathrm{P}(T,\phi) - \int \phi\,d\mu.$$

or in other words, $\mathrm{h}_\mu(T) + \int \phi\,d\mu \geq \mathrm{P}(T,\phi)$ which just means that μ is an equilibrium state. ♣

SECTION 4.2. TRANSFER OPERATOR AND ITS CONJUGATE. MEASURE WITH PRESCRIBED JACOBIAN.

Suppose first that we are in the situation of Chapter 1., i.e. T is a measurable map. Suppose that m is backward quasi-invariant with respect to T, i.e.

$$T_*(m) = m \circ T^{-1} \prec m. \tag{4.2.1}$$

Then by the Radon-Nikodym Theorem there exists an m-integrable function Φ such that for every measurable set $A \subset X$ we have $m(T^{-1}(A)) = \int_A \Phi dm$. One writes $d(m \circ T^{-1})/dm = \Phi$. In the situation of this Chapter, where T is a local homeomorphism (one does not need expanding yet) if T^{-1} has d branches on a ball $B(x, \xi)$ mapping the ball onto $U_1, ..., U_d$ respectively, then $\Phi = \sum_{j=1}^d \Phi_j$ where $\Phi_j := d(m \circ (T|_{U_j})^{-1})/dm$. If we consider measures absolutely continuous with respect to a backward quasi-invariant "reference measure" m then the transformation $\mu \mapsto T_*(\mu)$ can be rewritten in the language of densities with respect to m as

$$d\mu/dm \mapsto d(T_*\mu)/dm = \sum_{j=1}^{d} \big((d\mu/dm) \circ (T|_{U_j})^{-1}\big)\Phi_j. \tag{4.2.1a}$$

It is comfortable to define $\Psi(z) = \frac{d(m \circ (T|_{U_j})^{-1})}{dm}(T(z))$, i.e. $\Psi = \Phi_j \circ T$ for $z \in U_j$. Notice that Ψ is defined on a set whose T-image has full measure (a set maybe larger than a set of full measure), see Sec.6 for further discussion.

The transformation in (4.2.1a) can be considered as a linear operator $\mathcal{L}_m : L^1(m) \to L^1(m)$,

$$\mathcal{L}_m(u)(x) = \sum_{\overline{x} \in T^{-1}(x)} u(\overline{x})\Psi(\overline{x}).$$

This makes sense, because if we change u on a set A of measure 0, then even if $m(T(A)) > 0$, we have $\Phi_j|_{T(A) \cap B(x,\xi)} = 0$ m-a.e., hence $\mathcal{L}_m(u)$ does not depend on u on $T(A)$.
We have the convention that if u is not defined (on a set of measure 0) and $\Psi = 0$, then $u\Psi = 0$.

Thus we obtain the following characterization of probability T-invariant measures absolutely continuous with respect to m.

Proposition 4.2.0. A probability measure $\mu = hm$, $h \geq 0$, is T-invariant if and only if

$$\mathcal{L}_m(h) = h.$$

After this introduction, the appearence of the following linear operator, called the Perron-Frobenius-Ruelle or Ruelle or Araki or also *transfer operator*, is not surprising:

$$\mathcal{L}_\phi(u)(x) = \sum_{\overline{x} \in T^{-1}(x)} u(\overline{x}) \exp(\phi(\overline{x})).$$

If the function ϕ is fixed we omit sometimes the subscript ϕ at $\mathcal{L}$. The function ϕ is often called a *potential* function.

The transfer's conjugate operator will be our tool to find a quasi-invariant measure m such that Ψ will be a scalar multiple of $\exp\phi$, hence $\mathcal{L}_m$ will be a scalar multiple of $\mathcal{L}_\phi$. Then in turn we will look for fixed points of $\mathcal{L}_m$ to find invariant measures. Restricting our attention to $\exp\phi$, we restrict considerations to Ψ strictly positive defined everywhere. One sometimes allows ϕ to have the value $-\infty$, but we do not consider this case in our book.

Let us be now more precise. Consider $\mathcal{L}_\phi$ acting on the Banach space of continuous functions $\mathcal{L}_\phi : C(X) \to C(X)$. It is a continuous linear operator and its norm is equal to $\sup_x \sum_{\overline{x}\in T^{-1}(x)} \exp(\phi(\overline{x})) = \sup \mathcal{L}_\phi(1\!\!1)$ as this is a positive operator i.e. it maps real non-negative functions to real non-negative functions (see Ch.2.1). Consider the conjugate operator $\mathcal{L}_\phi^* : C^*(X) \to C^*(X)$. Note that as conjugate to a positive operator it is also positive, i.e. transforms measures into measures.

Lemma 4.2.1. For every $\mu \in C^*(X)$ and every Borel set $A \subset X$ on which T is injective

$$\mathcal{L}_\phi^*(\mu)(A) = \int_{T(A)} \exp(\phi \circ (T|_A)^{-1}) d\mu \tag{4.2.2}$$

Proof. It is sufficient to prove (4.2.2) for $A \subset B(x,r)$ with any $x \in X$ and r such that T is injective on $B(x,2r)$ (say $r = \eta$). Approximate in pointwise convergence the indicator function χ_A by uniformly bounded continuous functions with support in $B = B(x,2r)$. We have for any such function f

$$\mathcal{L}_\phi^*(\mu)(f) = \mu(\mathcal{L}_\phi(f)) = \int_{T(B)} (f \exp(\phi)) \circ (T|_B)^{-1} d\mu.$$

We used here the fact that the only branch of T^{-1} mapping $T(B)$ to the support of f is that one leading $T(B)$ to B. Passing with f to the limit χ_A on both sides (Lebesgue convergence theorem) gives (4.2.2). ♣

Observe that whereas $\mathcal{L}_\phi$ transports measure from the past, $\mathcal{L}_\phi^*$ pulls it back from the future with Jacobian $\exp\phi$. This is the right operator to use, to look for the missing "reference measure" m.

Definition 4.2.2. Recall from Chapter 1 (Def.1.9.4) that a measurable function $J : X \to [0,\infty)$ is called the *Jacobian* or the*strong Jacobian* of a map $T : X \to X$ with respect to a measure μ if for every Borel set $A \subset X$ on which T is injective $\mu(T(A)) = \int_A J d\mu$. In particular μ is forward quasi-invariant

J is called the *weak Jacobian* if $J : X \to [0,\infty)$ and there exists a Borel set $E \subset X$ such that $\mu(E) = 0$ and for every Borel set $A \subset X$ on which T is injective, $\mu(T(A \setminus E)) = \int_A J d\mu$.

Notice that if μ is backward quasi-invariant then the condition that J is the weak Jacobian translates to $\mu(A) = \int_{T(A)} \frac{1}{J\circ(T|_A)^{-1}} d\mu$.

Corollary 4.2.3. If a probability measure μ satisfies $\mathcal{L}_\phi^*(\mu) = c\mu$ (i.e. μ is an eigenmeasure of $\mathcal{L}_\phi^*$ corresponding to a positive eigenvalue c), then $c\exp(-\phi)$ is the Jacobian of T with respect to μ.

Proof. Substitute $c\mu$ in place of $\mathcal{L}^*(\mu)$ in (4.2.2). It then follows that μ is backward quasi-invariant and $c\exp(-\phi)$ is the weak Jacobian of T with respect to μ. Since $\frac{1}{\exp(-\phi)} = \exp\phi$, it is positive everywhere, hence $c\exp(-\phi)$ is the strong Jacobian of T. ♣

Theorem 4.2.4. Let $T : X \to X$ be a local homeomorphism of a compact metric space X and let $\phi : X \to I\!\!R$ be continuous. Then there exists a probability measure $m = m_\phi$ and a constant $c > 0$, such that $\mathcal{L}_\phi^*(m) = cm$. The function $c\exp(-\phi)$ is the strong Jacobian for T with respect to the measure m.

Proof. Consider the map $l(\mu) := \frac{\mathcal{L}^*(\mu)}{\mathcal{L}^*(\mu)(1\!\!1)}$ on the convex set of probability measures on X, i.e. on $M(X)$, endowed with the weak* topology (Ch.2.1). The transformation l is continuous in this topology since $\mu_n \to \mu$ weak* implies for every $u \in C(X)$ that $\mathcal{L}^*(\mu_n)(u) = \mu_n(\mathcal{L}(u)) \to \mu(\mathcal{L}(u)) = \mathcal{L}^*(\mu)(u)$. As $M(X)$ is weak* compact (see Th.2.1.6) we can use Theorem 2.1.7 (Schauder-Tychonoff fixed point theorem) to find $m \in M(X)$ such that $l(m) = m$. Hence $\mathcal{L}^*(m) = cm$ for $c = \mathcal{L}^*(m)(1\!\!1)$. Thus T has the Jacobian equal to $c\exp(-\phi)$, by Corollary 4.2.3. ♣

Note again that we write $\exp\phi$ in order to guarantee it never vanishes, so that there exists the Jacobian for T with respect to m. To find an eigen-measure m for $\mathcal{L}^*$ (i.e. with a weak Jacobian being a multiple of $\exp(-\phi)$) we could perfectly allow $\exp\phi = 0$.

We have the following complementary fact in case Jacobian J exists.

Proposition 4.2.4a. If $T : X \to X$ is a local homeomorphism of a compact metric space X and a Jacobian J with respect to a probability measure m exists, then for every Borel set A

$$\frac{1}{d}\int_A J\,dm \le m(T(A)) \le \int_A J\,dm.$$

where d is the degree of T ($d := \sup_{x\in X} \sharp T^{-1}(\{x\})$). In particular if $m(A) = 0$, then $m(T(A)) = 0$.

Proof. Let us partition A into finitely many Borel sets, say $A_1, A_2, \dots, A_n$, of diameters so small that T restricted to each of them is injective. Then, on one hand,

$$m(T(A)) = m\big(\bigcup_{i=1}^n T(A_i)\big) \le \sum_{i=1}^n m(T(A_i)) = \sum_{i=1}^n \int_{A_i} J\,dm = \int_A J\,dm.$$

and on the other hand, since the multiplicity of the family $\{T(A_i) : 1 \le i \le n\}$ does not exceed d,

$$m(T(A)) = m\big(\bigcup_{i=1}^n T(A_i)\big) \ge \frac{1}{d}\sum_{i=1}^n m(T(A_i)) = \frac{1}{d}\sum_{i=1}^n \int_{A_i} J\,dm = \frac{1}{d}\int_A J\,dm.$$

The proof is finished. ♣

Let us go back to T, an open distance expanding topologically transitive map.

Proposition 4.2.5. The measure m is positive on non-empty open sets. Moreover for every $r > 0$ there exists $\alpha = \alpha(r) > 0$ such that for every $x \in X$, $m(B(x,r)) \geq a$.

Proof. For every open $U \subset X$ there exists $n \geq 0$ such that $\bigcup_{j=0}^n T^j(U) = X$ (Theorem 3.3.9). So, by Proposition 4.2.4a, $m(U) = 0$ would imply that $1 = m(X) \leq \sum_{j=0}^n m(T^j(U)) = 0$, a contradiction.

Finally let $x_1, ..., x_m$ be an $r/2$-net in X and $\alpha := \min_{1 \leq j \leq m}\{m(B(x_j, r/2))\}$. Since for every $x \in X$ there exists j such that $\rho(x, x_j) \leq r/2$, hence $m(B(x,r)) \supset B(x_j, r/2)$. Thus it is enough to set $\alpha(r) := \alpha$. ♣

Proposition 4.2.6. The measure m is a Gibbs state of ϕ and $\log c = \mathrm{P}(T, \phi)$.

Proof. We have for every $x \in X$ and every integer $n \geq 0$,

$$m(B(x,\xi)) = \int_{T_x^{-n}(B(T^n(x),\xi))} c^n \exp(-S_n\phi)\, dm.$$

Since, by Lemma 3.4.2, the ratio of the supremum and infimum of the integrand of the above integral is bounded from above by a constant $C > 0$ and from below by C^{-1}, we obtain

$$1 \geq m(B(x,\xi)) \geq C^{-1} c^n \exp(-S_n\phi(x)) m(T_x^{-n}(B(T^n(x),\xi)))$$

and

$$\alpha(\xi) \leq m(B(x,\xi)) \leq C c^n \exp(-S_n\phi(x)) m(T_x^{-n}(B(T^n(x),\xi))).$$

Hence

$$\alpha(\xi) C^{-1} \leq \frac{m(T_x^{-n}(B(T^n(x),\xi)))}{\exp(S_n\phi(x) - n\log c)} \leq C$$

and therefore m is a Gibbs state. That $\log c = \mathrm{P}(T,\phi)$ follows now from Proposition 4.1.1. ♣

We now also give a simple direct proof of equality $\log c = \mathrm{P}(T,\phi)$. First note that by the definition of $\mathcal{L}_\phi$ and a simple inductive argument, for every integer $n \geq 0$

$$\mathcal{L}_\phi^n(u)(x) = \sum_{\overline{x} \in T^{-n}(x)} u(\overline{x}) \exp(S_n\phi(\overline{x})). \tag{4.2.2a}$$

The estimate (3.4.3) translates to

$$C^{-1} \leq \mathcal{L}^n(1\!\!1)(x)/\mathcal{L}^n(1\!\!1)(y) \leq C \quad \text{for every} \quad x, y \in X. \tag{4.2.3}$$

Now $c^n = c^n m(1\!\!1) = (\mathcal{L}^*)^n(m)(1\!\!1) = m(\mathcal{L}^n(1\!\!1))$ and hence

$$\log c = \lim_{n\to\infty} \frac{1}{n} \log m(\mathcal{L}^n(1\!\!1)) = P(T,\phi).$$

The latter equality follows from (4.2.3) and Proposition 3.4.3.

Note that in the latter equality we used the property that m is a measure (positive). For m a signed eigen-measure and c a complex eigenvalue for $\mathcal{L}^*$ we would obtain only $\log|c| \le P(T,\phi)$ (one should consider a function u such that $\sup|u| = 1$ and $m(u) = 1$ rather than the function $1\!\!1$) and indeed usually the point spectrum of $\mathcal{L}^*$ is big(ref ??????).

We are in the position to prove already some ergodic properties of Gibbs states:
Theorem 4.2.7. If T is topologically exact, then the system (T, m) is exact in the measure theoretic sense, namely for every A of positive measure $m(T^n(A)) \to 1$ as $n \to \infty$, see 1.10.3.

The topological counterpart of this Theorem is the fact that topological mixing implies topological exactness, Th.3.3.10.

Proof. Let E be an arbitrary Borel set with $m(E) > 0$. By the regularity of m we can find a compact set $A \subset E$ such that $m(A) > 0$. Fix an arbitrary $\varepsilon > 0$. As in the proof of Proposition 4.1.1, we find for every n large enough, a covering of A by sets D_ν of the form $T_x^{-n}(B(x_i,\xi)), x \in X(i), i = 1,...,l$ such that $m(\bigcup_\nu D_\nu) \le m(A)+\varepsilon$. Hence $m(\bigcup_\nu(D_\nu \setminus A)) \le \varepsilon$. Since the multiplicity of this covering is at most l, we have

$$\sum_\nu m(D_\nu \setminus A) \le l\varepsilon.$$

Hence

$$\frac{\sum_\nu m(D_\nu \setminus A)}{\sum_\nu m(D_\nu)} \le \frac{l\varepsilon}{m(A)}.$$

Therefore for all n large enough there exists $D = D_\nu = T_x^{-n}(B)$, for some $B = B(x_i,\xi))$, $1 \le i \le l$, such that

$$\frac{m(D \setminus A)}{m(D)} \le \frac{l\varepsilon}{m(A)}.$$

Hence

$$\frac{m(B \setminus T^n(A))}{m(B)} \le \frac{\int_{D\setminus A} c^n \exp(-S_n\phi)dm}{\int_D c^n \exp(-S_n\phi)dm} \le C\frac{m(D \setminus A)}{m(D)} \le C\frac{l\varepsilon}{m(A)}$$

with C as in Proof of Proposition 4.2.6. By the topological exactness of T, there exists $N \ge 0$ such that for every j we have $T^N(B(x_j,\xi)) = X$. In particular $T^N(B) = X$. So, using Proposition 4.2.4a, we get

$$m(X \setminus T^N(T^n(A))) \le m(T^N(B \setminus T^n(A))) \le c^N(\inf\exp\phi)^{-N}\frac{Cl\varepsilon}{m(A)}.$$

Letting $\varepsilon \to 0$ we obtain $m(X \setminus T^N(T^n(A))) \to 0$ as $n \to \infty$. Hence $m(T^{N+n}(A)) \to 1$. ♣

We have considered here a special Gibbs measure $m = m_\phi$. Notice however that by Proposition 4.1.1 the assertion of Theorem 4.2.7 holds for every Gibbs measure associated to T and ϕ.

Corollary 4.2.8. If T is a topologically transitive, open, distance expanding map, then for every Hölder potential ϕ, each corresponding Gibbs measure is ergodic.

Proof. By Th.3.3.4 and Th.3.3.9 there exists a positive integer N such that T^N is topologically mixing on a T^N-invariant closed-open set $Y \subset X$, where $\bigcup_{j=0,\dots,N-1} T^j(Y) = X$. So our $T^N|_Y$, being also an open expanding map, is exact in the measure-theoretic sense by Theorem 3.7. So if $m(E) > 0$ then for every $j = 0, ..., N-1$ we have $m(T^{Nn}T^j(E)) \to m(T^j(Y))$, hence $m(\bigcup_{n\geq 0} T^n(E)) \to 1$. For E being T-invariant this yields $m(E) = 1$. This implies ergodicity. ♣

With the use of Proposition 1.2.4 we get the following fact promised in Section 4.1.

Corollary 4.2.9. If T is a topologically transitive, open, distance expanding map, then for every Hölder potential ϕ, there is at most one corresponding invariant Gibbs measure.

SECTION 4.3. ITERATION OF TRANSFER OPERATOR. EXISTENCE OF GIBBS STATES.

It is comfortable to consider the operator $\mathcal{L}_{\overline{\phi}}$ for $\overline{\phi} = \phi - \mathrm{P}(T,\phi)$. That is $\mathcal{L}_{\overline{\phi}} = \mathrm{e}^{-\mathrm{P}(T,\phi)}\mathcal{L}_\phi$. (Recall that $\mathrm{P}(T,\phi) = \log c$.) Then for the reference measure $m = m_\phi$ satisfying $\mathcal{L}^*_\phi(m) = \mathrm{e}^{\mathrm{P}(\phi)}m$ we have $\mathcal{L}^*_{\overline{\phi}}(m) = m$ i.e.

$$\int u dm = \int \mathcal{L}_{\overline{\phi}}(u) dm \quad \text{for every} \quad u \in C(X). \tag{4.3.1}$$

For fixed ϕ we often denote $\mathcal{L}_{\overline{\phi}}$ by $\mathcal{L}_0$. By (4.2.3) for every $x, y \in X$, and non-negative integer n

$$\mathcal{L}_0^n(1\!\!1)(x)/\mathcal{L}_0^n(1\!\!1)(y) \le C. \tag{4.3.2}$$

Multiplying this inequality by $\mathcal{L}_0^n(1\!\!1)(y)$ and then integrating with respect to the variable x and y we get respectively the first and the third of the following inequalities below

$$C^{-1} \le \inf \mathcal{L}_0^n(1\!\!1) \le \sup \mathcal{L}_0^n(1\!\!1) \le C. \tag{4.3.3}$$

By (3.4.2) for every $x, y \in X$ such that $x \in B(y,\xi)$ we have an inequality more refined than (3.4.3). Namely
(4.3.4)

$$\frac{\mathcal{L}_\phi^n(1\!\!1)(x)}{\mathcal{L}_\phi^n(1\!\!1)(y)} = \frac{\sum_{\overline{x}\in T^{-n}(x)} \exp S_n\phi(\overline{x})}{\sum_{\overline{y}\in T^{-n}(y)} \exp S_n\phi(\overline{y})} \le \sup_{\overline{x}\in T^{-n}(x)} \frac{\exp S_n\phi(\overline{x})}{\exp S_n\phi(y_n(\overline{x}))} \le \exp(C_1\rho(x,y)^\alpha),$$

where $C_1 = \frac{\vartheta_\alpha(\phi)}{1-\lambda^{-\alpha}}$ and $y_n(\overline{x}) := T_{\overline{x}}^{-n}(y)$. By this estimate and by (4.3.3) we get for all $n \ge 1$ and all $x, y \in X$ such that $x \in B(y,\xi)$

$$\begin{aligned} \Big|\mathcal{L}_0^n(1\!\!1)(x) - \mathcal{L}_0^n(1\!\!1)(y)\Big| &= \Big|\frac{\mathcal{L}_0^n(1\!\!1)(x)}{\mathcal{L}_0^n(1\!\!1)(y)} - 1\Big|\mathcal{L}_0^n(1\!\!1)(y) \le \\ C|\exp(C_1\rho(x,y)^\alpha) - 1| &\le C_2\rho(x,y)^\alpha \end{aligned} \tag{4.3.6}$$

with C_2 depending on C, C_1 and ξ.

Proposition 4.3.1. There exists a positive function $u_\phi \in \mathcal{H}_\alpha(X)$ such that $\mathcal{L}_0(u_\phi) = u_\phi$ and $\int u_\phi \, dm = 1$.

Proof. By (4.3.6) and (4.3.3) the functions $\mathcal{L}_0^n(1\!\!1)$ have uniformly bounded norms in Hölder space $\mathcal{H}_\alpha(X)$, see Ch.3.4. Hence by Arzela-Ascoli theorem there exists a limit $u_\phi \in C(X)$ for a subsequence of $u_n = \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{L}_0^j(1\!\!1)$, $n = 1, \ldots$. Of course $u_\phi \in \mathcal{H}_\alpha(X)$, $C^{-1} \le u_\phi \le C$, and using (4.3.3), a straightforward computation shows that $\mathcal{L}_0(u_\phi) = u_\phi$ (compare 2.1.14). Also $\int u_\phi \, dm = \lim_{n\to\infty} \int u_n \, dm = \int 1\!\!1 \, dm = 1$. The proof is finished. ♣

Combining this proposition, Proposition 4.2.0, Proposition 4.2.6 and Corollary 4.2.9 we get the following.

Theorem 4.3.2. For every Hölder continuous function $\phi : X \to I\!R$ there exists a unique invariant Gibbs state associated to T and ϕ, namely $\mu_\phi = u_\phi m_\phi$.

In the rest of this Section we provide a detailed study of iteration of $\mathcal{L}_0$ on the real or complex Banach spaces $C(X)$ and $\mathcal{H}_\alpha$.

Definition 4.3.3. We call a continuous linear operator $Q : B \to B$ on a Banach space B *almost periodic* if for every $b \in B$ the sequence $Q^n(b), n = 0, 1, ...$ is relatively compact, i.e. its closure in B is compact (in the norm topology).

Proposition 4.3.4. The operators $\mathcal{L}_0^n$ on $C(X)$ have uniformly bounded norms for all $n = 1, 2, ...$

Proof. By the definition of $\mathcal{L}$ and by (4.3.3) for every $u \in C(X)$:

$$\sup |\mathcal{L}_0^n(u)| \le \sup |u| \sup \mathcal{L}_0^n(1\!\!1) \le C \sup |u| \tag{4.3.8}$$

♣

Remark that instead of referring to the form of $\mathcal{L}$ one can only refer to the fact that $\mathcal{L}$ is a positive operator, hence its norm is attained on $1\!\!1$.

Theorem 4.3.5. The operator $\mathcal{L}_0$ is almost periodic on $C(X)$. Moreover, all the functions $\mathcal{L}_0^n(u)$ are equicontinuous and have uniformly bounded absolute values, provided $\mathcal{L}_0$'s are associated to ϕ belonging to a bounded set in $\mathcal{H}_\alpha$ and u taken from a family of equicontinuous functions, of uniformly bounded absolute values.

Proof. For every $x \in X$ and $n > 0$ denote $\exp(S_n\overline{\phi}(x))$ by $E_n(x)$. Consider arbitrary points $x \in X, y \in B(x, \xi)$. Use the notation $y_n(\overline{x}) := T_{\overline{x}}^{-n}(y)$, the same as in (4.3.4). We have for every $u \in C(X)$

$$|\mathcal{L}_0^n(u)(x) - \mathcal{L}_0^n(u)(y)| = \Big| \sum_{\overline{x} \in T^{-n}(x)} u(\overline{x})E_n(\overline{x}) - u(y_n(\overline{x}))E_n(y_n(\overline{x})\Big|$$

$$\le \Big| \sum_{\overline{x} \in T^{-n}(x)} u(\overline{x})(E_n(\overline{x}) - E_n(y_n(\overline{x})))\Big| + \Big| \sum_{\overline{x} \in T^{-n}(x)} E_n(y_n(\overline{x}))(u(\overline{x}) - u(y_n(\overline{x}))\Big|$$

$$\le (\sup |u|)C_2\rho(x, y)^\alpha + C \sup_{\overline{x} \in T^{-n}(x)} |u(\overline{x}) - u(y_n(\overline{x}))| \tag{4.3.9}$$

by (4.3.6) and (4.3.3). Denote a modulus of uniform continuity of u by h, i.e. consider an increasing function $h : I\!R^+ \to I\!R^+$ such that $\lim_{\varepsilon \searrow 0} h(\varepsilon) = 0$ and for every $z_1, z_2 \in X$ $|u(z_1) - u(z_2)| \le h(\rho(z_1, z_2))$. (4.3.9) gives:

$$|\mathcal{L}_0^n(u)(x) - \mathcal{L}_0^n(u)(y)| \le (\sup|u|)C_2\rho(x,y)^\alpha + Ch(\rho(x,y)) := g(\rho(x,y)) \tag{4.3.10}$$

We conclude that all functions $\mathcal{L}_0^n(u)$ have the same modulus of continuity g, depending on h, $\sup|\phi|$ and $||\phi||_{\mathcal{H}_\alpha}$. They are also uniformly bounded by Proposition 4.3.4. Hence by Arzela-Ascoli theorem the sequence $\mathcal{L}_0^n(u)$ is relatively compact.

If we consider a family of functions u rather than one function, we set h a modulus of continuity of the family.

♣

For $u \in \mathcal{H}_\alpha$ we obtain the fundamental estimate (4.3.11):

Theorem 4.3.6. There exist constants $C_3, C_4 > 0$ such that for every $u \in \mathcal{H}_\alpha$ all $n = 1, 2, ...$and $\lambda > 1$ from the expanding property of t

$$\vartheta_\alpha(\mathcal{L}_0^n(u)) \le C_3\lambda^{-n\alpha}\vartheta_\alpha(u) + C_4\|u\|_\infty, \tag{4.3.11}$$

Proof. Continuing the third line of (4.3.9) we obtain

$$|\mathcal{L}_0^n(u)(x) - \mathcal{L}_0^n(u)(y)| \le \|u\|_\infty C_2\rho(x,y)^\alpha + C\vartheta_{\alpha,\xi}(u)\lambda^{-n\alpha}\rho(x,y)^\alpha$$

We have applied here the inequality $\rho(\overline{x}, y_n(\overline{x})) \le \lambda^{-n}\rho(\overline{x}, \overline{y})$

This proves (4.3.11), provisionally with $\vartheta_{\alpha,\xi}$ rather than ϑ_α, with $C_3 = C$ from (3.4.3) and (4.3.3) and with $C_4 = C_2$ (recall that the latter constant is of order CC_1 where C_1 appeared in (4.3.4)). Passing to ϑ_α changes C_4 to $\max\{C_4, 2C/\xi^\alpha\}$, see (3.3.8) and Ch.3.4. ♣

Corollary 4.3.7. There exist an integer $N > 0, \tau < 1, C_5 > 0$ such that for every $u \in \mathcal{H}_\alpha$

$$\|\mathcal{L}_0^N(u)\|_{\mathcal{H}_\alpha} \le \tau\|u\|_{\mathcal{H}_\alpha} + C_5\|u\|_\infty \tag{4.3.12}$$

Proof. This Corollary immediately follows from (4.3.11) and Proposition 4.3.4. ♣

In fact (4.3.12) together with (4.3.8) imply a similar fact for iterates of $\mathcal{L}^N$, which resembles back (3.3.11). Namely the following holds

Proposition 4.3.8.

$$\exists C_6 > 0 \ \forall n = 1, 2, ... \quad \|\mathcal{L}_0^{nN}(u)\|_{\mathcal{H}_\alpha} \le \tau^n\|(u)\|_{\mathcal{H}_\alpha} + C_6\|u\|_\infty \tag{4.3.13}$$

Proof. Substitute in (4.3.12) $\mathcal{L}_0^N(u)$ in place of u etc. n times using $\|\mathcal{L}_0^j(u)\|_\infty \le C\|u\|_\infty$. You obtain (4.1.13) with $C_6 = CC_5/(1-\tau)$. ♣

In Appendix we prove a general theorem by Ionescu-Tulcea and Marinescu (abbr.: ITM), which under assumptions (4.3.8), (4.3.12) gives an information about the spectrum of $\mathcal{L}_0$. Ch.3 Sec.5 is devoted to this. This occurs useful in other than expanding and Hölder cases. Here, in the next Section, assuming topological mixing of T, we shall proceed directly, not referring to ITM Theorem.

Analogously to $\mathcal{L}_0^n$ considered on $C(X)$ the convergence theorem below is a special case of a general theory of almost periodic operators, see Sec.5.

SECTION 4. CONVERGENCE OF $\mathcal{L}^n$. MIXING PROPERTIES OF GIBBS MEASURES

Recall that by Proposition 4.3.1 there exists a positive function $u_\phi \in \mathcal{H}_\alpha(X)$ such that $\mathcal{L}_0(u_\phi) = u_\phi$.

It is convienient to replace the operator $\mathcal{L}_0$ by $\hat{\mathcal{L}}(u) = \frac{1}{u_\phi}\mathcal{L}_0(uu_\phi)$.

If we denote the operator of multiplication by a function w by the same symbol w then we can write $\hat{\mathcal{L}}(u) = u_\phi^{-1} \circ \mathcal{L}_0 \circ u_\phi$. Since $\hat{\mathcal{L}}$ and $\mathcal{L}_0 = \mathcal{L}_{\overline{\phi}}$ are conjugate by the operator u_ϕ, their spectra are the same. In addition, as this operator is positive, non-negative functions go to non-negative functions. Hence measures are mapped to measures by the conjugate operator.

Proposition 4.4.2. $\hat{\mathcal{L}} = \mathcal{L}_\psi$ where $\psi = \overline{\phi} + \log u_\phi - \log u_\phi \circ T = \phi - \mathrm{P}(T, \phi) + \log u_\phi - \log u_\phi \circ T$.

Proof. $\hat{\mathcal{L}}(u)(x) = \frac{1}{u_\phi(x)} \sum_{T(\overline{x})=x} u(\overline{x}) u_\phi(\overline{x}) \exp \phi(\overline{x}) = \sum_{T(\overline{x})=x} u(\overline{x}) \exp(\phi(\overline{x})$
$+ \log u_\phi(\overline{x}) - \log u_\phi(x))$ ♣

Note that the eigenfunction u_ϕ for $\mathcal{L}_0$ has changed to the eigenfunction $1\!\!1$ for $\hat{\mathcal{L}}$. In other words we have the following.

Proposition 4.4.3. $\hat{\mathcal{L}}(1\!\!1) = 1\!\!1$, i.e. for every $x \in X$

$$\sum_{\overline{x} \in T^{-1}(x)} \exp \psi(\overline{x}) = 1 \tag{4.4.1}$$

♣

Note that Jacobian of T with respect to the Gibbs measure $\mu = u_\phi m$ (see Th. 4.3.2) is $(u_\phi \circ T)(\exp(-\overline{\phi}))u_\phi^{-1} = \exp(-\psi)$. So for ψ the reference measure (with Jacobian $\exp(-\psi)$) and the invariant Gibbs measure coincide.

Note that passing from $\mathcal{L}_\phi$, through $\mathcal{L}_{\overline{\phi}}$, to $\mathcal{L}_\psi$ we have been replacing ϕ by cohomological (up to a constant) functions. By Proposition 4.1.3. this does not change the set of Gibbs states.

One can think of the transformation $u \mapsto u/u_\phi$ as new coordinates on $C(X)$ or $\mathcal{H}_\alpha(X)$ (real or complex-valued functions). $\mathcal{L}_0$ changes in these coordinates to $\mathcal{L}_\psi$ and the functional $m(u)$ to $m(u_\phi u)$. The latter, denote it by m_ψ, is the eigenmeasure for $\mathcal{L}_\psi^*$ with the eigenvalue 1. It is positive because the operator u_ϕ is positive (see the comment above). So $\exp(-\psi)$ is the Jacobian for m_ψ by Corollary 4.2.3. Hence by (4.4.1) m_ψ is T-invariant. This is our invariant Gibbs measure μ.

Proposition 4.3.4 applied to $\hat{\mathcal{L}}$ takes the form.

Proposition 4.4.4. $\|\hat{\mathcal{L}}\|_\infty = 1$.

Proof. $\sup|\hat{\mathcal{L}}(u)| \le \sup|u|$ because $\hat{\mathcal{L}}$ is an operator of "taking an average" of u from the past (by Proposition 4.4.3). The equality follows from $\hat{\mathcal{L}}(1\!\!1) = 1\!\!1$. ♣

The topological exactness of T gives a stronger result:

Lemma 4.4.5. If T is topologically exact, then given any increasing function $g : I\!\!R^+ \to I\!\!R^+$ such that $\lim_{\varepsilon\searrow 0} g(\varepsilon) = 0$, $(\forall\delta_1 > 0$ and $K > 0)$ $(\exists\delta_2 = \delta_2(g, \delta_1, K) > 0$ and a positive integer $n = n(g, \delta_1, K) > 0)$ such that for all $\phi \in \mathcal{H}_\alpha$ with $||\phi||_{\mathcal{H}_\alpha} \le K$ and $u \in C(X, I\!\!R)$ with modulus of continuity g) (i.e. for every $z_1, z_2 \in X$ $|u(z_1) - u(z_2)| \le g(\rho(z_1, z_2))$, and such that $\int u d\mu = 0$ and $\sup|u| \ge \delta_1$, we have

$$\sup|\hat{\mathcal{L}}^n(u)| \le \sup|u| - \delta_2.$$

Proof. Fix $\varepsilon > 0$ so small that $g(\varepsilon) < \delta_1/2$. Let n be ascribed to ε according to Proposition 3.3.10, namely $(\forall x)T^n(B(x, \varepsilon)) = X)$. Since $\int u d\mu = 0$, there exist $y_1, y_2 \in X$ such that $u(y_1) \le 0$ and $u(y_2) \ge 0$. For an arbitrary $x \in X$ choose $x' \in B(y_1, \varepsilon) \cap T^{-n}(x)$ (it exists by the definition of n). We have $u(x') \le \delta_1/2$. So

$$\hat{\mathcal{L}}^n(u)(x) = u(x') \exp S_n\psi(x') + \sum_{\overline{x}\in T^{-n}(x)\setminus\{x'\}} u(\overline{x}) \exp S_n\psi(\overline{x})$$

$$\le (\sup|u| - \delta_1/2) \exp S_n\psi(x') + \sup|u| \sum_{\overline{x}\in T^{-n}(x)\setminus\{x'\}} \exp S_n\psi(\overline{x})$$

$$\le \sup|u|\Big(\sum_{\overline{x}\in T^{-n}(x)} \exp S_n\psi(\overline{x})\Big) - (\delta_1/2) \exp S_n\psi(x') = \sup|u| - (\delta_1/2) \exp S_n\psi(x').$$

Similarly for $x' \in B(y_2, \varepsilon) \cap T^{-n}(x)$

$$\hat{\mathcal{L}}^n(u)(x) \ge -\sup|u| + (\delta_1/2) \exp S_n\psi(x').$$

Thus we proved our Lemma, with $\delta_2 := (\delta_1/2) \inf_{x\in X} \exp S_n\psi(x)$.

Note that we used here the existence of a uniform bound, $\sup|\psi| \le \sup|\phi| + 2\sup|\log(u_\phi)|$ and $\sup|\log(u_\phi)| \le \log C$, where C depends on K, see (4.2.3), (3.4.2), (3.4.3). ♣

We shall prove now a theorem which completes Proposition 4.3.4 and Theorem 4.3.5.

Theorem 4.4.6. For every $u \in C(X, \mathbb{C})$ and T, topologically exact open expanding map, we have

$$c^{-n}\mathcal{L}^n_\phi(u) - m_\phi(u)u_\phi \to 0 \quad \text{(coverges uniformly)} \quad \text{as} \quad n \to \infty \tag{4.4.2}$$

In particular if $\int u d\mu = 0$ then

$$\hat{\mathcal{L}}^n(u) \to 0 \tag{4.4.3}$$

Moreover the convergences in (4.4.2) and (4.4.3) are uniform in every set of equicontinuous functions u of uniformly bounded absolute values, and ϕ in a bounded set in $\mathcal{H}_\alpha(X)$.

Proof. For real-valued u, with $\int u d\mu = 0$, the sequence $a_n(u) := \sup |\hat{\mathcal{L}}^n(u)|$ is decreasing by Proposition 4.4.4. Suppose that $\lim_{n\to\infty} a_n = a > 0$. By Theorem 4.3.5 all the iterates $\hat{\mathcal{L}}^n(u)$ have a common modulus of continuity g. So applying Lemma 4.4.5 with this g and $\delta_1 = a$ we find n_0, δ_2 such that $\sup |\hat{\mathcal{L}}^{n_0}(\hat{\mathcal{L}}^n(u))| \leq \sup |\hat{\mathcal{L}}^n(u)| - \delta_2$ for every n. So for n such that $\sup |\hat{\mathcal{L}}^n(u)| < a + \delta_2$ we obtain $\sup |\hat{\mathcal{L}}^{n+n_0}(u)| < a$, a contradiction with the definition of a.

This proves (4.4.3). For an arbitrary $u \in C(X, I\!\!R)$ we obtain from (4.4.3) due to $\hat{\mathcal{L}}(1\!\!1) = 1\!\!1$

$$\hat{\mathcal{L}}^n(u) - \mu(u)1\!\!1 = \hat{\mathcal{L}}^n(u - \mu(u)1\!\!1) \to 0.$$

Change now coordinates on $C(X)$ to go back to $\mathcal{L}_0$ and next replace it by $c^{-1}\mathcal{L}_\phi$. One obtains (4.4.2). Given a complex-valued u decompose it into sum of real and imaginary part.

If we allow u and ϕ to vary we modify the proof. The point is that by Lemma 4.4.5, for every $\delta_1 > 0$, for every $m \geq \sup |u| n(g, \delta_1, K)/\delta_2(g, \delta_1, K)$, we get in $\sup |u|/\delta_2(g, \delta_1, K)$ steps, $\sup |\hat{\mathcal{L}}^m(u)| \leq \delta_1$, where g is the modulus of continuity for the family $\{\hat{\mathcal{L}}^n(u)\}$ provided by Theorem 4.3.5, and K bounds the norm in $\mathcal{H}_\alpha$ of the functions ϕ. Letting $\delta_1 \to 0$ proves the Theorem. ♣

Note that (4.4.2) means weak*-convergence of measures

$$\lim_{n\to\infty} \sum_{\overline{x} \in T^{-n}(x)} \delta_{\overline{x}} \times \exp(S_n\phi(\overline{x}))/c^n \to u_\phi(x) m_\phi$$

for every $x \in X$. Using (4.4.2) also for $u = 1\!\!1$ we obtain

$$\lim_{n\to\infty} \sum_{\overline{x} \in T^{-n}(x)} \delta_{\overline{x}} \times \exp(S_n\phi(\overline{x}))/\mathcal{L}_\phi^n(1\!\!1)(x) \to m_\phi \tag{4.4.3'}$$

In the sequel one can consider either $C(X, I\!\!R)$ or $C(X, \mathbb{C})$. Let us decide for $C(X, \mathbb{C})$.

Note that by $\mathcal{L}_\phi^*(m_\phi) = c m_\phi$ we have the $\mathcal{L}$-invariant decomposition

$$C(X) = \text{span}(u_\phi) \oplus \ker(m_\phi).$$

For $u \in \mathrm{span}(u_\phi)$ we have $\mathcal{L}_\phi(u) = cu$. On $\ker(m_\phi)$, by Th.4.4.6., $c^{-n}\mathcal{L}_\phi^n \to 0$ in strong topology. Denote $(\mathcal{L}_\phi)|_{\ker(m_\phi)}$ by $\mathcal{L}_{\ker,\phi}$. For $\mathcal{L}_{\ker,\phi}$ restricted to $\mathcal{H}_\alpha$ we can say more on the above convergence:

Theorem 4.4.7. There exists an integer $n > 0$ such that

$$\|c^{-n}\mathcal{L}^n_{\ker,\phi}\|_{\mathcal{H}_\alpha} < 1.$$

Proof. Again it is sufficient to consider real u with $\mu(u) = 0$ and the operator $\hat{\mathcal{L}}$. Set $\delta = \min\{1/8C_4, 1/4\}$, with C_4 from (4.3.11). By Th.4.3.6. for u such that $\|u\|_{\mathcal{H}_\alpha} \le 1$ all functions $\hat{\mathcal{L}}^n(u)$ have the same modulus of continuity $g(\varepsilon) = C_7\varepsilon^\alpha$ for $C_7 = C_3 + C_4 > 0$. Hence from Theorem 4.4.6. we conclude that $(\exists n_1)(\forall n \ge n_1)(\forall u : \|u\|_{\mathcal{H}_\alpha} \le 1)$

$$\|\hat{\mathcal{L}}^n(u)\|_\infty \le \delta. \tag{4.4.4}$$

Next, for n_2 satisfying $C_3\lambda^{-n_2\alpha}C_7 + C_4\delta \le 1/4$, again by Th.4.3.6. we obtain

$$\vartheta_\alpha(\hat{\mathcal{L}}^{n_2}(\hat{\mathcal{L}}^{n_1}(u)) \le 1/4.$$

Hence $||\hat{\mathcal{L}}^{n_1+n_2}(u)||_{\mathcal{H}_\alpha} \le 1/2$. Theorem has been proved with $n = n_1 + n_2$. ♣

Note that Theorem 4.4.6. could be deduced from Theorem 4.4.7 by approximation of continuous functions uniformly by Hölder ones, and using Proposition 4.3.4.

Corollary 4.4.8. The convergences in Theorem 4.4.6 for $u \in \mathcal{H}_\alpha$ are exponential. Namely there exist $0 < \tau < 1$ and $C \ge 0$ such that for every function $u \in \mathcal{H}_\alpha$

$$\|c^{-n}\mathcal{L}^n_\phi(u) - m_\phi(u)u_\phi\|_\infty \le \|c^{-n}\mathcal{L}^n_\phi(u) - m_\phi(u)u_\phi\|_{\mathcal{H}_\alpha} \le C\|u - m_\phi(u)u_\phi\|_{\mathcal{H}_\alpha}\tau^n. \tag{4.4.4'}$$

In particular if $\int u d\mu = 0$ then

$$\|\hat{\mathcal{L}}^n(u)\|_\infty \le \|\hat{\mathcal{L}}^n(u)\|_{\mathcal{H}_\alpha} \le C\|u\|_{\mathcal{H}_\alpha}\tau^n. \tag{4.4.4''}$$

Now we can study "mixing" properties of the system (T, μ) for our invariant Gibbs measure μ. Roughly speaking the speed of mixing is related to the speed of convergence of $\mathcal{L}^n_{\ker,\phi}$ to 0.

The first dynamical (mixing) consequence of Theorem 4.4.8 is the following result known as the exponential decay of correlations, see the definition in Ch.1.11.

Theorem 4.4.10. There exists $C \ge 1$ and $\rho < 1$ such that for all $f \in \mathcal{H}_\alpha,\ g \in L^1(\mu)$

$$C_n(f, g) \le C\rho^n\|f - Ef\|_{\mathcal{H}_\alpha}\|g - Eg\|_1.$$

Proof. Write $F = f - Ef, G = g - Eg$. We obtain

$$C_n(f,g) = \Big|\int F \cdot (G \circ T^n)\, d\mu\Big| = \Big|\int \hat{\mathcal{L}}^n\big(F \cdot (G \circ T^n)\big)\, d\mu\Big|$$

$$= \Big|\int G \cdot \hat{\mathcal{L}}^n(F)\, d\mu\Big| \le \|G\|_1 C\tau^n \|F\|_{\mathcal{H}_\alpha}.$$

We have used here a very important identity true for arbitrary F, G ($EF, EG = 0$ did not matter), that

$$\mathcal{L}^n\big(F \cdot (G \circ T^n)\big) = G \cdot \mathcal{L}^n(F), \tag{4.4.5}$$

which follows immediately from the definition of $\mathcal{L}$ with an arbitrary potential ϕ. Namely

$$\mathcal{L}_\phi{}^n\big(F \cdot (G \circ T^n)\big)(x) = \sum_{\overline{x} \in T^{-n}(x)} G(x) F(\overline{x}) \exp S_n\psi(\overline{x}) = G(x)\mathcal{L}_\phi^n(F)(x)$$

♣

Exercise. Prove that for every μ square integrable functions f, g one has $\int f \cdot (g \circ T^n)\, d\mu \to Ef \cdot Eg$. (Hint: approximate f and g by Hölder functions. Of course the information on the speed of convergence would become lost.)

The convergence in the exercise is one of equivalent definitions of the *mixing* property, see Ch.1.10. We proved however earlier the stronger property: measure-theoretical exactness, Th. 4.2.7.

We can however make a better use of the exponential convergence in Theorem 4.4.10.

Theorem 4.4.11. Let (X,T) be a topologically mixing topological one-sided Markov chain with T the shift to the left and $d \ge 2$ symbols, see Ch.0. Let $\mathcal{F}$ be the σ-algebra generated by the partition $\mathcal{A}$ into sets with fixed 0-th coordinate, namely by $\mathcal{A} = \{X_1, ..., X_d\}$ where $X_j = \{(a_0, a_1, ...) \in X : a_0 = j\}$. For every $0 \le k \le l$ write $\mathcal{F}_k^l$ for the σ-algebra generated by $\mathcal{A}_k^l = \{\bigvee_{j=k}^l T^{-j}(\mathcal{A})$ i.e. by the sets with fixed $k, k+1, ..., l$'th coordinates. Let $\phi : X \to I\!R$ be Hölder continuous.

Then there exist $0 < \rho < 1, C > 0$ such that for every $k \ge 0, f : X \to I\!R$ measurable in $\mathcal{F}_0^k$ and g being μ_ϕ-integrable

$$|\int f \cdot (g \circ T^n)\, d\mu_\phi - Ef \cdot Eg| \le C\rho^{n-k} \|f - Ef\|_1 \|g - Eg\|_1. \tag{4.4.6}$$

Proof. Assume $Ef = Eg = 0$. By Theorem 4.4.10

$$|\int f \cdot (g \circ T^n)\, d\mu| = |\int g \cdot \hat{\mathcal{L}}^{n-k}(\hat{\mathcal{L}}^k(f))\, d\mu| \le \|g\|_1 C\rho^{n-k} \|\hat{\mathcal{L}}^k(f)\|_{\mathcal{H}_\alpha}. \tag{4.4.7}$$

Decompose f into real and imaginary parts and represent each one by the difference of nowhere negative functions. This allows in the estimates which follow to assume that $f \geq 0$.

Notice that for every cylinder $A \in \mathcal{A}$ and $x \in A$, in the expression

$$\hat{\mathcal{L}}^k(f)(x) = \sum_{T^k(y)=x} f(y) \exp S_k\psi(y)$$

there is no dependence of $f(y)$ on $x \in A$ because f is constant on cylinders of $\mathcal{A}_0^k$. So

$$\frac{\sup_A \hat{\mathcal{L}}^k(f)}{\inf_A \hat{\mathcal{L}}^k(f} \leq \sup_{B \in \mathcal{A}_0^k} \sup_{y,y' \in B} \exp\big(S_k\psi(y) - S_k\psi(y')\big) \leq C$$

a constant C resulting from Ch.3.4. So

$$\sup_A \hat{\mathcal{L}}^k(f) \leq \frac{C}{\mu(A)} \int \hat{\mathcal{L}}^k(f)\, d\mu = \frac{C}{\mu(A)} \|f\|_1 \leq \big(\frac{C}{\inf_{A \in \mathcal{A}} \mu(A)}\big) \|f\|_1 = C'\|f\|_1,$$

where the latter equality defines C'.

It is left yet to estimate the $\vartheta_{\alpha,\xi}$ and ϑ_α pseudonorms of $\hat{\mathcal{L}}^k(f)$, cf.Ch.3.4. We assume that ξ is less than the minimal distance between the cylinders in $\mathcal{A}$. We have similarly to (4.3.6), for x, y belonging to the same cylinder $A \in \mathcal{A}$,

$$\Big|\hat{\mathcal{L}}^k(f)(x) - \hat{\mathcal{L}}^k(f)(y)\Big| = \Big|\big(\frac{\hat{\mathcal{L}}^k(f)(x)}{\hat{\mathcal{L}}^k(f)(y)} - 1\big)\Big| |\hat{\mathcal{L}}^k(f)(y)|$$

$$\leq (\exp C_1\rho(x,y)^\alpha - 1)\|C'\|f\|_1 \leq C''\rho(x,y)^\alpha\|f\|_1.$$

for a constant C''.

Hence, $\vartheta_{\alpha,\xi}(\hat{\mathcal{L}}^k(f)) \leq \|f\|_1 C''$ and, passing to ϑ_α as in Ch.3.4,

$$\vartheta_\alpha(\hat{\mathcal{L}}^k(f)) \leq \|f\|_1 \max\{C'', 2C'\xi^{-\alpha}\}).$$

Thus, continuing (4.4.7), we obtain for a constant C

$$C_n(f,g) \leq \|f\|_1\|g\|_1 C\rho^{n-k}.$$

♣

An immediate corollary from Theorem 4.4.11 is that for every $B_1 \in \mathcal{F}_0^k$ and Borel B_2 (i.e. $B_2 \in \mathcal{F}_0^\infty$)

$$|\mu(B_1 \cap T^{-n}(B_2)) - \mu(B_1)\mu(B_2)| \leq C\rho^{n-k}\mu(B_1)\mu(B_2), \tag{4.4.8}$$

compare (1.11.9). Therefore, for every non-negative integer t and every $A \in \mathcal{F}_0^k$ for the conditional measures with respect to A

$$\sum_{B \in \mathcal{A}_0^t} |\mu(T^{-n}(B)|A) - \mu(B)| \le C\rho^{n-k}.$$

This means that $\mathcal{A}$ satisfies the weak Bernoulli property, hence the natural extension $(\tilde{X}, \tilde{T}, \tilde{\mu})$ is measure-theoretically isomorphic to a two-sided Bernoulli shift, see Ch.1.11.

Corollary 4.4.12. Every topologically exact, open, distance expanding map T, with invariant Gibbs measure $\mu = \mu_\phi$ for a Hölder continuous function ϕ, has the natural extension $(\tilde{X}, \tilde{T}, \tilde{\mu})$ measure-theoretically isomorphic to a two-sided Bernoulli shift.

Proof. Let $\pi : \Sigma_A \to X$ be the coding map from a one-sided topological Markov chain, due to a Markov partition, see Ch.3.5. Since π is Hölder, the function $\phi \circ \pi$ is also Hölder continuous, hence we can discuss the invariant Gibbs measure $\mu_{\phi\circ\pi}$. For this measure we can apply Theorem 4.4.11 and its consequences. Recall also that by Theorem 3.5.5 π yields a measure-theoretical isomorphism between $\mu_{\phi\circ\pi}$ and $\mu_{\phi\circ\pi} \circ \pi^{-1}$, Therefore to end the proof it is enough to prove the following.

Lemma 4.4.13. The measures μ_ϕ and $\mu_{\phi\circ\pi} \circ \pi^{-1}$ coincide.

Proof. The function $\exp(-\phi \circ \pi + P - h)$ for $h := \log u_{\phi\circ\pi} + \log u_{\phi\circ\pi} \circ \sigma)$, is the strong Jacobian for the shift map σ and the measure $\mu_{\phi\circ\pi}$, where P is the pressure for both $(\sigma, \phi \circ \pi)$ and (T, ϕ), see Theorem 3.5.4. Since π yields a measure-theoretical isomorphism between $\mu_{\phi\circ\pi}$ and $\mu_{\phi\circ\pi} \circ \pi^{-1}$, the measure $\mu_{\phi\circ\pi} \circ \pi^{-1}$ is forward quasi-invariant under T and has the strong Jacobian $\exp(-\phi + P - h \circ \pi^{-1})$. The same up to a bounded function factor is the Jacobian of μ_ϕ. Therefore both measures are equivalent, hence as ergodic they coincide. ♣

Chapter 3, Section 5: More on almost periodic operators

version of June 9, 1997

In this Section we show how to deduce Theorem 4.4.6 (on convergence) and Theorem 4.4.7 and Corollary 4.4.8 (exponential convergence) from general functional analysis theorems. We do not need this later on in this book, but the theorems are useful in other important situations .....

Recall (Def.4.3.3) that $Q : F \to F$ a continuous linear operator of a Banach space is called almost periodic if for every $b \in F$ the sequence $Q^n(b)$ is relatively compact. By Banach-Steinhaus theorem there is a constant $C \geq 0$ such that $\|Q^n\| \leq C$ for every $n \geq 0$.

Theorem 4.5.1. If $Q : F \to F$ is an almost periodic operator on a complex Banach space F, then

$$F = F_0 \oplus F_{\mathrm{u}}, \tag{4.5.1}$$

where $F_0 = \{x \in F : \lim_{n\to\infty} A^n(x) = 0\}$ and F_{u} is the closure of the subspace of F generated by eigenfunctions of eigenvalues of modulus 1. Adding additional assumptions one gains additional information on this decomposition.

Definition 4.5.2. Let $F = C(X)$ and suppose $Q : F \to F$ is *positive*, namely $f \geq 0$ implies $Q(f) \geq 0$. Then Q is called *primitive* if for every $f \in C(X), f \geq 0, f \not\equiv 0$ there exists $n \geq 0$ such that for every $x \in X$ it holds $Q^n(f)(x) > 0$. If we change the order of the quantificators to: ... for every x there exists n ..., then we call Q *nondecomposable.*

Theorem 4.5.3 For $Q : C(X) \to C(X)$ (real or complex) linear positive primitive operator of spectral radius equal to 1 we have dim span$(C(X)_u = 1$ in the decomposition (4.5.1), the eigenvalue corresponding to $C(X)_u$ is equal to 1 and the eigenfunction is positive (everywhere > 0). More precisely there exists a probability measure m_Q on X and a positive function u_Q such that for every $u \in C(X)$ we have strong convergence

$$Q^n(u) \to u_Q \int u\, dm.$$

Proof. This is just a repetition of considerations of Sections 2-4. First find a probability measure m such that $Q^*(m) = m$ as in Th.4.2.4. (we leave a proof that the eigenvalue is equal to 1, to the reader). Next find for Q an eigenfunction $u_Q \geq 0$ as $\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} Q^j(1\!\!1)$. We have $u_Q = Q(u_Q) > 0$ because Q is nondecomposable. Finally for $\hat{Q}(u) := Q(uu_Q)u_Q^{-1}$ we have $\hat{Q}(1\!\!1) = 1\!\!1$ (a positive operator with this property is called *stochastic*) and we repeat Proof of Th. 4.4.6, replacing the property of topological exactness by primitivity. ♣

Notice that this yields Theorem 4.4.6 because of

Proposition 4.5.4. If an open expanding map T is topologically exact then for every continuous function ϕ the transfer operator $Q = \mathcal{L}_{\overline{\phi}}$ is primitive.

The proof is easy, it is in fact contained in Proof of Lemma 4.4.5.

Assume now only that T is topologically transitive. Let Ω^k denote the sets from spectral decomposition $X = \Omega = \bigcup_{k=1}^{n} \Omega^k$ as in Th.3.3.4. Write $u_Q \in C(X)$ for an eigenfunction of the operator Q as before. Notice now (exercise!) that the space F_{u} for the operator $Q = \mathcal{L}_{\overline{\phi}}$ is spanned by n eigenfunctions $v_t = \sum_{k=1}^{n} \chi_{\Omega^k} \lambda^{-tk} u_Q, \quad t = 1, ..., n$, where χ means indicator functions, with $\lambda = \varepsilon^{2\pi i/n}$. Each v_t corresponds to the eigenvalue λ^t. Thus the set of these eigenvalues is a cyclic group.

It is also an easy exercise to describe F_{u} if $X = \Omega = \bigcup_{j=1}^{J} \bigcup_{k=1}^{k(j)} \Omega_j^k$. The set of eigenvalues is the union of J cyclic groups. It is harder to understand F_{u} and the corresponding set of eigenvalues for T open expanding, without assuming $\Omega = X$.

References to the above theory are:

[LL] M. Yu. Lyubich, Yu. I. Lyubich: Perron-Frobenius theory for almost periodic operators and semigroups representations. Teoria Funkcii 46 (1986), 54-72.

[L] M. Yu. Lyubich: Entropy properties of rational endomorphisms of the Riemann sphere. ETDS (1983), 351-385.

A general theorem related to Theorem 4.4.7 and Corollary 4.4.8 is the following.

Theorem 4.5.5 (Ionescu–Tulcea and Marinescu) Let $(F, |\cdot|)$ be a Banach space equipped with a norm $|\cdot|$ and let $E \subset F$ be its linear subspace. Moreover the linear space E is assumed to be equipped with a norm $\|\cdot\|$ which satisfies the following two conditions.

(1) Any bounded subset of the Banach space E with the norm $\|\cdot\|$ is relatively compact as a subset of the Banach space F with the norm $|\cdot|$.

(2) If $\{x_n : n = 1, 2, \ldots\}$ is a sequence of points in E such that $\|x_n\| \le K_1$ for all $n \ge 1$ and some constant K_1, and if $\lim_{n\to\infty} |x_n - x| = 0$ for some $x \in F$, then $x \in E$ and $\|x\| \le K_1$.

Let $Q : F \to F$ be a bounded linear operator which preserves E, whose restriction to E is also bounded with respect to the norm $\|\cdot\|$ and which satisfies the following two conditions.

(3) There exists a constant K such that $|Q^n| \le K$ for all $n = 1, 2, \ldots$.

(4) $\exists N \ge 1 \quad \exists \tau < 1 \quad \exists K_2 > 0 \quad \|Q^N(x)\| \le \tau\|x\| + K_2|x|$ for all $x \in E$.

Then

(5) There exists at most finitely many eigenvalues of $Q : F \to F$ of modulus 1, say $\gamma_1, \ldots, \gamma_p$.

(6) Let $F_i = \{x \in F : Q(x) = \gamma_i x\}$, $i = 1, \ldots, p$. Then $F_i \subset E$ and $\dim(F_i) < \infty$.

(7) The operator $Q : F \to F$ can be represented as

$$Q = \sum_{i=1}^{p} \gamma_i Q_i + S$$

where Q_i and S are bounded, $Q_i(F) = F_i$, $\sup_{n\geq 1} |S^n| < \infty$ and

$$Q_i^2 = Q_i, \quad Q_i Q_j = 0 \ (i \neq j), \quad Q_i S = S Q_i = 0$$

Moreover

(8) $S(E) \subset E$ and $S|_E$ considered as a linear operator on $(E, \|\cdot\|)$, is bounded and there exist constants $K_3 > 0$ and $0 < \tilde{\tau} < 1$ such that

$$\|S^n|_E\| \leq K_3 \tilde{\tau}^n$$

for all $n \geq 1$.

The proof of this theorem can be found in [...] in the case $N = 1$ (see assumpion 4). Its validity for any $N \geq 1$ is mentiond in Section 9, p.145 of this paper. In Appendix ... we give a complete proof.

Now, in view of Theorem 3.4.1 and Corollary 4.3.7, Theorem 4.5.5 applies to the operator $Q = \mathcal{L}_{\overline{\phi}} : C(X) \to C(X)$ if one substitutes $F = C(X)$, $E = \mathcal{H}_\alpha(X)$. If T is topologically exact and in concequence Q is primitive on $C(X)$, then $\dim(\oplus F_i) = 1$ and the corresponding eigenvalue is equal to 1, as in Theorem 4.5.3.

Example of application Lasota-Yorke, Rychlik: functions of bounded variation ....

Sec. 4.6. UNIQUENESS OF EQUILIBRIUM STATES

We proved already the existence (Th.4.3.2) and uniqueness (Cor.4.2.9) of invariant Gibbs states and proved that invariant Gibbs states are equilibrium states (Prop.4.1.5). Here we shall give 3 different proofs of the uniqueness of equilibrium states.

Let ν be a T-invariant measure and let a finite real function J_ν be the corresponding Jacobian in the weak sense, J_ν is defined ν-a.e. . By the invariance of ν we have $\nu(E) = 0 \Rightarrow \nu(T^{-1}(E)) = \nu(E) = 0$, i.e. ν is backward quasi-invariant. At the beginning of Sec.2 we defined in this situation $\Psi = \Phi_x \circ T$ with $\Phi_x = \frac{d\nu \circ T_x^{-1}}{d\nu}$ defined for ν-a.e. point in the domain of a branch T_x^{-1}. (In Sec.2 we used notation Φ_j for Φ_x.) Φ_x is strong Jacobian for T_x^{-1}.

Notice that for ν-a.e. z

$$(J_\nu \circ T_x^{-1}) \cdot \Phi_x(z) = \begin{cases} 1, & \text{if } \Phi_x(z) \neq 0; \\ 0, & \text{if } \Phi_x(z) = 0. \end{cases} \tag{4.6.1}$$

Indeed, after removal of $\{z : \Phi_x(z) = 0\}$ the measures ν and $\nu \circ T^{-1}$ are equivalent, hence Jacobians of T and T_x^{-1} are mutual inverses. We can fix J_ν arbitrary, bounded, on $T^-(\{z : \Phi_x(z) = 0\})$.

Recall that we have defined $\mathcal{L}_\nu : L^1(\nu) \to L^1(\nu)$, the transfer operator associated with the measure ν as follows

$$\mathcal{L}_\nu(g)(x) = \sum_{y \in T^{-1}(x)} g(y)\Psi(y).$$

Remind that if T maps a set A of measure 0 to a set of positive measure, then Ψ is specified, equal to 0, on a subset of A that is mapped by T to a set of full measure ν in $T(A)$.

Then since ν is T-invariant, $\mathcal{L}_\nu(1\!\!1) = 1\!\!1$ and for every ν-integrable g we have $\int \mathcal{L}_\nu(g)\, d\nu = \int g\, d\nu$.

Lemma 4.6.1. Let $\psi : X \to I\!\!R$ be a continuous function such that $\mathcal{L}_\psi(1\!\!1) = 1\!\!1$, i.e. for every x, $\sum_{y \in T^{-1}(x)} \exp \psi(y) = 1$, and let ν be an ergodic equilibrium state for ψ. Then J_ν is strong Jacobian and $J_\nu = \exp(-\psi)$ ν-almost everywhere .

Proof. The proof is based on the following computation using the inequality $1 + \log(x) \le x$, with the equality only for $x = 1$.

$$1 = \int 1\!\!1\, d\nu \ge \int \mathcal{L}_\nu(J_\nu \exp \psi)\, d\nu = \int J_\nu \exp \psi\, d\nu$$

$$\ge \int \big(1 + \log(J_\nu \exp \psi)\big)\, d\nu = 1 + \int \psi d\nu + \int \log J_\nu\, d\nu = 1 + \int \psi d\nu + h_\nu(T) \ge 1.$$

To obtain the first inequality, write $\mathcal{L}_\nu(J_\nu \exp \psi)(x) = \sum_{y \in T^{-1}(x)} J_\nu(y)(\exp \psi(y))\Psi(y)$ which is equal to 1 if $(\forall y \in T^{-1}(x))\Psi(y) > 0$ or < 1 otherwise, by (4.6.1) and by $\sum_{y \in T^{-1}(x)} \exp \psi(y) = 1$.

The last inequality follows from

$$\int \psi d\nu + h_\nu(T) = P(\psi) \geq \limsup_{n\to\infty} \frac{1}{n} \log \sum_{y\in T^{-n}(x)} \exp S_n\psi(y) = 0,$$

true, see Th.2.2.10, since all points in $T^{-n}(x)$ are (n,η)-separated, η defined in Ch.3.

Therefore all the inequalities in this proof must become equalities. Therefore the Jacobian $\Phi_x \neq 0$ for each branch T_x^{-1} and $J_\nu = \exp(-\psi)$, ν- a.e. ♣

Notice that we have not assumed ψ is Hölder above. Now we shall assume Hölder.

Theorem 4.6.2. There exists exactly one equilibrium state for each Hölder continuous potential ϕ.

Proof. Let ν be an equilibrium state for ϕ. As in Sec.4 set $\psi = \phi - P(T,\phi) + \log u_\phi \circ T - \log u_\phi$ and ν is also equilibrium state for ψ. Then by Lemma 4.6.1

$$\nu\big(T_z^{-n}(B(T^n(z),\xi))\big) = \int_{B(T^n(z),\xi)} \exp\, \big(S_n\psi(T_z^{-n}(x))\big)\, d\nu(x) =$$

$$\int_{B(T^n(z),\xi)} \frac{u_\phi(x)}{u_\phi(T^n(x))} \exp(S_n\phi - nP(T,\phi))(T_z^{-n}(x))\, d\nu(x).$$

So, by *pre-bounded distortion lemma* (Lemma 3.4.2),

$$\frac{\inf|u_\phi|}{\sup|u_\phi|} BC^{-1} \leq \frac{\nu\big(T_z^{-n}(B(T^n(z),\xi))\big)}{\exp(S_n\phi - nP(T,\phi))(z)} \leq \frac{\sup|u_\phi|}{\inf|u_\phi|} C,$$

where $B = \inf\{\nu(B(y,\xi)\}$. It is positive by Proposition 4.2.5.

Therefore ν is an invariant Gibbs state for ϕ; unique by Corollary 4.2.9. ♣

Remark 4.6.3. In fact already the knowledge that $\exp(-\psi)$ is weak Jacobian implies automatically that it is a strong Jacobian. Indeed by the invariance of ν we have

$$\sum_{y\in T^{-1}(x)} \Phi_y(y) = 1 = \sum_{y\in T^{-1}(x)} \exp\psi(y)$$

and each non-zero summand on the left is equal to a corresponding summand on the right. So there are no summands equal to 0.

Uniqueness: Proof II. We shall provide a new proof of Lemma 4.6.1. It is not so elementary as the previous one, but it exhibits a relation with the finite case, the *prototype lemma* in Introduction.

For every $y \in X$ denote $A(y) := T^{-1}\Big(T\big(\{y\}\big)\Big)$. Let $\{\nu_A\}$ denote the canonical system of conditional measures for the partition of X into the sets $A = A(y)$, see Ch.1.6. Since there exists a finite one-sided generator, see Lemma 2.4.5, with the use of Theorem 1.9.7 we obtain

$$0 = P(T,\, \psi) = h_\nu(T) + \int \psi \, d\nu = H_\nu\Big(\varepsilon \mid T^{-1}(\varepsilon)\Big) + \int \psi \, d\nu =$$
$$= \int \Big(\sum_{z \in A(y)} \nu_{A(y)}(\{z\})\big(-\log\big(\nu_{A(y)}(\{z\})\big) + \psi(z)\big)\Big) \, d\nu(y).$$

The latter expression is always negative except for the case $\nu_{A(y)}(z) = \exp \psi(z)$ ν-a.e. by the prototype lemma. So for a set $Y = T^{-1}\big(T(Y)\big)$ of full measure ν, for every $y \in Y$ we have

$$\nu_{A(y)}(\{y\}) = \exp \psi(y), \quad \text{in particular} \quad \nu_{A(y)}(\{y\}) \neq 0. \tag{4.6.2}$$

So for every Borel set $B \subset Y$ such that T is 1–to–1 on it, since B intersects each $A(y) \subset T^{-1}(T(B))$ at precisely one point, we obtain

$$\nu\big(T(B)\big) = \nu\Big(T^{-1}\big(T(B)\big)\Big) =$$
$$= \int_{T^{-1}\left(T(B)\right)} \Big(\int_{A(y)} 1\!\!1_B(z)/\nu_{A(y)}(\{z\}) \, d\nu_{A(y)}(z) \Big) \, d\nu(y) =$$
$$= \int_{T^{-1}\left(T(B)\right)} 1\!\!1_B(y)/\nu_{A(y)}(\{y\}) \, d\nu(y) = \int_B 1/\nu_{A(y)}(\{y\}) \, d\nu(y)$$

Notice that we have proved in this computation a general useful fact that $1/\nu_{A(y)}(\{y\}))$ is weak Jacobian for T and ν. In absence of the property (4.6.2) that $\nu_{A(y)}(\{y\}) \neq 0$ we should have subtracted the set $E = \{y : \nu_{A(y)}(y) = 0\}$ of measure 0 under the integrals.

Let us go back to our situation. By (4.6.2) this Jacobian is equal to $\exp -\psi$. Observe also that $\nu\big(T(X \setminus Y)\big) = 0$ because $X \setminus Y = T^{-1}\big(T(X \setminus Y)\big)$ and ν is T-invariant. So $\exp -\psi$ is strong Jacobian.

Uniqueness. Proof III. Due to Corollary 2.5.7 it is sufficient to prove the differentiability of the pressure $P(T, \phi)$ as a function of continuous function ϕ at Hölder ϕ in a set of directions dense in the weak topology on $C(X)$.

Lemma 4.6.4. Let $\phi : X \to I\!R$ be a Hölder continuous function and μ_ϕ denote the invariant Gibbs measure. Let $F : X \to I\!R$ be continuous. Then, for an arbitrary $x \in X$,

$$\lim_{n \to \infty} \frac{1}{n} \frac{\sum_{y \in T^{-n}(x)} S_n F \exp(S_n \phi)(y)}{\sum_{y \in T^{-n}(x)} \exp(S_n \phi)(y)} = \int F \, d\mu_\phi. \tag{4.6.3}$$

The convergence is uniform for an equicontinuous family of F's and ϕ's in a bounded set in the Banach space of Hölder functions $\mathrm{H}_\alpha(X)$.

Proof.

The above left hand side expression can be written in the form:

$$\lim_{n\to\infty} \frac{\frac{1}{n}\sum_{j=0}^{n-1} \mathcal{L}_\phi^n(F\circ T^j)(x)}{\mathcal{L}_\phi^n(1\!\!1)(x)} = \lim_{n\to\infty} \frac{\frac{1}{n}\sum_{j=0}^{n-1} \mathcal{L}^{n-j}(F\cdot \mathcal{L}^j(1\!\!1))(x)}{\mathcal{L}^n(1\!\!1)(x)}. \tag{4.6.4}$$

where $\mathcal{L} = \mathcal{L}_0 = e^{-P(T,\phi)}\mathcal{L}_\phi$, compare the beginning of Sec.3.
Since $F\cdot \mathcal{L}^j(1\!\!1)$ is an equicontinuous family of functions we obtain

$$\mathcal{L}^{n-j}(F\cdot \mathcal{L}^j(1\!\!1))(x) \to u_\phi(x)\int F\cdot \mathcal{L}^j(1\!\!1)\, dm_\phi$$

as $n-j\to\infty$, see Remark 4.4.6a.

Therefore continuing (4.6.4) we obtain

$$\lim_{n\to\infty} \frac{\frac{1}{n}\sum_{j=0}^{n-1} u_\phi(x)\int F\cdot \mathcal{L}^j(1\!\!1)\, dm_\phi}{u_\phi(x)} = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1}\int F\cdot \mathcal{L}^j(1\!\!1)\, dm_\phi = \int F\, d\mu_\phi$$

since $\mathcal{L}^j(1\!\!1)$ uniformly converges to u_ϕ and $\mu_\phi = u_\phi m_\phi$. ♣

Now we shall calculate the derivative $d\mathrm{P}(T,\phi+t\gamma)/dt$ for every Hölder ϕ and γ at every t. In particular, this will give differentiability at $t=0$. Thus our dense set of directions is spanned by Hölder functions γ.

Theorem 4.6.5. We have

$$\frac{d}{dt}\mathrm{P}(T,\phi+t\gamma) = \int \gamma\, d\mu_{\phi+t\gamma}$$

for all $t\in I\!\!R$.

Proof. Write

$$P_n(t) = \frac{1}{n}\log \sum_{y\in T^{-n}(x)} \exp(S_n(\phi+t\gamma))(y);$$

$$Q_n(t) := (dP_n/dt)(t) = \frac{\frac{1}{n}\sum_{y\in T^{-n}(x)} S_n\gamma(y)\exp(S_n(\phi+t\gamma))(y)}{\sum_{y\in T^{-n}(x)}\exp(S_n(\phi+t\gamma)(y)}. \tag{4.6.5}$$

By Lemma 4.6.4 $\lim_{n\to\infty} Q_n(t) = \int\gamma\, d\mu_{\phi+t\gamma}$ and the convergence is uniform with respect to t. Since, in addition, $\lim_{n\to\infty} P_n(t) = \mathrm{P}(t)$, we conclude that $P(T,\phi+t\gamma) = \lim_{n\to\infty} P_n(t)$ is differentiable and the derivative is equal to the limit of derivatives: $\lim_{n\to\infty} Q_n(t) = \int\gamma\, d\mu_{\phi+t\gamma}$, ♣

Notice that the differential (Gateaux) operator $\gamma\mapsto\int\gamma\, d\mu_\phi$ is indeed that one from Proposition 2.5.6. Notice also that *a posteriori*, by Cor.2.5.7, we proved that for ϕ Hölder continuous, $P(T,\phi)$ is differentiable in direction of every continuous function. This is by the way obvious in general: two different supporting functionals are different restricted to any dense subspace.

§4.7. Probability laws and $\sigma^2(u,v)$.

Exponential convergences in §4.4 allow to prove the probability laws.

Theorem 4.7.1. Let $T : X \to X$ be an open distance expanding topologically exact map and μ the invariant Gibbs measure for a Hölder function $\phi : X \to I\!R$. Then if $g : X \to I\!R$ satisfies

$$\sum_{n=0}^{\infty} \|\hat{\mathcal{L}}^n(g - \mu(g))\|_2 < \infty, \tag{4.7.1}$$

in particular if g is Hölder continuous, it satisfies CLT. If g is Hölder continuous it satisfies LIL.

Proof. First show how CLT can be deduced from Theorem 1.11.5. We can assume $\mu(g) = 0$. Let $(\tilde{X}\tilde{\mathcal{F}}, \tilde{\mu})$ be the natural extension (see Ch.1.7). Recall that $\tilde{X}$ can be viewed as the set of all T-trajectories $(x_n)_{n\in\boldsymbol{Z}}$ (or backward trajectories), $\tilde{T}((x_n)) = (x_{n+1})$ and $\pi_n((x_n)) = x_n$. It is sufficient now to check (1.11.12) for the automorphism $\tilde{T}$ the function $\tilde{g} = g \circ \pi_0$ and $\tilde{\mathcal{F}}_0 = \pi^{-1}(\mathcal{B})$ for the completed Borel σ-algebra $\mathcal{B}$. Since $\tilde{g}$ is measurable with respect to $\tilde{\mathcal{F}}_0$ it is also measurable with respect to all $\tilde{\mathcal{F}}_n = \tilde{T}^{-n}(\tilde{\mathcal{F}}_0)$ for $n \le 0$ hence $\tilde{g} = E(\tilde{g}|\tilde{\mathcal{F}}_n)$. So we need only to prove $\sum_{n\ge 0}^{\infty} \|E(\tilde{g}|\tilde{\mathcal{F}}_n)\|_2 < \infty$.

Let us start with a general fact concerning an arbitrary probability space $(X, \mathcal{F}, \mu)$ and a μ-preserving endomorphism T.

Lemma 4.7.2. Let U denote the unitary operator on $L^2(X, \mathcal{F}, \mu)$ associated to T, namely $U(f) = f \circ T$. Then for every $k \ge 0$ the operator $U^k U^{*k}$ is the orthogonal projection of $H_0 = L^2(X, \mathcal{F}, \mu)$ to $H_k = L^2(X, T^{-k}(\mathcal{F}), \mu)$.

Proof. U^* is the operator in the space conjugate to H_0 which is H_0 itself (a Hilbert space). $U^k(u) = u \circ T^k$ is measurable with respect to $T^{-k}(\mathcal{F})$, so the range of $U^k U^{*k}$ is indeed in $H_k = L^2(X, T^{-k}(\mathcal{F}), \mu)$.

For any $u, v \in H_0$ write $\int u \cdot v \, d\mu =< u, v >$, the scalar product of u and v. For arbitrary $f, g \in H_0$ we calculate

$$< U^k U^{*k}(f), g \circ T^k >=< U^k U^{*k}(f), U^k(g) >$$

$$=< U^{*k}(f), g >=< f, U^k(g) >=< f, g \circ T^k > .$$

It is clear that all functions in $H_k = L^2(X, T^{-k}(\mathcal{F}), \mu)$ are represented by $g \circ T^k$ for $g \in L^2(X, \mathcal{F}, \mu)$. Therefore by the above equality for all $h \in H_k$ we obtain

$$< f - U^k U^{*k}(f), h >=< f, h > - < f, h >= 0. \tag{4.7.2}$$

In particular for $f \in H_k$ we conclude from (4.7.2) for $h = f - U^k U^{*k}(f)$, that $< f - U^k U^{*k}(f), f - U^k U^{*k}(f) >= 0$ hence $U^k U^{*k}(f) = f$. Therefore $U^k U^{*k}$ is a projection to H_k, which is orthogonal by (4.7.2). ♣

Since the conditional expectation value $f \to E(f|T^{-k}(\mathcal{F}))$ is the orthogonal projection to H_k we conclude that $E(f|T^{-k}(\mathcal{F})) = U^k U^{*k}(f)$. Now, let us pass to our special situation of Theorem 4.7.1.

Lemma 4.7.3. For every $f \in L^2(X, \mathcal{F}, \mu)$ we have $U^*(f) = \hat{\mathcal{L}}(f)$.
Proof. $< U^* f, g >=< f, Ug >= \int f \cdot (g \circ T)\, d\mu = \int \hat{\mathcal{L}}(f \cdot (g \circ T))\, d\mu =$ $\int (\hat{\mathcal{L}}(f)) \cdot g\, d\mu =< \hat{\mathcal{L}}(f), g >$. ♣

Proof of Theorem 4.7.1. Conclusion. We can assume that $\mu(g) = 0$. We have

$$\sum_{n\ge 0}^{\infty} \|E(\tilde{g}|\tilde{\mathcal{F}}_n)\|_2 = \sum_{n\ge 0}^{\infty} \|U^n U^{*n}(g)\|_2 = \sum_{n\ge 0}^{\infty} \|\hat{\mathcal{L}}^n(g)\|_2 < \infty,$$

the latter has been assumed in (4.7.1). Thus CLT has been proved by applying Theorem 1.11.5. If g is Hölder continuous it satisfies (4.7.1). Indeed $\hat{\mathcal{L}}^k(g)$ converges to 0 in the sup norm exponentially fast as $k \to \infty$ by Corollary 4.4.8 (see (4.4.4")). This implies the same convergence in L^2 hence the convergence of the above series. ♣

Now let us prove CLT and LIL with the use of Theorem 1.11.1 for Hölder continuous g. As in Proof of Corollary 4.4.12, let $\pi : \Sigma_A \to X$ be a coding map from a 1-sided topological Markov chain of d symbols due to a Markov partition, see Ch.3.5. Since π is Hölder continuous, if g and ϕ are Hölder continuous, then the compositions $g \circ \pi, \phi \circ \pi$ are Hölder continuous. π is an isomorphism between the measures $\mu_{\phi\circ\pi}$ on Σ_A and μ_ϕ on X, see Ch.3.5 and Lemma 4.4.13. The function $g \circ \pi$ satisfies the assumptions of Theorem 1.11.1 with respect to the σ-algebra $\mathcal{F}$ associated to the partition of Σ_A into 0-th cylinders, see Theorem 4.4.11. ϕ-mixing follows from (4.4.7) and the estimate in (1.11.7) is exponential with an arbitrary δ due to the Hölder property of $g \circ \pi$. Hence, by Theorem 1.11.1, $g \circ \pi$ and therefore g satisfy CLT and LIL.

In Section 4.6 we computed the first derivative of the pressure function. Here using the same method we compute the second derivative and see that it is a respective dispersion (asymptotic variance) σ^2, see Ch.1.11.

Theorem 4.7.4. For every $\phi, u, v : X \to I\!R$ Hölder continuous functions there exists the second derivative

$$\frac{\partial^2}{\partial s \partial t} P(T, \phi + su + tv)|_{s=t=0} = \lim_{n\to\infty} \frac{1}{n} \int S_n(u - \mu_\phi u) S_n(v - \mu_\phi v)\, d\mu_\phi, \tag{4.7.2}$$

where μ_ϕ is the invariant Gibbs measure for ϕ. In particular

$$\frac{\partial^2}{\partial t^2} P(T, \phi + tv)|_{t=0} = \sigma^2_{\mu_\phi}(u)$$

(where the latter is the asymptotic variance discussed in CLT, Ch1.11). In addition, the function $(s,t) \mapsto P(T, \phi + su + tv)$ is C^2-smooth.

Proof. By Ch.4.6, see (4.6.3), (4.6.5),

$$\frac{\partial^2}{\partial s \partial t} P(T, \phi + su + tv)|_{t=0} = \frac{\partial}{\partial s} \lim_{n\to\infty} \frac{\frac{1}{n}\sum_{y\in T^{-n}(x)} S_n v(y) \exp S_n(\phi+su)(y)}{\sum_{y\in T^{-n}(x)} \exp S_n(\phi+su)(y)}. \tag{4.7.3}$$

Now we change the order of $\partial/\partial s$ and lim. This will be justified if we prove the uniform convergence of the resulting derivative functions.

Fixed $x \in X$ and n we abbreviate in the further notation $\sum_{y\in T^{-n}(x)}$ to $\sum_y$ and compute

$$F_n(s) := \frac{\partial}{\partial s}\Big(\frac{\sum_y S_n v(y) \exp S_n(\phi+su)(y)}{\sum_y \exp S_n(\phi+su)(y)}\Big) =$$

$$\frac{\sum_y S_n u(y) S_n v(y) \exp S_n(\phi+su)(y)}{\sum_y \exp S_n(\phi+su)(y)} -$$

$$\frac{\big(\sum_y S_n u(y) \exp S_n(\phi+su)(y)\big)\big(\sum_y S_n v(y) \exp S_n(\phi+su)(y)\big)}{\big(\sum_y \exp S_n(\phi+su)(y)\big)^2} =$$

$$\frac{\mathcal{L}^n\big((S_n u)(S_n v)\big)(x)}{\mathcal{L}^n(1\!\!1)(x)} - \frac{\mathcal{L}^n(S_n u)(x)}{\mathcal{L}^n(1\!\!1)(x)}\frac{\mathcal{L}^n(S_n v)(x)}{\mathcal{L}^n(1\!\!1)(x)}.$$

As in Section 6 we write here $\mathcal{L} = \mathcal{L}_0 = e^{-P(T,\phi+su)}\mathcal{L}_{\phi+su}$. It is useful to write the later expression for $F_n(s)$ in the form

$$F_n(s) = \int (S_n u)(S_n v)\, d\mu_{s,n} - \int (S_n u)\, d\mu_{s,n} \int (S_n v)\, d\mu_{s,n} \tag{4.7.4}$$

or

$$F_n(s) = \sum_{i,j=0}^{n-1} \int (u\circ T^i - \mu_{s,n}(u\circ T^i))(v\circ T^j - \mu_{s,n}(v\circ T^j))\, d\mu_{s,n}, \tag{4.7.5}$$

where $\mu_{s,n}$ is the probability measure distributed on $T^{-n}(x)$ according to the weights $\exp(S_n(\phi+su))(y)/\sum_y \exp S_n(\phi+su)(y)$.

Note that $\frac{1}{n}F_n(s)$ with $F_n(s)$ as in the formula (4.7.5) resembles already (4.7.2) because $\mu_{s,n} \to m_{\phi+su}$ in the weak*-topology, see (4.4.3'). However we still need to work a little bit.

For each i, j denote the respective summand in (4.7.5) by $K_{i,j}$. To simplify notation denote $u\circ T^i$ by u_i and $v\circ T^j$ by v_j. We have

$$K_{i,j} = \frac{\mathcal{L}^n\big((u_i - \mu_{s,n}u_i)(v_j - \mu_{s,n}v_j)\big)(x)}{\mathcal{L}^n(1\!\!1)(x)}$$

and for $0 \leq i \leq j < n$, using (4.4.5) twice,

$$K_{i,j} = \frac{\mathcal{L}^{n-j}\Big(\big(\mathcal{L}^{j-i}((u - \mu_{s,n}u_i)\mathcal{L}^i(1\!\!1))\big)\big(v - \mu_{s,n}v_j\big)\Big)(x)}{\mathcal{L}^n(1\!\!1)(x)}. \tag{4.7.6}$$

By Corollary 4.4.8 for $\tau < 1$ and Hölder norm $||\cdot||_{\mathcal{H}_\alpha}$ for an exponent $\alpha > 0$, transforming the integral as in Proof of Theorem 4.4.10, we get

$$||\mathcal{L}^{j-i}((u - \mu_{s,n}u_i)\mathcal{L}^i(1\!\!1)) - u_{\phi+su}\Big(\int u_i \, dm_{\phi+su} - \mu_{s,n}u_i)\Big)||_{\mathcal{H}_\alpha} \leq C\tau^{j-i}$$

where C depends only on Hölder norms of u and $\phi + su$. The difference in the large parentheses, denote it by $D_{i,n}$, is bounded by $C\tau^{n-i}$ in the Hölder norm, again by Corollary 4.4.8.

We conclude that for all j the functions

$$L_j := \sum_{i \leq j} \mathcal{L}^{j-i}((u - \mu_{s,n}u_i)\mathcal{L}^i(1\!\!1))$$

are uniformly bounded in the Hölder norm $||\cdot||_{\mathcal{H}_\alpha}$ by a constant C depending again only on $||u||_{\mathcal{H}_\alpha}$ and $||\phi + su||_{\mathcal{H}_\alpha}$. Hence summing over $i \leq j$ in (4.7.6) and applying $\mathcal{L}^{n-j}$ we obtain

$$\Big|\Big|\sum_{i=0}^{j} K_{i,j} - \sum_{i=0}^{j}\int (u_i - \mu_{s,n}u_i)(v_j - \mu_{s,n}v_j)\, dm_{\phi+su}\Big|\Big|_\infty \leq C\tau^{n-j}.$$

Here C depends also on $||v||_{\mathcal{H}_\alpha}$. We can replace the first sum by the second sum without changing the limit in (4.7.3) since after summing over $j = 0, 1, ..., n-1$, dividing by n and passing with n to ∞, they lead to the same result. Let us show now that $\mu_{s,n}$ can be replaced by $m_{\phi+su}$ in the above estimate without changing the limit in (4.7.3). Indeed, using the formula $ab - a'b' = (a - a')b' + a(b - b')$, we obtain

$$\begin{aligned}\Big|\int (u_i - m_{\phi+su}u_i)(v_j - m_{\phi+su}v_j)\, dm_{\phi+su} - \int (u_i - \mu_{s,n}u_i)(v_j - \mu_{s,n}v_j)\, dm_{\phi+su}\Big| \leq \\ |(\mu_{s,n}u_i - m_{\phi+su}u_i)\cdot(m_{\phi+su}v_j - \mu_{s,n}v_j) + \\ + \Big|\int (u_i - m_{\phi+su}u_i)\cdot(\mu_{s,n}v_j - m_{\phi+su}v_j)\, dm_{\phi+su}\Big|.\end{aligned}$$

Since $D_{i,n} \leq C\tau^{n-i}$ and $D_{j,n} \leq C\tau^{n-j}$, the first summand is bounded above by $\tau^{n-i}\tau^{n-j}$. Note that the second summand is equal to 0. Thus, our replacememnt is justified.

The last step is to replace $m = m_{\phi+su}$ by the invariant Gibbs measure $\mu = \mu_{\phi+su}$. Similarly as above we can replace m by μ in mu_i, mv_i. Indeed,

(4.7.7)

$$|mu_i - \mu u_i| = |\int u\cdot\mathcal{L}^i(1\!\!1)\, dm - \int uu_{\phi+su}\, dm| = |\int u\cdot(\mathcal{L}^i(1\!\!1) - u_{\phi+su})\, dm| \leq Cm(u)\tau^i.$$

Thus the resulting difference is bounded by $Cm(u)m(v)\tau^i\tau^j$. Finally we justify the replacement of m by μ at the second integral in the previous formula. To simplify notation write $F = u - \mu u, G = v - \mu v$. Since $j \geq i$, using (4.7.7), we can write

$$
\begin{aligned}
| \int (F \circ T^i)(G \circ T^j)\, dm - \int (F \circ T^i)(G \circ T^j)\, d\mu | = \\
&= | \int (F \cdot (G \circ T^{j-i})) \circ T^i \, dm - \int (F \cdot (G \circ T^{j-i})) \circ T^i \, d\mu| \\
&\leq C\tau^i \int |F \cdot (G \circ T^{j-i})|\, dm \leq Cm(F)m(G)\tau^i\tau^{j-i} = Cm(F)m(G)\tau^j
\end{aligned}
$$

by Theorem 4.4.10 (exponential decay of correlations), the latter C depending again on the Hölder norms of $u, v, \phi + su$. Summing over all $0 \leq i \leq j < n$ gives the bound $Cm(F)m(G)\sum_{j=0}^{n-1} j\tau^j$ and our replacement is justified. For $i > j$ we do the same replacements changing the roles of u and v. The C^2-smoothness follows from the uniformity of the convergence of the sequence of the functions $F_n(s)$, for $\phi + tv$ in place of ϕ, with respect to the variables (s, t), resulting from the proof. ♣

Exercises.

Exercise 1. Prove that (4.1.1) with an arbitrary $0 < \xi' \leq \xi$ in place of ξ implies (4.1.1) for every $0 < \xi' \leq \xi$ (with C depending on ξ') .

CHAPTER 6

FRACTAL DIMENSIONS

In the first section of this chapter we we provide a more complete treatment of outer measure indirectly begun in Chapter 1. The rest of this chapter is devoted to present basic definitions of pressure related to Hausdorff and packing measures, Hausdorff and packing dimensions of sets and measures and ball-counting dimensions.

§6.1 OUTER MEASURES

In Section 1.1 we have introduced the abstract notion of measure. In the beginning of this section we want to show how to construct measures starting with functions of sets called *outer measures* which are required to satisfy much weaker conditions. Our exposition of this material is brief and the reader should find its complete treatment in all handbooks of geometric measure theory (see for ex. [Falconer, 1985], [Ma] or [Pe]).

Definition 6.1.1 An *outer measure* on a set X is a function μ defined on all subsets of X taking values in $[0, \infty]$ such that

$$\mu(\emptyset) = 0, \tag{6.1.1}$$

$$\mu(A) \le \mu(B) \text{ if } A \subset B \tag{6.1.2}$$

and

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) \le \sum_{n=1}^{\infty} \mu(A_n) \tag{6.1.3}$$

for any countable family $\{A_n : n = 1, 2, \ldots\}$ of subsets of X.

A subset A of X is called *μ-measurable* or simply *measurable with respect to the outer measure μ* if and only if

$$\mu(B) \ge \mu(B \cap A) + \mu(B \setminus A) \tag{6.1.4}$$

for all sets $B \subset X$. Check that the opposite inequality follows immediately from (6.1.3). Check also that if $\mu(A) = 0$ then A is μ-measurable.

Theorem 6.1.2. If μ is an outer measure on X, then the family $\mathcal{F}$ of all μ-measurable sets is a σ-algebra and the restriction of μ to $\mathcal{F}$ is a measure.

Proof. Obviously $X \in \mathcal{F}$. By symmetry of (6.1.4), $A \in \mathcal{F}$ if and only if $A^c \in \mathcal{F}$. So, the conditions (1.1.1) and (1.1.2) of the definition of σ-algebra are satisfied. To check condition (1.2.3) that $\mathcal{F}$ is closed under countable union, suppose that $A_1, A_2, \ldots \in \mathcal{F}$ and let $B \subset X$ be any set. Applying (6.1.4) in turn to $A_1, A_2, \ldots$ we get for all $k \geq 1$

$$\begin{aligned}
\mu(A) &\geq \mu(B \cap A_1) + \mu(B \setminus A_1) \\
&\geq \mu(B \cap A_1) + \mu((B \setminus A_1) \cap A_2) + \mu(B \setminus A_1 \setminus A_2) \\
&\geq \ldots \\
&\geq \sum_{j=1}^{k} \mu\left((B \setminus \bigcup_{i=1}^{j-1} A_i) \cap \mathcal{A}_j \right) + \mu(B \setminus \bigcup_{j=1}^{k} A_j) \\
&\geq \sum_{j=1}^{k} \mu\left((B \setminus \bigcup_{i=1}^{j-1} A_i) \cap \mathcal{A}_j \right) + \mu(B \setminus \bigcup_{j=1}^{\infty} A_j)
\end{aligned}$$

and therefore

$$\mu(A) \geq \sum_{j=1}^{\infty} \mu\big((B \setminus \bigcup_{i=1}^{j-1} A_i) \cap \mathcal{A}_j\big) + \mu(B \setminus \bigcup_{j=1}^{\infty} A_j) \tag{6.1.5}$$

Since

$$B \cap \bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} (B \setminus \bigcup_{i=1}^{j-1} A_i) \cap A_j$$

using (6.1.3) we thus get

$$\mu(A) \geq \mu\big(\bigcup_{j=1}^{\infty} (B \setminus \bigcup_{i=1}^{j-1} A_i) \cap A_j\big) + \mu(B \setminus \bigcup_{j=1}^{\infty} A_j)$$

Hence condition (1.1.3) is also satisfied and $\mathcal{F}$ is a σ-algebra. To see that μ is a measure on $\mathcal{F}$ i.e. that condition (1.1.4) is satisfied, consider mutually disjoint sets $A_1, A_2, \ldots \in \mathcal{F}$ and apply (6.1.5) with $B = \bigcup_{j=1}^{\infty} A_j$ to get

$$\mu(\bigcup_{j=1}^{\infty} A_j) \geq \sum_{j=1}^{\infty} \mu(A_j)$$

Combining this with (6.1.3) we conclude that μ is a measure on $\mathcal{F}$. ♣

Now, let (X, ρ) be a metric space. An outer measure μ on X is said to be a *metric outer measure* if and only if

$$\mu(A \cup B) = \mu(A) + \mu(B) \tag{6.1.6}$$

for all *positively separated* sets $A, B \subset X$ that is satisfying the following condition

$$\rho(A, B) = \inf\{\rho(x, y) : x \in A, y \in B\} > 0$$

Recall that the Borel σ-algebra on X is that generated by open, or equivalently closed, sets. We want to show that if μ is a metric outer measure then the family of all μ-measurable sets contains this σ-algebra. The proof is based on the following version of Carathéodory's lemma.

Lemma 6.1.3. Let μ be a metric outer measure on (X, ρ). Let $\{A_n : n = 1, 2, \ldots\}$ be an increasing sequence of subsets of X and denote $A = \bigcup_{n=1}^{\infty} A_n$. If $\rho(A_n, A \setminus A_{n+1}) > 0$ for all $n \geq 1$, then $\mu(A) = \lim_{n\to\infty} \mu(A_n)$.

Proof. By (6.1.3) it is enough to show that

$$\mu(A) \leq \lim_{n\to\infty} \mu(A_n) \tag{6.1.7}$$

If $\lim_{n\to\infty} \mu(A_n) = \infty$, there is nothing to prove. So, suppose that

$$\lim_{n\to\infty} \mu(A_n) = \sup_n \mu(A_n) < \infty \tag{6.1.8}$$

Let $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for $n \geq 2$. If $n \geq m + 2$, then $B_m \subset A_m$ and $B_n \subset A \setminus A_{n-1} \subset A \setminus A_{m+1}$. Thus B_m and B_n are positively separated and applying (6.1.6) we get for every $j \geq 1$

$$\mu\Big(\bigcup_{i=1}^{j} B_{2i-1}\Big) = \sum_{i=1}^{j} \mu(B_{2i-1}) \ \text{ and } \ \mu\Big(\bigcup_{i=1}^{j} B_{2i}\Big) = \sum_{i=1}^{j} \mu(B_{2i}) \tag{6.1.9}$$

We have also for every $n \geq 1$

$$\begin{aligned} \mu(A) &= \mu\Big(\bigcup_{k=n}^{\infty} A_k\Big) = \mu\Big(A_n \cup \bigcup_{k=n+1}^{\infty} B_k\Big) \\ &\leq \mu(A_n) + \sum_{k=n+1}^{\infty} \mu(B_k) \leq \lim_{l\to\infty} \mu(A_l) + \sum_{k=n+1}^{\infty} \mu(B_k) \end{aligned} \tag{6.1.10}$$

Since the sets $\bigcup_{i=1}^{j} B_{2i-1}$ and $\bigcup_{i=1}^{j} B_{2i}$ appearing in (6.1.9) are both contained in A_{2j}, it follows from (6.1.8) and (6.1.9) that the series $\sum_{k=1}^{\infty} \mu(B_k)$ converges. Therefore (6.1.7) follows immediately from (6.1.10). The proof is finished. ♣

Theorem 6.1.4. If μ is a metric outer measure on (X, ρ) then all Borel subsets of X are μ-measurable.

Proof. Since the Borel sets form the least σ-algebra containing all closed subsets of X, it follows from Theorem 6.1.2 that it is enough to check (6.1.4) for every closed set $A \subset X$ and every $B \subset X$. For all $n \geq 1$ let $B_n = \{x \in B \setminus A : \rho(x, A) \geq 1/n\}$. Then $\rho(B \cap A, B_n) \geq 1/n$ and by (6.1.6)

$$\mu(B \cap A) + \mu(B_n) = \mu((B \cap A) \cup B_n) \leq \mu(B) \tag{6.1.10}$$

The sequence $\{B_n\}_{n=1}^{\infty}$ is increasing and, since A is closed, $B \setminus A = \bigcup_{n=1}^{\infty} B_n$. In order to apply Lemma 6.1.3 we shall show that

$$\rho(B_n, (B \setminus A) \setminus B_{n+1}) > 0$$

for all $n \geq 1$. And indeed, if $x \in (B \setminus A) \setminus B_{n+1}$, then there exists $z \in A$ with $\rho(x, z) < 1/(n+1)$. Thus, if $y \in B_n$, then

$$\rho(x, y) \geq \rho(y, z) - \rho(x, z) > 1/n - 1/(n+1) = \frac{1}{n(n+1)}$$

and consequently $\rho(B_n, (B \setminus A) \setminus B_{n+1}) > 1/n(n+1) > 0$. Applying now Lemma 6.1.3 with $A_n = B_n$ shows that $\mu(B \setminus A) = \lim_{n\to\infty} \mu(B_n)$. Thus (6.1.4) follows from (6.1.10). The proof is finished. ♣

§6.2 HAUSDORFF MEASURES

Let $\phi : [0, \infty) \to [0, \infty)$ be a non-decreasing function continuous at 0, positive on $(0, \infty)$ and such that $\phi(0) = 0$. Let (X, ρ) be a metric space. For every $\delta > 0$ define

$$\Lambda_\phi^\delta(A) = \inf\{\sum_{i=1}^{\infty} \phi(\mathrm{diam}(U_i))\} \tag{6.2.1}$$

where the infimum is taken over all countable covers $\{U_i : i = 1, 2, \ldots\}$ of A of diameter not exceeding δ. Conditions (6.1.1) and (6.1.2) are obviously satisfied with $\mu = \Lambda_\phi^\delta$. To check (6.1.3) let $\{A_n : n = 1, 2, \ldots\}$ be a countable family of subsets of X. Given $\varepsilon > 0$ for every $n \geq 1$ we can find a countable cover $\{U_i^n : i = 1, 2, \ldots\}$ of A_n of diameter not exceeding δ such that $\sum_{i=1}^{\infty} \phi(\mathrm{diam}(U_i^n)) \leq \Lambda_\phi^\delta(A_n) + \varepsilon/2^n$. Then the family $\{U_i^n : n \geq 1, i \geq 1\}$ covers $\bigcup_{n=1}^{\infty} A_n$ and

$$\Lambda_\phi^\delta(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \phi(\mathrm{diam}(U_i^n)) \leq \sum_{n=1}^{\infty} \Lambda_\phi^\delta(A_n) + \varepsilon$$

Thus, letting $\varepsilon \to 0$, (6.1.3) follows proving that Λ_ϕ^δ is an outer measure. Define

$$\Lambda_\phi(A) = \lim_{\delta\to 0} \Lambda_\phi^\delta(A) = \sup_{\delta>0} \Lambda_\phi^\delta(A) \tag{6.2.2}$$

The limit exists, but may be infinite, since $\Lambda_\phi^\delta(A)$ increases as δ decreases. Since all Λ_ϕ^δ are outer measures, the same argument also shows that Λ_ϕ is an outer measure. Moreover Λ_ϕ turns out to be a metric outer measure, since if A and B are two positively separated sets in X, then no set of diameter less than $\rho(A, B)$ can intersect both A and B. Consequently

$$\Lambda_\phi^\delta(A \cup B) = \Lambda_\phi^\delta(A) + \Lambda_\phi^\delta(B)$$

for all $\delta < \rho(A, B)$ and letting $\delta \to 0$ we get the same formula for Λ_ϕ which is just (6.1.6) with $\mu = \Lambda_\phi$. The metric outer measure Λ_ϕ is called the *Hausdorff outer measure* associated to the function ϕ. Its restriction to the σ-algebra of Λ_ϕ-measurable sets, which by Theorem 6.1.4 includes all the Borel sets, is called the *Hausdorff measure* associated to the function ϕ.

As an immediate consequence of the definition of Hausdorff measure and the properties of the function ϕ we get the following.

Proposition 6.2.1 The Hausdorff measure Λ_ϕ is non-atomic.

Remark 6.2.2. A particular role is played by functions ϕ of the form $t \to t^\alpha$, $t, \alpha > 0$ and in this case the corresponding outer measures are denoted by Λ_α^δ and Λ_α.

Remark 6.2.3. Note that if ϕ_1 is another function but such that ϕ_1 and ϕ restrected to an interval $[0, \varepsilon)$, $\varepsilon > 0$, are equal, then the outer measures Λ_{ϕ_1} and Λ_ϕ are also equal. So, in fact, it is enough to define the function ϕ only on an arbitrarily small interval $[0, \varepsilon)$.

Remark 6.2.4. Notice that we get the same values for $\Lambda_\phi^\delta(A)$, and consequently also for $\Lambda_\phi(A)$, if the infimum in (6.2.1) is taken only over covers consisting of sets contained in A. This means that the Hausdorff outer measure $\Lambda_\phi(A)$ of A is its intrinsic property, i.e. does not depend on in which space the set A is contained. If we treated A as the metric space $(A, \rho|_A)$ with the metric $\rho|_A$ induced from ρ, we would get the same value for the Hausdorff outer measure.

If we however took the infimum in (6.2.1) only over covers consisting of balls, we could get different "Hausdorff measure" which (dependently on ϕ) would need not be even equivalent with the Hausdorff measure just defined. To assure this last property ϕ is from now on assumed to satisy the following condition.

There exists a function $C : (0, \infty) \to [1, \infty)$ such that for every $a \in (0, \infty)$ and every $t > 0$ sufficiently small (dependently on a)

$$C(a)^{-1}\phi(t) \le \phi(at) \le C(a)\phi(t) \tag{6.2.3}$$

Since $(ar)^t = a^t r^t$, all functions ϕ of the form $r \to r^t$, considered in Remark 6.2.2, satisfy (6.2.3) with $C(a) = a^t$. Check that all functions $r \to r^t \exp(c\sqrt{\log\log 1/t \log\log\log 1/r}$, $c \ge 0$ also satisfy (6.2.3) with a suitable function C.

Definition 6.2.5. A countable collection $\{(x_i, r_i) : i = 1, 2, \ldots\}$ of pairs $(x_i, r_i) \in X \times (0, \infty)$ is said to *cover* a subset A of X if $A \subset \bigcup_{i=1}^{\infty} B(x_i, r_i)$, and is said to be *centered* at the set A if $x_i \in A$ for all $i = 1, 2, \ldots$. The *radius* of this collection is defined as $\sup_i r_i$ and its diameter as the diameter of the family $\{B(x_i, r_i) : i = 1, 2, \ldots\}$.

For every $A \subset X$ and every $r > 0$ let

$$\Lambda_\phi^{Br}(A) = \inf\{\sum_{i=1}^{\infty} \phi(r_i)\} \tag{6.2.4}$$

where the infimum is taken over all collections $\{(x_i, r_i) : i = 1, 2, \ldots\}$ centered at the set A, covering A and of radii not exceeding r. Let

$$\Lambda_\phi^{B}(A) = \lim_{r \to 0} \Lambda_\phi^{Br}(A) = \sup_{r>0} \Lambda_\phi^{Br}(A) \tag{6.2.5}$$

The limit exists by the same argument as used for the limit in (6.2.2). We shall prove the following.

Lemma 6.2.6. For every set $A \subset X$

$$1 \le \frac{\Lambda_\phi(A)}{\Lambda_\phi^B(A)} \le C(2)$$

Proof. Since the diameter of any ball does not exceed its double radius, since the diameter of any collection $\{(x_i, r_i) : i = 1, 2, \ldots\}$ also does not exceed its double radius and since the function ϕ is non- decreasing and satisfies (6.2.3), we see that for every $r > 0$ small enough

$$\sum_{i=1}^{\infty} \phi(\mathrm{diam}(B(x_i, r_i))) \le \sum_{i=1}^{\infty} \phi(2r_i) \le C(2) \sum_{i=1}^{\infty} \phi(r_i)$$

and therefore $\Lambda_\phi^{2r}(A) \le C(2)\Lambda_\phi^{Br}(A)$. Thus, letting $r \to 0$,

$$\Lambda_\phi(A) \le C(2)\Lambda_\phi^{B}(A) \tag{6.2.6}$$

On the other hand, let $\{U_i : i = 1, 2, \ldots\}$ be a countable cover of A consisting of subsets of A. For every $i \ge 1$ choose $x_i \in U_i$ and put $r_i = \mathrm{diam}(U_i)$. Then the collection $\{(x_i, r_i) : i = 1, 2, \ldots\}$ covers A, is centered at A and

$$\sum_{i=1}^{\infty} \phi(r_i) = \sum_{i=1}^{\infty} \phi(\mathrm{diam}(U_i))$$

which implies that $\Lambda_\phi^{B\delta}(A) \le \Lambda_\phi^{\delta}(A)$ for every $\delta > 0$. Thus $\Lambda_\phi^{B}(A) \le \Lambda_\phi(A)$ which combined with (6.2.6) finishes the proof. ♣

Remark 6.2.7. The function of sets Λ_ϕ^B need not to be an outer measure since condition (6.1.2) need not to be satisfied. Since we will be never interested in exact computation of Hausdorff measure, only in establishing its positiveness or finiteness or in comparing the ratio of its value with some other quantities up to bounded constants, we will be mostly dealing with $\Lambda_\phi^{B\delta}$ and Λ_ϕ^B using nevertheless always the symbols $\Lambda_\phi^\delta(A)$ and $\Lambda_\phi(A)$.

§6.3 PACKING MEASURES

Let, as in the previous section, $\phi : [0,\infty) \to [0,\infty)$ be a non-decreasing function such that $\phi(0) = 0$ and let (X, ρ) be a metric space. A collection $\{(x_i, r_i) : i = 1, 2, \dots\}$ centered at a set $A \subset X$ is said to be a *packing* of A if and only if for any pair $i \neq j$

$$\rho(x_i, x_j) \geq r_i + r_j$$

This property is not generally equivalent to requirement that all the balls $B(x_i, r_i)$ are mutually disjoint. It is obviously so if X is a Euclidean space. For every $A \subset X$ and every $r > 0$ let

$$\Pi_\phi^{*r}(A) = \sup\{\sum_{i=1}^{\infty} \phi(r_i)\} \tag{6.3.1}$$

where the supremum is taken over all packings $\{(x_i, r_i) : i = 1, 2, \dots\}$ of A of radius not exceeding r. Let

$$\Pi_\phi^*(A) = \lim_{r \to 0} \Pi_\phi^{*r}(A) = \inf_{r>0} \Pi_\phi^{*r}(A) \tag{6.3.2}$$

The limit exists since $\Pi_\phi^{*r}(A)$ decreases as r decreases. In opposite to Λ_ϕ^B the function Π_ϕ^* satisfies condition (6.1.2), however it also need not to be an outer measure since this time condition (6.1.3) need not to be satisfied. To obtain an outer measure we put

$$\Pi_\phi(A) = \inf\{\sum \Pi_\phi^*(A_i)\}, \tag{6.3.3}$$

where the supremum is taken over all covers $\{A_i\}$ of A. The reader will check easily, with similar arguments as in the case of Hausdorff measures, that Π_ϕ is already an outer measure and even more, a metric outer measure on X. It will be called the *outer packing measure* associated to the function ϕ. Its restriction to the σ-algebra of Π_ϕ-measurable sets, which by Theorem 6.1.4 includes all the Borel sets, will be called *packing measure* associated to the function ϕ.

Proposition 6.3.1. For every set $A \subset X$ it holds $\Lambda_\phi(A) \leq C(2)\Pi_\phi(A)$.

Proof. First we shall show that for every set $A \subset X$ and every $r > 0$

$$\Lambda_\phi^{2r}(A) \leq C(2)\Pi_\phi^{*r}(A) \tag{6.3.4}$$

Indeed, if there is no finite maximal (in the sense of inclusion) packing of the set A of the form $\{(x_i, r)\}$, then for every $k \ge 1$ there exists a packing $\{(x_i, r) : i = 1, \dots, k\}$ of A and therefore $\Pi_\phi^{*r}(A) \ge \sum_{i=1}^k \phi(r) = k\phi(r)$. Since $\phi(r) > 0$, this implies that $\Pi_\phi^{*r}(A) = \infty$ and (6.3.4) holds. Otherwise, let $\{(x_i, r) : i = 1, \dots, l\}$ be a maximal packing of A. Then the collection $\{(x_i, 2r) : i = 1, \dots, l\}$ covers A and therefore

$$\Lambda_\phi^{2r}(A) \le \sum_{i=1}^{l} \phi(2r) \le C(2) l \phi(r) \le C(2) \Pi_\phi^{*r}(A)$$

that is (6.3.4) is satisfied. Thus letting $r \to 0$ we get

$$\Lambda_\phi(A) \le C(2) \Pi_\phi^*(A) \tag{6.3.5}$$

So, if $\{A_n\}_{n \ge 1}$ is a countable cover of A then,

$$\Lambda_\phi(A) \le \sum_{n=1}^{\infty} \Lambda_\phi(A_i) \le C(2) \sum_{n=1}^{\infty} \Pi_\phi^*(A_i)$$

Hence, applying (6.3.3), the lemma follows. ♣

§6.4 DIMENSIONS

Let, similarly as in the two previous sections, (X, ρ) be a metric space. Recall (comp. Remark 6.2.2) that Λ_t, $t > 0$, is the Hausdorff outer measures on X associated to the function $r \to r^t$ and all Λ_t^δ are of corresponding meaning. Fix $A \subset X$. Since for every $0 < \delta \le 1$ the function $t \to \Lambda_t^\delta(A)$ is non-increasing, so is the function $t \to \Lambda_t(A)$. Furthermore, if $s < t$, then for every $0 < \delta$

$$\Lambda_s^\delta(A) \ge \delta^{s-t} \Lambda_t^\delta(A)$$

which implies that if $\Lambda_t(A)$ is positive, then $\Lambda_s(A)$ is infinite. Thus there is a unique value, $\mathrm{HD}(A)$, called the *Hausdorff dimension* of A such that

$$\Lambda_t(A) = \begin{cases} \infty & \text{if } 0 \le t < \mathrm{HD}(A) \\ 0 & \text{if } \mathrm{HD}(A) < t < \infty \end{cases} \tag{6.4.1}$$

Note that similarly as Hausdorff measures (comp. Remark 6.2.4), Hausdorff dimension is consequently also an intrinsic property of sets and does not depend on their complements. The following is an immediate consequence of the definitions of Hausdorff dimension and outer Hausdorff measures.

Theorem 6.4.1. The Hausdorff dimension is a monotonic function of sets, that is if $A \subset B$ then $\mathrm{HD}(A) \le \mathrm{HD}(B)$.

We shall prove the following.

Theorem 6.4.2. If $\{A_n\}_{n\geq 1}$ is a countable family of subsets of X then

$$\mathrm{HD}(\cup_n A_n) = \sup_n\{\mathrm{HD}(A_n)\}.$$

Proof. Inequality $\mathrm{HD}(\cup_n A_n) \geq \sup_n\{\mathrm{HD}(A_n)\}$ is an immediate consequence of Theorem 6.4.1. Thus, if $\sup_n\{\mathrm{HD}(A_n)\} = \infty$ there is nothing to prove. So, suppose that $s = \sup_n\{\mathrm{HD}(A_n)\}$ is finite and consider an arbitrary $t > s$. In view of (6.4.1), $\Lambda_t(A_n) = 0$ for every $n \geq 1$ and therefore, since Λ_t is an outer measure, $\Lambda_t(\cup_n A_n) = 0$. Hence, by (6.4.1) again, $\mathrm{HD}(\cup_n A_n) \leq t$. The proof is finished. ♣

As an immediate consequence of this theorem, Proposition 6.2.1 and formula (6.4.1) we get the following.

Proposition 6.4.3. The Hausdorff dimension of any countable set is equal to 0.

In exactly the same way as Hausdorff dimension HD one can define *packing* dimension* PD* and *packing dimension* PD using respectively $\Pi_t^*(A)$ and $\Pi_t(A)$ instead of $\Lambda_t(A)$. The reader can check easily that results analogous to Theorem 6.4.1, Theorem 6.4.2 and Proposition 6.4.3 are also true in these cases. As an immediate consequence of these definitions and Proposition 6.3.1 we get the following.

Lemma 6.4.4. $\mathrm{HD}(A) \leq \mathrm{PD}(A) \leq \mathrm{PD}^*(A)$ for every set $A \subset X$.

Now we shall define the third basic dimension – *ball-counting dimension* frequently also called box-counting dimension, Minkowski dimension or capacity. Let A be an arbitrary subset of the metric space (X,ρ). We first need the following.

Definition 6.4.5. For every $r > 0$ consider the family of all collections $\{(x_i, r_i)\}$ (see Definition 6.2.5) of radius not exceeding r which cover A and are centered at A. Put $N(A,r) = \infty$ if this family is empty. Otherwise define $N(A,r)$ to be the minimum of all cardinalities of elements of this family. Note that one gets the same number if one considers the subfamily of collections of radius exactly r and even only its subfamily of collections of the form $\{(x_i, r)\}$.

Now the *lower ball-counting dimensions* and *upper ball-counting dimension* of A are defined respectively by

$$\text{(6.4.2)} \qquad \underline{\mathrm{BD}}(A) = \liminf_{r\to 0}\frac{\log N(A,r)}{-\log r} \ \text{ and } \ \overline{\mathrm{BD}}(A) = \limsup_{r\to 0}\frac{\log N(A,r)}{-\log r}.$$

If $\underline{\mathrm{BD}}(A) = \overline{\mathrm{BD}}(A)$, the common value is called simply ball-counting dimension and is denoted by $\mathrm{BD}(A)$. The reader will easily prove the next theorem which explains the reason of the name box-counting dimension. The other names will not be discussed here.

Proposition 6.4.6. Fix $n \geq 1$. For every $r > 0$ let $\mathcal{L}(r)$ be any lattice in $I\!R^n$ consisting of cubes of sides of length r. For any set $A \subset I\!R^n$ let $L(A, r)$ denotes the number of cubes in $\mathcal{L}(r)$ which intersect A. Then

$$\underline{\mathrm{BD}}(A) = \liminf_{r \to 0} \frac{\log L(A, r)}{-\log r} \text{ and } \overline{\mathrm{BD}}(A) = \limsup_{r \to 0} \frac{\log L(A, r)}{-\log r}$$

Remark 6.4.7. Ball-counting dimension has properties which distinguish it qualitatively from Hausdorff and packing dimensions. For instance $\underline{\mathrm{BD}}(\overline{A}) = \underline{\mathrm{BD}}(A)$ and $\overline{\mathrm{BD}}(\overline{A}) = \overline{\mathrm{BD}}(A) =$. So, in particular there exist countable sets of positive ball-counting dimension, for example the set of rational numbers in the interval $[0, 1]$. Even more, there exist compact countable sets with this property like the set $\{1, 1/2, 1/3, \ldots, 0\} \subset I\!R$. On the other hand in many cases (see Theorem 6.6.6) all these dimensions coincide.

Now we shall provide other characterizations of ball-counting dimension, which in particular will be used to prove Lemma 6.4.8 and consequently Theorem 6.4.9 which establishes most general relations between the dimensions considered in this section.

Let $A \subset X$. For every $r > 0$ define $P(A, r)$ to be the supremum of cardinalities of all packings of the set A of the form $\{(x_i, r)\}$. First we shall prove the following.

Lemma 6.4.7. For every set $A \subset I\!R^n$ and every $r > 0$

$$N(A, 2r) \leq P(A, r) \text{ and } P(A, r) \leq N(A, r).$$

Proof. Let us start with the proof of the first inequality. If $P(A, r) = \infty$, there is nothing to prove. Otherwise, let $\{(x_i, r) : i = 1, \ldots, k\}$ be a packing of A with $k = P(A, r)$. Then this packing is maximal in the sense of inclusion and therefore the collection $\{(x_i, 2r) : i = 1, \ldots, l\}$ covers A. Thus $N(A, 2r) \leq l = P(A, r)$. The first part of Lemma 6.4.7 is proved.

If $N(A, r) = \infty$, the second part is obvious. Otherwise consider a finite packing $\{(x_i, r) : i = 1, \ldots, k\}$ of A and a finite cover $\{(y_j, r) : j = 1, \ldots, l\}$ of A centered at A. Then for every $1 \leq i \leq k$ there exists $1 \leq j = j(i) \leq l$ such that $x_i \in B(y_j(i), r)$ and every ball $B(y_j, r)$ can contain at most one element of the set $\{x_i : i = 1, \ldots, k\}$. So, the function $i \to j(i)$ is injective and therefore $k \leq l$. The proof is finished. ♣

As an immediate consequence of Lemma 6.4.7 we get the following.

$$\underline{\mathrm{BD}}(A) = \liminf_{r \to 0} \frac{\log P(A, r)}{-\log r} \text{ and } \overline{\mathrm{BD}}(A) = \limsup_{r \to 0} \frac{\log P(A, r)}{-\log r}. \tag{6.4.3}$$

Now we are in a position to prove the following.

Lemma 6.4.8 For every set $A \subset X$ we have $\mathrm{PD}^*(A) = \overline{\mathrm{BD}}(A)$.

Proof. Take $t < \overline{\mathrm{BD}}(A)$. In view of (6.4.3) there exists a sequence $\{r_n : n = 1, 2, \ldots\}$ of positive reals converging to zero and such that $P(A, r_n) \geq r_n^{-t}$ for every $n \geq 1$. Then

$\Pi_t^{*r_n}(A) \ge r^t P(A, r_n) \ge 1$ and consequently $\Pi_t^*(A) \ge 1$. Hence $t \le \mathrm{PD}^*(A)$ and therefore $\overline{\mathrm{BD}}(A) \le \mathrm{PD}^*(A)$.

In order to prove the converse inequality consider $s < t < \mathrm{PD}^*(A)$. Then $\Pi_t^*(A) = \infty$ and therefore for every $n \ge 1$ there exists a finite packing $\{(x_{n,i}, r_{n,i}) : i = 1, \ldots, k(n)\}$ of A of radius not exceeding 2^{-n} and such that

$$\sum_{i=1}^{k(n)} r_{n,i}^t > 1 \tag{6.4.4}$$

Now for every $m \ge n$ let

$$l_m = \#\{i \in \{1, \ldots, k(n)\} : 2^{-(m+1)} < r_{n,i} \le 2^{-m}\}$$

Then by (6.4.4)

$$\sum_{m=n}^{\infty} l_m 2^{-nt} > 1 \tag{6.4.5}$$

Suppose that $l_m < 2^{ns}(1 - 2^{(s-t)})$ for every $m \ge n$. Then

$$\sum_{m=n}^{\infty} l_m 2^{-nt} < (1 - 2^{(s-t)}) \sum_{m=1}^{\infty} 2^{n(s-t)} = 1$$

This contradicts (6.4.5) and shows that for every $n \ge 1$ there exists $m = m(n) \ge n$ such that

$$l_m \ge 2^{ns}(1 - 2^{(s-t)})$$

Hence $P(A, 2^{-(m+1)}) \ge 2^{ns}(1 - 2^{(s-t)})$, whence

$$\frac{\log P(A, 2^{-(m+1)})}{(m+1)\log 2} \ge \frac{sk \log 2}{(m+1)\log 2}$$

Thus, letting $n \to \infty$ (then also $m = m(n) \to \infty$) we obtain $\overline{\mathrm{BD}}(A) \ge s$. So, we are done. ♣

Combining now Lemma 6.4.4 and Lemma 6.4.8 and checking easily that $\mathrm{HD}(A) \le \underline{\mathrm{BD}}(A)$ we obtain the following main general relation connecting all the dimensions under consideration.

Theorem 6.4.9. For every set $A \subset X$

$$\mathrm{HD}(A) \le \min\{\mathrm{PD}(A), \underline{\mathrm{BD}}(A)\} \le \max\{\mathrm{PD}(A), \underline{\mathrm{BD}}(A)\} \le \overline{\mathrm{BD}}(A) = \mathrm{PD}^*(A)$$

We finish this section with the following definition.

Definition 6.4.10. Let μ be a Borel measure on (X, ρ). Then the *Hausdorff dimension* $\mathrm{HD}(\mu)$ of the measure μ is defined as

$$\mathrm{HD}(\mu) = \inf\{\mathrm{HD}(Y) : \mu(X \setminus Y) = 0\}$$

an analogous definition can be formulated for packing dimension.

§6.5 BESICOVITCH COVERING THEOREM

In this section we prove only one result, the Besicovitch covering theorem. Although this theorem seems to be almost always omitted in the classical geometric measure theory, we however consider it as one of most powerful geometric tools when dealing with some aspects of fractal sets. We refer the reader to Section 6.6 to verify our opinion.

Theorem 6.5.1. (Besicovitch covering theorem) Let $n \geq 1$ be an integer. Then there exists a constant $b(n) > 0$ such that the following claim is true.

If A is a bounded subset of $I\!R^n$ then for any function $r : A \to (0, \infty)$ there exists $\{x_k : k = 1, 2, \ldots\}$ a countable subset of A such that the collection $\mathcal{B}(A, r) = \{(x_k, r(x_k)) : k \geq 1\}$ covers A and can be decomposed into $b(n)$ packings of A.

In particular it follows from Theorem 6.5.1 that $\#\{B \in \mathcal{B} : x \in B\} \leq b(n)$. Exactly the same proof (world by world) goes if open balls in Theorem 6.5.1 are replaced by closed ones.

For any $x \in I\!R^n$, any $0 < r \leq \infty$ and any $0 < \alpha < \pi$ by $\mathrm{Con}(x, \alpha, r)$ we will denote any solid central cone with vertex x, radius r and angle (Lebesgue measure on the unit sphere S^{n-1}) α. The proof of Theorem 6.5.1 is based on the following obvious geometric observation.

Observation 6.5.2. Let $n \geq 1$ be an integer. Then there exists $\alpha(n) > 0$ so small that the following holds.

If $x \in I\!R^n$, $0 < r < \infty$, if $z \in B(x, r) \setminus B(x, r/3)$ and $x \in \mathrm{Con}(z, \alpha(n), \infty)$ then the set $\mathrm{Con}(z, \alpha(n), \infty) \setminus B(x, r/3)$ consists of two connected components (one of z and one of "∞") and that containing z is contained in $B(x, r)$.

Proof of Theorem 6.5.1. We will construct the sequence $\{x_k : k = 1, 2, \ldots\}$ inductively. Let

$$a_0 = \sup\{r(x) : x \in A\}$$

If $a_0 = \infty$ then one can find $x \in A$ with $r(x)$ so large that $B(x, r(x) \supset A$ and the proof is finished.

If $a_0 < \infty$ choose $x_1 \in A$ so that $r(x_1) > a_0/2$. Fix $k \geq 1$ and assume that the points $x_1, x_2, \ldots, x_k$ have been already chosen. If $A \subset B(x_1, r(x_1)) \cup \ldots \cup B(x_k, r(x_k))$ then the selection process is finished. Otherwise put

$$a_k = \sup\{r(x) : x \in A \setminus \big(B(x_1, r(x_1)) \cup \ldots \cup B(x_k, r(x_k))\big)\}$$

and take

$$x_{k+1} \in A \setminus \big(B(x_1, r(x_1)) \cup \ldots \cup B(x_k, r(x_k))\big) \tag{6.5.1}$$

such that

$$r(x_{k+1}) > a_k/2 \tag{6.5.2}$$

In order to shorten notation from now on throughout this proof we will write r_k for $r(x_k)$. By (6.5.1) we have $x_l \notin B(x_k, r_k)$ for all pairs k, l with $k < l$. Hence

$$\|x_k - x_l\| \geq r(x_k) \tag{6.5.3}$$

It follows from the construction of the sequence (x_k) that

$$r_k > a_{k-1}/2 \geq r_l/2 \tag{6.5.4}$$

and therefore $r_k/3 + r_l/3 < r_k/3 + 2r_k/3 = r_k$. Joining this and (6.5.3) we obtain

$$B(x_k, r_k/3) \cap B(x_l, r_l/3) = \emptyset \tag{6.5.5}$$

for all pairs k, l with $k \neq l$ since then either $k < l$ or $l < k$.

Now we shall show that the balls $\{B(x_k, r_k) : k \geq 1\}$ cover A. Indeed, if the selection process stops after finitely many steps this claim is obvious. Otherwise it follows from (6.5.5) that $\lim_{k\to\infty} r_k = 0$ and if $x \notin \bigcup_{k=1}^{\infty} B(x_k, r_k)$ for some $x \in A$ then by construction $r_k > a_{k-1}/2 \geq r(x)$ for every $k \geq 1$. The contradiction obtained proves that $\bigcup_{k=1}^{\infty} B(x_k, r_k) \supset A$.

The main step of the proof is given by the following.

Claim. For every $z \in \mathbb{R}^n$ and any cone $\mathrm{Con}(z, \alpha(n), \infty)$ ($\alpha(n)$ given by Observation 6.5.2)

$$\#\{k \geq 1 : z \in B(x_k, r_k) \setminus B(x_k, r_k/3) \text{ and } x_k \in \mathrm{Con}(z, \alpha(n), \infty)\} \leq 1 + 16^n$$

Denote by Q the set of integers whose cardinality is to be estimated. If $Q = \emptyset$, there is nothing to prove. Otherwise let $i = \min Q$. If $k \in Q$ and $k \neq i$ then $k > i$ and therefore $x_k \notin B(x_i, r_i)$. In view of this, Observation 6.5.2 applied with $x = x_i$, $r = r_i$, and the definition of Q, we get $\|z - x_k\| \geq 2r_i/3$, whence

$$r_k \geq \|z - x_k\| \geq 2r_i/3 \tag{6.5.6}$$

On the other hand by (6.5.4) we have $r_k < 2r_i$ and therefore $B(x_k, r_k/3) \subset B(z, 4r_k/3) \subset B(z, 8r_i/3)$. Thus, using (6.5.5), (6.5.6) and the fact that the n-dimensional volume of balls in $\mathbb{R}^n$ is proportional to the n^{th} power of radii we obtain $\#Q \leq (8r_i/3)^n/(2r_i/9)^n = 12^n$. The proof of the claim is finished.

Clearly there exists an integer $c(n) \geq 1$ such that for every $z \in I\!\!R^n$ the space $I\!\!R^n$ can be covered by at most $c(n)$ cones of the form $\mathrm{Con}(z, \alpha(n), \infty)$. Therefore it follows from the claim that for every $z \in I\!\!R^n$

$$\#\{k \geq 1 : z \in B(x_k, r_k) \setminus B(x_k, r_k/3)\} \leq c(n)12^n$$

Thus applying (6.5.5)

$$\#\{k \geq 1 : z \in B(x_k, r_k) \leq 1 + c(n)12^n \tag{6.5.7}$$

Since the ball $\overline{B}(0, 3/2)$ is compact, it contains a finite subset P such that $\bigcup_{x \in P} B(x, 1/2) \supset \overline{B}(0, 3/2)$. Now for every $k \geq 1$ consider the composition of the map $I\!\!R^n \ni x \to r_k x \in I\!\!R^n$ and the translation determined by the vector from 0 to x_k. Call by P_k the image of P under this translation. Then $\#P_k = \#P$, $P_k \subset \overline{B}(x_k, 3r_k/2)$ and

$$\bigcup_{x \in P_k} B(x, r_k/2) \supset \overline{B}(0, 3r_k/2) \tag{6.5.8}$$

Consider now two integers $1 \leq k < l$ such that

$$B(x_k, r_k) \cap B(x_l, r_l) \neq \emptyset \tag{6.5.9}$$

Let $y \in I\!\!R^n$ be the only point lying on the interval joining x_l and x_0 at the distance $r_k - r_l/2$ from x_k. As $x_l \notin B(x_k, r_k)$, by (6.5.9) we have $\|y - x_l\| \leq r_l + r_l/2 = 3r_l/2$ and therefore by (6.5.8) there exists $z \in P_l$ such that $\|z - y\| < r_l/2$. Consequently $z \in B(x_k, r_l/2 + r_k - r_l/2) = B(x_k, r_k)$. Thus applying (6.5.7) with z being the elements of P_l, we obtain the following

$$\#\{1 \leq k \leq l-1 : B(x_k, r_k) \cap B(x_l, r_l) \neq \emptyset\} \leq \#P(1 + c(n)12^n) \tag{6.5.10}$$

for every $l \geq 1$.

Putting $b(n) = \#P(1 + c(n)12^n) + 1$ this property allows us to decompose the set $I\!\!N$ of positive integers into $b(n)$ subsets $I\!\!N_1, I\!\!N_2, \ldots, I\!\!N_{b(n)}$ in the following inductive way. For every $k = 1, 2, \ldots, b(n)$ set $I\!\!N_k(b(n)) = \{k\}$ and suppose that for every $k = 1, 2, \ldots, b(n)$ and some $j \geq b(n)$ mutually disjoint families $I\!\!N_k(j)$ have been already defined so that

$$I\!\!N_1(j) \cup I\!\!N_{b(n)}(j) = \{1, 2, \ldots, j\}$$

Then by (6.5.10) there exists at least one $1 \leq k \leq b(n)$ such that $B(x_{j+1}, r_{j+1}) \cap B(x_i, r_i) = \emptyset$ for every $i \in I\!\!N_k(j)$. We set $I\!\!N_k(j+1) = I\!\!N_k(j) \cup \{j+1\}$ and $I\!\!N_l(j+1) = I\!\!N_l(j)$ for all $l \in \{1, 2, \ldots, b(n)\} \setminus \{k\}$. Putting now for every $k = 1, 2, \ldots, b(n)$

$$I\!\!N_k = I\!\!N_k(b(n)) \cup I\!\!N_k(b(n)+1) \cup \ldots$$

we see from the inductive construction that these sets are mutually disjoint, that they cover $I\!\!N$ and that for every $k = 1, 2, \ldots, b(n)$ the families of balls $\{B(x_l, r_l) : l \in I\!\!N_k\}$ are also mutually disjoint. The of proof the Besicovitch covering theorem is finished. ♣

We would like to emphasize here once more that the same statement remains true if open balls are replaced by closed ones. Also if instead of balls one considers n-dimensional cubes. Then although the proof is based on the same idea, however technically is considerably easier.

§6.6 VOLUME LEMMAS

In this section a function $\phi : [0, \infty) \to [0, \infty)$ is assumed to satisfy the same conditions as in Section 6.2 including (6.2.3) and moreover is assumed to be continuous. We start with the following.

Theorem 6.6.1. Let $n \geq 1$ be an integer and let $b(n)$ be the constant claimed in Theorem 6.5.1 (Besicovitch covering theorem). Assume that μ is a Borel probability measure on $I\!R^n$ and A is a bounded Borel subset of $I\!R^n$. If there exists $C \in (0, \infty]$, $(1/\infty = 0)$, such that

(a) for all (but countably many maybe) $x \in A$

$$\limsup_{r \to 0} \frac{\mu(B(x,r))}{\phi(r)} \geq C$$

then $\Lambda_\phi(E) \leq \frac{b(n)}{C}\mu(E)$ for every Borel set $E \subset A$. In particular $\Lambda_\phi(A) < \infty$.
or

(b) for all $x \in A$

$$\limsup_{r \to 0} \frac{\mu(B(x,r))}{\phi(r)} \leq C < \infty$$

then $\mu(E) \leq C\Lambda_\phi(E)$ for every Borel set $E \subset A$.

Proof. (a) In view of Proposition 6.2.1 we can assume that E does not intersect the exceptional countable set. Fix $\varepsilon > 0$ and $r > 0$. Since μ is a regular measure, there exists an open set $G \supset E$ such that $\mu(G) \leq \mu(E) + \varepsilon$. By openness of G and by assumption (a), for every $x \in E$ there exists $0 < r(x) < r$ such that $B(x, r(x)) \subset G$ and $(1/C + \varepsilon)\mu(B(x,r)) \geq \phi(r)$. Let $\{(x_k, r(x_k)) : k \geq 1\}$ be the cover of E obtained by applying Theorem 6.5.1. (Besicovitch covering theorem) to the set E. Then

$$\begin{aligned}\Lambda_\phi^r(E) &\leq \sum_{k=1}^\infty \phi(r(x_k)) \leq \sum_{k=1}^\infty (C^{-1} + \varepsilon)\mu(B(x_k, r(x_k))) \\ &\leq b(n)(C^{-1} + \varepsilon)\mu\Big(\bigcup_{k=1}^\infty B(x_k, r(x_k))\Big) \leq b(n)(C^{-1} + \varepsilon)(\mu(E) + \varepsilon)\end{aligned}$$

Letting $r \to 0$ we thus obtain $\Lambda_\phi(E) \leq b(n)(1/C + \varepsilon)(\mu(E) + \varepsilon)$ and therefore letting $\varepsilon \to 0$ the part (a) follows (note that the proof is correct with $C = \infty$!).

(b) Fix an arbitry $s > C$. Since for every $r > 0$ the function $x \to \mu(B(x,r))/\phi(r)$ is measurable and since the supremum of a countable sequence of measurable functions is also a measurable function, we conclude that for every $k \geq 1$ the function $\psi_k : A \to I\!R$ is measurable, where

$$\psi_k(x) = \sup\left\{ \frac{\mu(B(x,r))}{\phi(r)} : r \in Q \cap (0, 1/k] \right\}$$

and Q denotes the set of rational numbers. For every $k \geq 1$ let $A_k = \psi_k^{-1}((0,s])$. In view of measurability of the functions ψ_k all the sets A_k are measurable. Take an arbitrary $r \in (0, 1/k]$. Then there exists a sequence $r_j : j = 1, 2, \ldots\}$ of rational numbers converging to r from above. Since the function ϕ is continuous and the function $t \to \mu(B(x,t))$ is non-decreasing, we have for every $x \in A_k$

$$\frac{\mu(B(x,r))}{\phi(r)} \leq \lim_{j\to\infty} \frac{\mu(B(x,r_j))}{\phi(r_j)} \leq s$$

So, if $F \subset A_k$ is a Borel set and if $\{(x_i, r_i) : i = 1, 2, \ldots\}$ is a collection centered at the set F, covering F and of radius not exceeding $0 < r \leq 1/k$, then

$$\sum_{i=1}^{\infty} \phi(r_i) \geq s^{-1} \sum_{i=1}^{\infty} \mu(B(x_i, r_i)) \geq s^{-1}\mu(F)$$

Hence, $\Lambda_\phi^r(F) \geq s^{-1}\mu(F)$ and letting $r \to 0$ we get

$$\Lambda_\phi(E) \geq \Lambda_\phi(F) \geq s^{-1}\mu(F)$$

By the assumption of (b), $\cup_k A_k = A$ and therefore, putting $B_k = A_k \backslash (A_1 \cup A_2 \cup \ldots \cup A_{k-1})$, $k \geq 1$, we see that the family $\{B_k : k \geq 1\}$ is a countable partition of A into Borel sets. Therefore, if $E \subset A$ then

$$\Lambda_\phi(E) = \sum_{k=1}^{\infty} \Lambda_\phi(E \cap A_k) \geq s^{-1} \sum_{k=1}^{\infty} \mu(E \cap A_k) = s^{-1}\mu(E)$$

So, letting $s \searrow C$ finishes the proof. ♣

In an analogous way one can prove the following.

Theorem 6.6.2. Let $n \geq 1$ be an integer and let $b(n)$ be the constant claimed in Theorem 6.5.1 (Besicovitch covering theorem). Assume that μ is a Borel probability measure on $I\!R^n$ and A is a bounded subset of $I\!R^n$. If there exists $C \in (0, \infty]$, $(1/\infty = 0)$, such that

(a) for all $x \in A$

$$\liminf_{r\to 0} \frac{\mu(B(x,r))}{\phi(r)} \leq C$$

then $\mu(E) \le b(n)C\Pi_\phi(E)$ for every Borel set $E \subset A$.
or

(b) for all $x \in A$

$$\liminf_{r\to 0} \frac{\mu(B(x,r))}{\phi(r)} \ge C < \infty$$

then $\Pi_\phi(E) \le C^{-1}\mu(E)$ for every Borel set $E \subset A$. In particular $\Pi_\phi(A) < \infty$.

Note that each Borel measure μ defined on a Borel subset B of R^n can be in a canonical way considered as a measure on $I\!R^n$ by putting $\mu(A) = \mu(A \cap B)$ for every Borel set $A \subset I\!R^n$.
As a simple consequence of Theorem 6.6.1 we shall prove the following.

Theorem 6.6.3. Suppose that μ is a Borel probability measure on $I\!R^n$, $n \ge 1$, and A is a bounded Borel subset of $I\!R^n$.

(a) If $\mu(A) > 0$ and there exists θ_1 such that for every $x \in A$

$$\liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \ge \theta_1$$

then $\mathrm{HD}(A) \ge \theta_1$.
(b) If there exists θ_2 such that for every $x \in A$

$$\liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \le \theta_2$$

then $\mathrm{HD}(A) \le \theta_2$.
Proof. (a) Take any $0 < \theta < \theta_1$. Then, by the assumption, $\limsup_{r\to 0} \mu(B(x,r))/r^\theta \le 1$. Therefore applying Theorem 6.6.1(b) with $\phi(t) = t^\theta$, we obtain $\Lambda_\theta(A) \ge \mu(A) > 0$. Hence $\mathrm{HD}(A) \ge \theta$ by definition (6.4.1) and consequently $\mathrm{HD}(A) \ge \theta_1$.

(b) Take now an arbitrary $\theta > \theta_2$. Then by the assumption $\limsup_{r\to 0} \mu(B(x,r))/r^\theta \ge 1$. Therefore applying Theorem 6.6.1(a) with $\phi(t) = t^\theta$ we obtain $\Lambda_\theta(A) < \infty$, whence $\mathrm{HD}(A) \le \theta$ and consequently $\mathrm{HD}(A) \le \theta_1$. The proof is finished. ♣

Recall that the Hausdorff dimension of a Borel measure has been defined in Definition 6.4.10. As a consequence of Theorem 6.6.3 we shall prove the following.

Corollary 6.6.4. Suppose that μ is a Borel probability measure on $I\!R^n$, $n \ge 1$.

(a) If there exists θ_1 such that for μ-a.e. $x \in I\!R^n$

$$\liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \ge \theta_1$$

then $\mathrm{HD}(\mu) \ge \theta_1$

(b) If there exists θ_2 such that for μ-a.e. $x \in I\!R^n$

$$\liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \le \theta_2$$

then $\mathrm{HD}(\mu) \le \theta_2$.

Proof. (a) Let $Y \subset \mathbb{R}^n$ be a Borel set such that $\mu(Y) = 1$. By the assumption there exists a bounded Borel subset $A \subset Y$ with $\mu(A) > 0$ such that for every $x \in A$

$$\liminf_{r \to 0} \frac{\log \mu(B(x,r))}{\log r} \geq \theta_1$$

Thus, applying Theorem 6.6.3(a) we get $\mathrm{HD}(Y) \geq \mathrm{HD}(A) \geq \theta_1$ and taking infimum, $\mathrm{HD}(\mu) \geq \theta_1$.

(b) Decompose now the space $\mathbb{R}^n$ into a countable union $\cup_k X_k$ of Borel bounded sets X_k and let $X \subset \mathbb{R}^n$ be a Borel set of measure 1 whose every point satisfies the assumptions of Corollary 6.6.4. Applying for every $k \geq 1$ Theorem 6.6.3(b) with $A = X \cap X_k$ we get $\mathrm{HD}(X \cap X_k) \leq \theta_2$ and we are done applying Theorem 6.4.2 since $\mathrm{HD}(\mu) \leq \mathrm{HD}(X)$. ♣

Definition 6.6.5. Let X be a Borel bounded subset of $\mathbb{R}^n$, $n \geq 1$. A Borel probability measure on X is said to be a *geometric measure* with an exponent $t \geq 0$ if and only if there exists a constant $C \geq 1$ such that

$$C^{-1} \leq \frac{\mu(B(x,r))}{r^t} \leq C$$

for every $x \in X$ and every $0 < r \leq 1$.

We shall prove the following.

Theorem 6.6.6. If X is a Borel bounded subset of $\mathbb{R}^n$, $n \geq 1$, and μ is a geometric measure on X with an exponent t, then $\mathrm{BD}(X)$ exists and

$$\mathrm{HD}(X) = \mathrm{PD}(X) = \mathrm{BD}(X) = t$$

Moreover the three measures μ, Λ_t and Π_t on X are equivalent with bounded Radon-Nikodym derivatives.

Proof. The last part of the theorem follows immediately from Theorem 6.6.1 and Theorem 6.6.2 applied for $A = X$. Consequently also $t = \mathrm{HD}(X) = \mathrm{PD}(X)$ and therefore, in view of Theorem 6.4.9, we only need to show that $\overline{\mathrm{BD}}(X) \leq t$. And indeed, let $\{(x_i, r) : i = 1, \ldots, k\}$ be a packing of X. Then

$$kr^t \leq C \sum_{i=1}^{k} \mu(B(x_i, r)) \leq C$$

and therefore $k \leq Cr^{-t}$. Thus $P(X,r) \leq Cr^{-t}$, whence $\log P(X,r) \leq \log C - t \log r$. Applying now formula (6.4.3) finishes the proof. ♣

In particular it follows from this theorem that every geometric measure admits exactly one exponent. Lots of examples of geometric measures will be provided in the next chapters.

BIBLIOGRAPHICAL NOTES

The history of notions and development of the geometric measure theory is very long, rich and complicated and its outline exceeds the scope of this book. We refer the interested reader to the books [Falconer, 1985] and [Mattila, 1995].

CHAPTER 7. CONFORMAL EXPANDING REPELLERS

Nov. 23, 2002

Conformal expanding repellers (abbreviation: CER's) were defined already in Chapter 5 and some basic properties of expanding sets and repellers in dimension 1 were discussed in Section 5.2. A more advanced geometric theory in the real 1-dimensional case was done in Section 5.6 ???. But now we have a new tool: Frostman Lemma and related facts from Chapter 6. Equipped with the theory of Gibbs measures and with the pressure function we are able to develop a geometric theory of CER's with Hausdorff measures and dimension playing the crucial role. We shall present this theory for $C^{1+\varepsilon}$ conformal expanding repellers in $I\!R^d$. Remind (Ch.5.2) that the assumed conformality forces for $d=2$ that f is holomorphic or antiholomorphic and for $d \geq 3$ that f is locally a Möbius map. Conformality for $d=1$ is meaningless, so we assume $C^{1+\varepsilon}$ in order to be able to rely on the Bounded Distortion for Iteration lemma.

We shall outline a theory of Gibbs measures from the point of view of multifractal spectra of dimensions (Sec.2) and pointwise fluctuations due to the Law of Iterated Logarithm (Sec.3) .

For $d=2$ we shall apply this theory to study the boundary of the immediate basin of attraction to a sink for a rational mapping of the Riemann sphere in the case the basin is simply-connected and the mapping on the boundary is expanding, for example for the mapping $z \mapsto z^2+c$ for $|c|$ small, for a quasicircle invariant under the action of a quasi-Fuchsian group ?????????????????????? and for the boundary of the "snowflake". In particular we study harmonic measure. We shall derive from this an information about the radial growth of the derivative of the Riemann mapping from the unit disc to the simply-connected domain under consideration.

Section 7.1.1. Pressure function and dimension.

Let $f: X \to X$ be a topologically mixing conformal expanding repeller in $I\!R^d$. As before we abbreviate notation of the pressure $\mathrm{P}(f,\phi)$, to $\mathrm{P}(\phi)$. We start with the following technical lemma.

Lemma 7.1.1. Let m be a Gibbs state (not necesserily invariant) on X and let $\phi: X \to I\!R$ be a Hölder continuous function. Assume $\mathrm{P}(\phi)=0$. Then there is a constant $E \geq 1$ such that for all r small enough and all $x \in X$ there exists $n=n(x,r)$ such that

$$\text{(7.1.0)} \qquad \frac{\log E + S_n\phi(x)}{-\log E - \log|(f^n)'(x)|} \leq \frac{\log m(B(x,r))}{\log r} \leq \frac{-\log E + S_n\phi(x)}{\log E - \log|(f^n)'(x)|}.$$

Proof. Take an arbitrary $x \in X$. Fix $r \in (0, C^{-1}\xi)$ and let $n=n(x,r) \geq 0$ be the largest integer so that

$$\text{(7.1.1)} \qquad |(f^n)'(x)| r C \leq \xi,$$

where $C = C_{\mathrm{MD}}$ is the multiplicative distortion constant (corresponding to the Hölder continuous function $\log|f'|$), as in the Distortion Lemma for Iteration (Theorem 4.2.1), see Notation 5.2.2.??? Then

$$f_x^{-n}(B(f^n(x),\xi)) \supset B(x,\xi|(f^n)'(x)|^{-1}C^{-1}) \supset B(x,r). \tag{7.1.2}$$

Now take n_0 such that $\lambda^{n_0-1} \geq C^2$. We then obtain

$$|(f^{n+n_0})'|rC^{-1} \geq \xi. \tag{7.1.3}$$

Hence, again by the Distortion Lemma for Iteration

$$f_x^{-n-n_0}(B(f^{n+n_0}(x),\xi)) \subset B\big(x,\xi|(f^{n+n_0})'(x)|^{-1}C\big) \subset B(x,r). \tag{7.1.4}$$

By the Gibbs property of the measure m, see (4.1.1), for a constant $E \geq 1$ (the constant C in (4.1.1)) we can write

$$E^{-1} \leq \frac{\exp S_n\phi(x)}{m(f_x^{-n}(B(f^n(x),\xi)))} \quad \text{and} \quad \frac{\exp S_{n+n_0}\phi(x)}{m(f_x^{-(n+n_0)}(B(f^{n+n_0}(x),\xi)))} \leq E.$$

Using this, (7.1.2), (7.1.4), the inequality $S_{n+n_0}\phi(x) \geq S_n\phi(x) + n_0 \inf\phi$, and finally increasing E so that the new $\log E$ is larger than the old $\log E - n_0\inf\phi$, we obtain

$$\log E + S_n\phi(x) \geq \log m(B(x,r)) \geq -\log E + S_n\phi(x). \tag{7.1.5}$$

Using now (7.1.1) and (7.1.3), denoting $L = \sup|f'|$, and applying logarithms, we obtain

$$\frac{\log E + S_n\phi(x)}{\log|(f^n)'(x)|^{-1} - n_0\log L + \log\xi} \leq \frac{\log m(B(x,r)}{\log r} \leq \frac{-\log E + S_n\phi(x)}{\log|(f^n)'(x)|^{-1}\xi}.$$

Increasing further E so that $\log E \geq n_0\log L - \log\xi$, we can rewrite it in the "symmetric" form of (7.1.0). ♣

When we studied the pressure function $\phi \mapsto \mathrm{P}(\phi)$ in Chapters 2 and 4 the linear functional $\psi \mapsto \int \psi d\mu_\phi$ appeared. This was the Gateaux differential of P at ϕ (Theorem 2.5.5, Proposition 2.5.6 and (4.6.5')). Here the presence of an ambient smooth structure (1-dimensional or conformal) distingushes ψ's of the form $-t\log|f'|$. We obtain a link between the ergodic theory and the geometry of the embedding of X into $\mathbb{R}^d$.

Definition 7.1.2. Let μ be an ergodic f-invariant probability measure on X. Then by Birkhoff's Ergodic Theorem, for μ-almost every $x \in X$, the limit $\lim_{n\to\infty}\frac{1}{n}\log|(f^n)'(x)|$ exists and is equal to $\int \log|f'|d\mu$. We call this number the *Lyapunov characteristic exponent* of the map f with respect to the measure μ and we denote it by $\chi_\mu(f)$. In our case of expanding maps considered in this Chapter we obviously have $\chi_\mu(f) > 0$.

This definition does not demand the expanding property. It makes sense for an arbitrary invariant subset X of $\mathbb{R}^d$ or the Riemann sphere $\bar{\mathbb{C}}$, for f conformal (or differentiable in the real case) defined on a neighbourhood of X. There is no problem with the integrability because $\log|f'|$ is upper bounded on X. We do not exclude the possibility that $\chi_\mu = -\infty$. The notion of a Lyapunov characteristic exponent will play a crucial role also in subsequent chapters where non-expanding invariant sets will be studied.

Theorem 7.1.3. (Volume Lemma, expanding map, Gibbs measure case). Let m be a Gibbs state for a topologically mixing conformal expanding repeller $X \in \mathbb{R}^d$ and a Hölder continuous potential $\phi : X \to \mathbb{R}$. Then for m-almost every point $x \in X$ there exists the limit

$$\lim_{r\to 0} \frac{\log m(B(x,r))}{\log r}.$$

Moreover, this limit is almost everywhere constant and is equal to $\mathrm{h}_\mu(f)/\chi_\mu(f)$, where μ denotes the only f-invariant probability measure equivalent to m.

Proof. We can assume that $\mathrm{P}(\phi) = 0$. We can achieve it by subtracting $\mathrm{P}(\phi)$ from ϕ; the Gibbs measure class will stay the same (see Proposition 4.1.4). In view of the Birkhoff Ergodic Theorem, for μ-a.e $x \in X$ we have.

$$\lim_{n\to\infty} \frac{1}{n} S_n\phi(x) = \int \phi\, d\mu \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \log|(f^n)'(x)| = \chi_\mu(f).$$

Combining these equalities with (7.8.0), along with the observation that $n = n(x,r) \to \infty$ as $r \to 0$, and using also the equality $\mathrm{h}_\mu(f) + \int \phi d\mu = \mathrm{P}(\phi) = 0$, we conclude that

$$\lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} = \frac{\mathrm{h}_\mu(f)}{\chi_\mu(f)}.$$

The proof is finished. ♣

As an immediate consequence of this lemma and Corolaries 6.6.4 and 6.6.4a we get the following.

Theorem 7.1.4. If μ is a Gibbs state for a conformal expanding repeller $X \in \mathbb{R}^d$ and a Hölder continuous potential ϕ on X, then there exist Hausdorff and packing dimensions of μ and

$$\mathrm{HD}(\mu) = \mathrm{PD}(\mu) = \mathrm{h}_\mu(f)/\chi_\mu(f).$$

Using the above technique we can find a formula for the Hausdorff dimension and other dimensions of the whole set X. This is the solution of the non-linear problem, corresponding to the formula for Hausdorff dimension of the linear Cantor sets, discussed in the introduction. As f is Lipschitz continuous (or as f is forward expanding), the function

$$\mathrm{P}(t) := \mathrm{P}(-t \log|f'|)$$

is finite (see comments at the beginning of Section 2.5). As $|f'| \geq \lambda > 1$, it follows directly from the definition that $\mathrm{P}(t)$ is strictly decreasing from $+\infty$ to $-\infty$. In particular there exists exactly one parameter t_0 such that $\mathrm{P}(t_0) = 0$. We prove first the following.

Theorem 7.1.5. (Existence of geometric measures). Let t_0 be defined by $\mathrm{P}(t_0) = 0$. Write ϕ for $-t_0 \log |f'|$ restricted to X. Then each Gibbs state m corresponding to the function ϕ is a geometric measure with the exponent t_0. In particular $\lim_{r\to 0} \frac{\log m(B(x,r))}{\log r} = t_0$ for every $x \in X$.

Proof. We put in (7.1.0) $\phi = -t_0 \log |f'|$. Then using (7.1.1) (7.1.3) and $\sup |f'| \le L$ to replace $|(f^n)'(x)|^{-1}$ by r we obtain

$$\frac{\log E + t_0 \log r}{-\log E + \log r} \le \frac{\log m(B(x,r))}{\log r} \le \frac{-\log E + t_0 \log r}{\log E + \log r}$$

with a corrected constant E. Hence

$$\frac{\log E + t_0 \log r}{\log r} \le \frac{\log m(B(x,r))}{\log r} \le \frac{-\log E + t_0 \log r}{\log r} \tag{7.1.6}$$

for further corected E. In consequence

$$t_0 \le \frac{\log\bigl(m(B(x,r))/E\bigr)}{\log r} \quad \text{and} \quad \frac{\log\bigl(Em(B(x,r))\bigr)}{\log r} \le t_0,$$

hence

$$m(B(x,r))/E \le r^{t_0} \ \text{ and } \ Em(B(x,r)) \ge r^{t_0}.$$

(In the denominators we passed in Proof of Theorem 7.1.1 from r to $|(f^n)'(x)|^{-1}$ and here we passed back, so at this point the proof could be shortened. Namely we could deduce (7.1.6) directly from (7.1.5). However we needed to pass from $|(f^n)'(x)|^{-1}$ to r also in numerators and this point could not be simplified). ♣

As an immediate consequence of this theorem and Theorem 5.6.6 we get the following.

Corollary 7.1.6. The Hausdorff dimension of X is equal to t_0. Moreover it is equal to the packing and Minkowski dimensions. All Gibbs states corresponding to the potential $\phi = -t_0 \log |f'|$, as well as t_0-dimensional Hausdorff and packing measures are mutually equivalent with bounded Radon-Nikodym derivatives.

More on Volume Lemma. We end this section with a version of the Volume Lemma for a Borel probability invariant measure on the expanding repeller (X, f). In Chapter 9 we shall prove this without the expanding assumption assuming only positivity of Lyapunov exponent (though assuming also ergodicity) and the proof will be difficult. So we prove first a simpler version, which will be needed already in the next section. We start with a simple fact following from Lebesgue Theorem of differentiability a.e. ([Lojasiewicz, Th.7.1.4]) We provide a proof since it is very much in the spirit of Chapter 6.

Lemma 7.1.7. Every non-decreasing function $k : I \to I\!R$ defined on a bounded closed interval $I \subset I\!R$ is Lipschitz continuous at Lebesgue almost every point in I. In other words, for every $\varepsilon > 0$ there exist $L > 0$ and a set $A \subset I$ such that $|I \setminus A| < \varepsilon$, where $|\cdot|$ is the Lebesque measure in $I\!R$, and at each $r \in A$ the function k is Lipschitz continuous with the Lipschitz constant L.

Proof. Suppose on the contrary, that

$$B = \{x \in I : \sup\{y \in I : x \neq y, \ \frac{|k(x) - k(y)|}{|x-y|}\} = \infty\}$$

has positive Lebesgue measure. Write $I = [a, b]$. We can assume, by taking a subset, that B is compact and contains neither a nor b. For every $x \in B$ choose $x' \in I, x' \neq x$ such that

$$\frac{|k(x) - k(x')|}{|x - x'|} > 2\frac{k(b) - k(a)}{|B|}. \tag{7.1.6'}$$

Replace each pair x, x' by y, y' with $(y, y') \supset [x, x']$, and y, y' so close to x, x' that (7.1.6') still holds for y, y' instead of x, x'. In case when x or x' equals a or b we do not make the replacement.) We shall use for y, y' the old notation x, x' assuming $x < x'$. Now from the family of intervals (x, x') choose a finite family $\mathcal{I}$ covering our compact set B. From this family it is possible to choose a subfamily of intervals whose union still covers B and which consists of two subfamilies $\mathcal{I}^1$ and $\mathcal{I}^2$ of pairwise disjoint intervals.

Indeed. Start with $I_1 = (x_1, x_1') \in \mathcal{I}$ with minimal possible $x = x_1$ and maximal in $\mathcal{I}$ in the sense of inclusion. Having found $I_1 = (x_1, x_1'), ..., I_n = (x_n, x_n')$ we choose I_{n+1} as follows. Consider $\mathcal{I}_{n+1} := \{(x, x') \in \mathcal{I} : x \in \bigcup_{i=1,\dots,n} I_i, x' > \sup_{i=1,\dots,n} x_i'\}$. If $\mathcal{I}_{n+1}$ is non-empty, we set (x_{n+1}, x'_{n+1}) so that $x'_{n+1} = \max\{x' : (x, x') \in \mathcal{I}_{n+1}\}$. If $\mathcal{I}_{n+1} = \emptyset$, we set (x_{n+1}, x'_{n+1}) so that x_{n+1} is minimal possible to the right of $\max\{x_i' : i = 1, ..., n\}$ or equal to it, and maximal in $\mathcal{I}$. In this construction the intervals (x_n, x_n') with even n are pairwise disjoint, since each (x_{n+2}, x'_{n+2}) has not been a member of $\mathcal{I}_{n+1}$. The same is true for odd n's. We define $\mathcal{I}^i$ for $i = 1, 2$ as the family of (x_n, x_n') for even, respectively odd, n. In view of the pairwise disjointness intervals of families $\mathcal{I}^1$ and $\mathcal{I}^2$, monotonicity of k and (7.1.6'), we get that

$$k(b) - k(a) \geq \sum_{n \in \mathcal{I}^1} k(x_n') - k(x_n) > 2\frac{k(b) - k(a)}{|B|} \sum_{n \in \mathcal{I}^1} (x_n' - x_n)$$

and the similar inequality for $n \in \mathcal{I}^2$. Hence, taking into account that $\mathcal{I}^1 \cup \mathcal{I}^2$ covers B, we get

$$2(k(b) - k(a)) > 2\frac{k(b) - k(a)}{|B|} \sum_{n \in \mathcal{I}^1 \cup \mathcal{I}^2} (x_n' - x_n) \geq 2\frac{k(b) - k(a)}{|B|}|B| = 2(k(b) - k(a)),$$

which is a contradiction finishing the proof. ♣

Corollary 7.1.8. For every Borel probability measure ν on a compact metric space (X, ρ) and for every $r > 0$ there exists a finite partition $\mathcal{P} = \{P_t, t = 1, ..., M\}$ of X into Borel sets of positive measure ν and with $\mathrm{diam}(\mathcal{P}) \leq r$ and there exists $C > 0$ such that for every $a > 0$

$$\nu(\partial_{\mathcal{P},a}) \leq Ca, \tag{7.1.7}$$

where $\partial_{\mathcal{P},a} := \bigcap_t \Big(\bigcup_{s \neq t} B(P_s, a)\Big)$.

Proof. Let $\{x_1, ..., x_N\}$ be a finite $r/4$-net in X. Fix $\varepsilon \in (0, r/4N)$. For each function $t \mapsto k_i(t) := \nu(B(x_i, t))$, $t \in I = [r/4, r/2]$, apply Lemma 7.1.7 and find appropriate L_i and A_i, for all $i = 1, ..., N$. Let $L = \max\{L_i, i = 1, ..., N\}$ and let $A = \bigcap_{i=1,...,N} A_i$. The set A has positive Lebesgue measure by the choice of ε. So, we can choose its point r_0 different from $r/4$ and $r/2$. Therefore, for all $a < a_0 := \min\{r_0 - r/4, r/2 - r_0\}$ and for all $i \in \{1, 2, \ldots, n\}$, we have $\nu(B(x_i, r_0 + a) \setminus B(x_i, r_0 - a)) \leq 2La$. Hence, putting

$$\Delta(a) = \bigcup_i B(x_i, r_0 + a) \setminus B(x_i, r_0 - a)),$$

we get $\nu(\Delta(a)) \leq 2LNa$. Define $\mathcal{P} = \{\bigcap_{i=1}^N B^{\kappa(i)}(x_i, r_0)\}$ as a family over all functions $\kappa : \{1, ..., N\} \to \{+, -\}$, where $B^+(x_i, r_0) := B(x_i, r_0)$ and $B^-(x_i, r_0) := X \setminus B(x_i, r_0)$, except κ yielding sets of measure 0, in particular except empty intersections. After removing from X of a set of measure 0, the partition $\mathcal{P}$ covers X. Since $r_0 \geq r/4$, the balls $B(x_i, r_0)$ cover X. Hence, for each non-empty $P_t \in \mathcal{P}$ at least one value of κ is equal to $+$. Hence $\mathrm{diam}(P_t) \leq 2r_0 < r$. Note now that $\partial_{\mathcal{P},a} \subset \Delta(a)$. Indeed, let $x \in \partial_{\mathcal{P},a}$. Since $\mathcal{P}$ covers X there exists t_0 such that $x \in P_{t_0}$ so $x \notin P_t$ for all $t \neq t_0$. However, since $x \in \bigcup_{t \neq t_0} B(P_t, a)$, there exists $t_1 \neq t_0$ such that $\mathrm{dist}(x, P_{t_1}) < a$. Let $B = B(x_i, r_0)$ be such that $P_{t_0} \subset B^+$ and $P_{t_1} \subset B^-$, or *vice versa*. In the case when $x \in P_{t_0} \subset B^+$, by the triangle inequality $\rho(x, x_i) > r_0 - a$ and since $\rho(x, x_i) < r_0$, we get $x \in \Delta(a)$. In the case $x \in P_{t_0} \subset B^-$ we have $x \in B(x_i, r_0 + a) \setminus B(x_i, r_0) \subset \Delta(a)$. We conclude that $\nu(\partial_{\mathcal{P},a})) \leq \nu(\Delta(a) \leq 2LNa$ for all $a < a_0$. For $a \geq a_0$ it suffices to take $C \geq 1/a_0$. So the corollary is proved, with $C = \max\{2LN, 1/a_0\}$. ♣

Remark. If X is embedded for example in a compact manifold Y, then we can view ν as a measure on Y, we find a partition $\mathcal{P}$ of Y and then $\partial_{\mathcal{P},a} = B(\bigcup_{t=1,,,.M} \partial P_t, a)$, provided $M \geq 2$. This justifies the notation $\partial_{\mathcal{P},a}$.

Corollary 7.1.9. Let ν be a Borel probability measure on a compact metric space (X, ρ) andlet $f : X \to X$ be an endomorphism measurable with respect to the Borel σ-algebra on X and preserving measure ν. Le for every $r > 0$ let $\mathcal{P} = \{P_t, t = 1, ..., M\}$ be the partition of X constructed in Corollary 7.1.8. In particular $\mathrm{diam}(\mathcal{P} \leq r$. Then for every $\delta > 0$ and ν-a.e. $x \in X$ there exists $n_0 = n_0(x)$ such that for every $n \geq n_0$

$$B(f^n(x), \exp(-n\delta)) \subset \mathcal{P}(f^n(x)) \tag{7.1.8}$$

Proof. Let $\mathcal{P}$ be the partition from Corollary 7.1.8. Fix an arbitrary $\delta > 0$. Then by Corollary 7.1.8

$$\sum_{n=0}^{\infty} \nu(\partial_{\mathcal{P},\exp(-n\delta)})) \leq \sum_{n=0}^{\infty} C \exp(-n\delta) < \infty.$$

Hence by the f-invariance of ν, we obtain

$$\sum_{n=0}^{\infty} \nu(f^{-n}(\partial_{\mathcal{P},\exp(-n\delta)})) < \infty.$$

Applying now the Borel-Cantelli lemma for the family $\{f^{-n}(\partial_{\mathcal{P},\exp(-n\delta)})\}_{=1}^{\infty}$ we conclude that for ν-a.e $x \in X$ there exists $n_0 = n_0(x)$ such that for every $n \geq n_0$ we have $x \notin f^{-n}(\partial_{\mathcal{P},\exp(-n\delta)})$, so $f^n(x) \notin \partial_{\mathcal{P},\exp(-n\delta)}$. Hence, by the definition of $\partial_{\mathcal{P},\exp(-n\delta)}$, if $f^n(x) \in P$ for some $P \in \mathcal{P}$, then $f^n(x) \notin \bigcup_{s\neq t} B\big(P_s, \exp(-n\delta)\big)$. Thus

$$B(f^n(x), \exp -n\delta) \subset P.$$

We are done. ♣

Theorem 7.1.10. (Volume Lemma, expanding map, any measure case). Let ν be an f-invariant Borel probability measure on a topologically exact conformal expanding repeller (X, f), where $X \subset I\!R^d$.Then

$$\mathrm{HD}_*(\nu) \leq \frac{\mathrm{h}_\nu(f)}{\chi_\nu(f)} \leq \mathrm{HD}^*(\nu).$$

If in addition ν is ergodic, then

$$\mathrm{HD}(\nu) = \frac{\mathrm{h}_\nu(f)}{\chi_\nu(f)}.$$

Proof. Fix the partition $\mathcal{P}$ coming from Lemma 7.1.8 with $r = \min\{\xi, \eta\}$, where $\eta >$ was defined in (3.1.1). Then, as we saw in Chapter 4

$$\mathcal{P}^{n+1}(x) \subset f_x^{-n}(B(f^n(x), \xi)). \tag{7.1.10}$$

for every $x \in X$ and all $n \geq$. We shall work now to get a sort of opposite inclusion. Consider an arbitrary $\delta > 0$ and x so that (7.1.8) from Corollary 7.1.9 is satisfied for al $n \geq n_0(x)$. For every $0 \leq i \leq n$ define $k(i) = [i\frac{\delta}{\log\lambda} + \frac{\log\xi}{\log\lambda}] + 1$, $\lambda > 1$ being the expanding constant for $f : X \to X$ (see (3.1.1)). Hence $\exp(-i\delta) \geq \xi\lambda^{-k}$ and therefore $f_{f^i(x)}^{-k}(B(f^{i+k}(x), \xi)) \subset B(f^i(x), \exp -i\delta)$. So, using (7.1.8) for i in place of n, we get

$$f_x^{-(i+k)}(B(f^{i+k}(x), \xi)) \subset f_x^{-i}(\mathcal{P}(f^i(x))$$

for all $i \geq n_0(x)$. From this estimate for all $n_0 \leq i \leq n$, we conclude that

$$f_x^{-(n+k(n))}(B(f^{n+k(n)}(x), \xi) \subset \mathcal{P}_{n_0}^{n+1}(x).$$

Notice that for ν-a.e. x there is $a > 0$ such that $B(x, a) \subset \mathcal{P}^{n_0}(x)$, by the definition of $\partial_{\mathcal{P},\cdot}$. Therefore for all n large enough

$$f_x^{-(n+k(n))}(B(f^{n+k(n)}(x), \xi)) \subset \mathcal{P}^n(x). \tag{7.1.11}$$

It follows from (7.1.11) and (7.1.10) with $n + k(n)$ in place of n, that

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} - \log\nu(\mathcal{P}^n(x)) &\leq \liminf_{n\to\infty} \frac{-\log\nu\big(f_x^{-(n+k(n))}(B(f^{n+k(n)}(x), \xi))\big)}{n} \\
&\leq \limsup_{n\to\infty} \frac{-\log\nu(f_x^{-(n+k(n))}(B(f^{n+k(n)}(x), \xi))}{n} \\
&\leq \lim_{n\to\infty} \frac{1}{n} - \log\nu(\mathcal{P}^n(x))(\mathcal{P}^{n+k(n)+1})(x).
\end{aligned}$$

The limits on the most left and most right-hand sides of these inequalities exist for ν-a.e. x by the Shennon-McMillan-Breiman Theorem (Theorem 1.5.4), see also (1.5.1), and their ratio is equal to 1. Letting $\delta \to 0$ we obtain the existence of the limit and the equality

$$\text{(7.1.12)} \quad \mathrm{h}_\nu(f, \mathcal{P}, x) := \lim_{n\to\infty} \frac{1}{n} - \log \nu(\mathcal{P}^n(x))(\mathcal{P}^n)(x) = \lim_{n\to\infty} \frac{-\log \nu(f_x^{-n}(B(f^n(x), \xi))}{n}.$$

In view of Birkhoff's Ergodic Theorem, thelimit

$$\text{(7.1.13)} \qquad \chi_\nu(f, x) := \lim_{n\to\infty} \frac{1}{n} \log |(f^n)'(x)|,$$

exists for ν-a.e. $x \in X$. Dividing side by side (7.1.12) by (7.1.13) and using (7.1.1)-(7.1.4), we get

$$\lim_{r\to 0} \frac{\log \nu(B(x,r))}{\log r} = \frac{\mathrm{h}_\nu(f, \mathcal{P}, x)}{\chi_\nu(f, x)}.$$

Since by the Shennon-McMillan-Breiman Theorem, and Birkhoff's Ergodic Theorem,

$$\frac{\int \mathrm{h}_\nu(f, \mathcal{P}, x)\, d\nu(x)}{\int \chi_\nu(f, x)\, d\nu} = \frac{\mathrm{h}_\nu(f, \mathcal{P})}{\chi_\nu(f)}. = \frac{\mathrm{h}_\nu(f)}{\chi_\nu(f)},$$

where the latter equality was written since f is expansive and diam$(\mathcal{P})$ is less than the expansiveenss constant of $f : X \to X$ which at least excesds η, there thus exists a positive measure set where $\frac{\mathrm{h}_\nu(f,\mathcal{P},x)}{\chi_\nu(f,x)} \le \frac{\mathrm{h}_\nu(f)}{\chi_\nu(f)}$ and a positive measure set where the opposite inequality holds. Therefore

$$\lim_{r\to 0} \frac{\log \nu(B(x,r))}{\log r} \le \frac{\mathrm{h}_\nu(f)}{\chi_\nu(f)}$$

and the opposite inequality also holds on a positive measure set. In view of definitions of HD_* and HD^* and by Corollary 6.6.4, this finishes the proof of the first part of our Theorem. In the ergodic case $\mathrm{h}_\nu(f, \mathcal{P}, x) = \mathrm{h}_\nu(f)$ and $\chi_\nu(f, x) = \chi_\nu(f)$ for ν-a.e. $x \in X$. So

$$\lim_{r\to 0} \frac{\log \nu(B(x,r))}{\log r} = \frac{\mathrm{h}_\nu(f)}{\chi_\nu(f)}$$

and we are done in this case as well. ♣

Section 7.8. Multifractal analysis of Gibbs state.

In the previous section we linked to a (Gibbs) measure only one dimension number, HD(m). Here one of our aims is to introduce 1-parameter families of dimensions, so-called spectra of dimensions. In these definitions we do not need the mapping f. Let ν be a Borel probability measure on a metric space X. Recall from Chapter 6.7 that given $x \in X$ we defined the lower and upper pointwise dimension of ν at x by putting respectively

$$\underline{d}_\nu(x) = \liminf_{r\to 0} \frac{\log \nu(B(x,r))}{\log r} \text{ and } \overline{d}_\nu(x) = \limsup_{r\to 0} \frac{\log \nu(B(x,r))}{\log r}.$$

If $\underline{d}_\nu(x) = \overline{d}_\nu(x)$, we call the common value the pointwise dimension of ν at x and we denote it by $d_\nu(x)$. The function d_ν is called the *dimension spectrum* of the measure ν. For any $\alpha \leq 0 \leq \infty$ write

$$X_\nu(\alpha) = \{x \in X : d_\nu(x) = \alpha\}.$$

The domain of d_ν namely the set $\bigcup_\alpha X_\nu(\alpha)$ is called *a regular part* of X and its complement $\hat{X}$ a *singular part.* The decomposition of the set X as

$$X = \bigcup_{0 \leq \alpha \leq \infty} X_\nu(\alpha) \cup \hat{X}.$$

is called the *multifractal decomposition with respect to the dimension spectrum.*

Define the $F_\nu(\alpha)$-*spectrum for dimensions* function related to Hausdorff dimension by

$$F_\nu(a) = \mathrm{HD}(X_\nu(\alpha)),$$

where we define the domain of F_ν as $\{\alpha : X_\nu(\alpha) \neq \emptyset\}$.

Note that by Theorem 7.1.5 if (X, f) is a topologically exact expanding conformal repeller and $\nu = \mu_{-\mathrm{HD}(X)\log|f'|}$ then all $X_\nu(\alpha)$ are empty except $X_\nu(\mathrm{HD}(X))$. In particular the domain of F_ν is in this case just one point $\mathrm{HD}(X)$.
Let for every real $q \neq 1$

$$R_q(\nu) := \frac{1}{q-1} \lim_{r \to 0} \frac{\log \sum_{i=1}^N \nu(B_i)^q}{\log r},$$

where $N = N(r)$ is the total number of boxes B_i of the form $B_i = \{(x_1, ..., x_d) \in I\!R^d : rk_j \leq x_j \leq r(k_j + 1),\ j = 1, ..., d\}$ for integers $k_j = k_j(i)$ such that $\nu(B_i) > 0$. This function is called *Rényi spectrum for dimensions*, provided the limit exists. It is easy to check (exercise 7.2.1) that it is equal to the *Hentschel-Procaccia spectrum*

$$HP_q(\nu) := \frac{1}{q-1} \lim_{r \to 0} \frac{\log \inf_{\mathcal{G}_r} \sum_{B(x_i,r) \in \mathcal{G}_r} \nu(B(x_i, r))^q}{\log r},$$

where infimum is taken over all $\mathcal{G}_r$ being finite or countable coverings of the (topological) support of ν by balls of radius r centered at $x_i \in X$, or

$$HP_q(\nu) := \frac{1}{q-1} \lim_{r \to 0} \frac{\log \int_X \nu(B(x, r))^{q-1} d\nu(x)}{\log r}$$

provided the limits exist. For $q = 1$ we define the *information dimension* $I(\nu)$ as follows. Set

$$H_\nu(r) = \inf_{\mathcal{F}_r} \Big(- \sum_{B \in \mathcal{F}_r} \nu(B) \log \nu(B) \Big),$$

where infimum is taken over all partitions $\mathcal{F}_r$ of a set of full measure ν into Borel sets B of diameter at most r. We define

$$I(\nu) = \lim_{r\to 0} \frac{H_\nu(r)}{-\log r}$$

provided the limit exists. A complement to Corollary 6.6.4 is that

$$\text{(7.8.0)} \qquad \mathrm{HD}_*(\nu) \le I(\nu) \le \mathrm{PD}^*(\nu).$$

For the proof see Exercise 7.2.5.Note that for Rényi and HP dimensions it does not make any difference whether we consider coverings of the topological support (the smallest closed set of full measure) of a measure or any set of full measure, since all balls have the same radius r, so we can always choose locally finite (number independent of r) subcovering. These are "box type" dimension quantities.

A priori there is no reason for the function $F_\nu(\alpha)$ to behave nicely. If ν is an f-invariant ergodic measure for (X, f), a topologically exact conformal expanding repeller, then at least we know that for $\alpha_0 = \mathrm{HD}(\nu)$, we have $d_\nu(x) = \alpha_0$ for ν-a.e. x (by the Volume Lemma: 7.1.3 and Theorem 7.1.4 for a Gibbs measure ν of a Hölder continuous function and by Theorem 7.1.10 in the general case). So, in particular we know at least that the domain of $F_\alpha(\nu)$ is nonempty. However for $\alpha \ne \alpha_0$ we have then $\nu(X_\nu(\alpha)) = 0$ so $X_\nu(\alpha)$ are not visible for the measure ν. Whereas the function $HP_q(\nu)$ can be determined by statistical properties of ν-typical (a.e.) trajectory, the function $F_\nu(\alpha)$ seems intractable. However if $\nu = \mu_\phi$ is an invariant Gibbs measure for a Hölder continuous function (potential) ϕ, then miraculously the above spectra of dimensions happen to be real-analytic functions and $-F_{\mu_\phi}(-p)$ and $HP_q(\mu_\phi)$ are mutual Legendre transforms. Compare this with the pair of Legendre-Fenchel transforms: pressure and -entropy, Remark 2.5.3. Thus fix an invariant Gibbs measure μ_ϕ corresponding to a Hölder continuous potential ϕ. We can assume without loosing generality that $\mathrm{P}(\phi) = 0$. Indeed, starting from an arbitrary ϕ, we can achieve this without changing μ_ϕ by subtracting from ϕ its topological pressure (as at the beginning of the proof of Lemma 7.1.3). Having fixed ϕ, in order to simplify notation, we denote $X_{\mu_\phi}(\alpha)$ by X_α and F_{μ_ϕ} by F. We define a two-parameter family of auxiliary functions $\phi_{q,t} : X \to I\!\!R$ for $q, t \in I\!\!R$, by setting

$$\phi_{q,t} = -t \log|f'| + q\phi.$$

Lemma 7.8.1. For every $q \in I\!\!R$ there exists a unique $t = T(q)$ such that $\mathrm{P}(\phi_{q,t}) = 0$.

Proof. This lemma follows from the fact the function $t \mapsto \mathrm{P}(\phi_{q,t})$ is decreasing from ∞ to $-\infty$ for every q (see comments preceding Theorem 7.1.5 and at the beginning of Section 2.5) and the Darboux theorem. ♣

We deal with invariant Gibbs measures $\mu_{\phi_{q,T(q)}}$ which we denote for abbreviation by μ_q and with the measure μ_ϕ so we need to know a relation between them. This is explained in the following.

Lemma 7.8.2. For every $q \in I\!\!R$ there exists $C > 0$ such that for all $x \in X$ and $r > 0$

$$C^{-1} \le \frac{\mu_q(B(x,r))}{r^{T(q)}\mu_\phi(B(x,r))^q} \le C. \tag{7.8.1}$$

Proof. Let $n = n(x,r)$ be defined as in Lemma 7.1.1. Then, by (7.1.5), (7.1.1) and (7.1.3), the ratios

$$\frac{\mu_\phi(B(x,r))}{\exp S_n\phi(x)}, \quad \frac{\mu_q(B(x,r))}{|(f^n)'(x)|^{-T(q)}\exp qS_n\phi(x)}, \quad \frac{r}{|(f^n)'(x)|^{-1}}$$

are bounded from below and above by positive constants independent of x, r. This yields the estimates (7.8.1) ♣

Let us prove the following.

Lemma 7.8.3. For any f-invariant ergodic probability measure τ on X and for τ-a.e. $x \in X$ we have

$$d_{\mu_\phi}(x) = \frac{\int \phi d\tau}{-\int \log|f'|d\tau}.$$

Proof. Using formula (7.1.0) in Lemma 7.1.1 and Birkhoff 's Ergodic Theorem, we get

$$d_{\mu_\phi}(x) = \lim_{n\to\infty} \frac{S_n\phi(x)}{\log|(f^n)'(x)|^{-1}} = \frac{\lim_{n\to\infty}\frac{1}{n}S_n\phi(x)}{\lim_{n\to\infty}\frac{1}{n}\log|(f^n)'(x)|^{-1}} = \frac{\int \phi d\tau}{-\int\log|f'|d\tau}.$$

♣

One can conclude from this, that the singular part $\hat{X}$ of X has zero measure for every f-invariant τ. Yet the set $\hat{X}$ is usually big, see Exercise 7.2.4.

On the Legendre transform. Let $k = k(q) : I \to I\!\!R$ be a convex function on $I = [\alpha_1(k), \alpha_2(k)]$ where $-\infty \le \alpha_1(k) \le \alpha_2(k) \le \infty$ (i.e. I is either a point or a closed interval or a semiline or $I\!\!R$). The Legendre transform of k is the function g of a new variable p defined by

$$g(p) = \sup_{q\in I}\{pq - k(q)\}$$

everywhere where a finite supremum exists. It can be easily proved (Exercise 7.2.2) that the domain of g is also either a point, or a closed interval or a semiline or $I\!\!R$. It is also easy to show that g is convex and that the Legendre transform is involutive. We then say that the functions k and g form a Legendre transform pair.

Proposition 7.8.4. If two convex functions k and g form a Legendre transform pair then $g(k'(q)) = qk'(q) - k(q)$, where $k'(q)$ is any number between the left and right hand side derivative of k at q, which are defined as $-\infty, \infty$ at $\alpha_1(k)$, $\alpha_2(k)$ respectively, if these end points are finite. We set $0 \cdot \pm\infty = 0$ in case $k' = \pm\infty$ at $q = \alpha_i(k) = 0$. If

$\alpha_2(k) = \infty$ (similarly if $\alpha_1(k) = -\infty$), then for $k'(\infty)$ defined as $\lim_{q\to\infty} k'(q)$, it holds $g(k'(\infty)) = \lim_{q\to\infty} g(k'(q))$.

Note that if k is C^2 with $k'' > 0$, therefore strictly convex, then also $g'' > 0$ at all points $k'(q)$ for $\alpha_1(k) < q < \alpha_2(k)$, therefore g is strictly convex on $[k'(\alpha_1(k)), k'(\alpha_2(k))]$. Outside this interval g is affine in its domain. If the domain of k is one point then g is affine on $I\!R$ and *vice versa.*
We are now in position to formulate our main theorem in this section gathering in particular some facts already proven.

Theorem 7.8.5.
(a) The pointwise dimension $d_{\mu_\phi}(x)$ exists for μ_ϕ-almost every $x \in X$ and

$$d_{\mu_\phi}(x) = \frac{\int \phi d\mu_\phi}{-\int \log|f'|d\mu_\phi} = \mathrm{HD}(\mu_\phi) = \mathrm{PD}(\mu_\phi).$$

(b) The function $q \mapsto T(q)$ for $q \in I\!R$, is real analytic, $T(0) = \mathrm{HD}(X)$, $T(1) = 0$, $T'(q) = \frac{\int \phi d\mu_q}{\int \log|f'|d\mu_q} < 0$ and $T''(q) \geq 0$.
(c) For all $q \in I\!R$ we have $\mu_q(X_{-T'(q)}) = 1$, where μ_q is the invariant Gibbs measure for the potential $\phi_{q,T(q)}$, and $\mathrm{HD}(\mu_q) = \mathrm{HD}(X_{-T'(q)})$.
(d) For every $q \in I\!R$, $F(-T'(q)) = T(q) - qT'(q)$, i.e. $p \mapsto -F(-p)$ is Legendre transform of $T(q)$.
If the measures μ_ϕ and $\mu_{-\mathrm{HD}(X)\log|f'|}$ (the latter discussed in Theorem 7.1.5 and Corollary 7.1.6) do not coincide, then $T'' > 0$ and $F'' < 0$, i.e. the functions T and F are respectively strictly convex on $I\!R$, and stricctly concave on $[-T'(\infty), -T'(-\infty)]$ which is a bounded interval in $I\!R^+ = \{\alpha \in I\!R : \alpha > 0\}$. If $\mu_\phi = \mu_{-\mathrm{HD}(X)\log|f'|}$ then T is affine and the domain of F is one point $-T'$.
(e) For every $q \neq 1$ the HP and Rényi spectra exist (i.e. limits in the definitions exist) and $\frac{T(q)}{1-q} = HP_q(\mu_\phi) = R_q(\mu_\phi)$. For $q = 1$ the information dimension $I(\mu_\phi)$ exists and

$$\lim_{q\to 1, q\neq 1} \frac{T(q)}{1-q} = -T'(1) = \mathrm{HD}(\mu_\phi) = \mathrm{PD}(\mu_\phi) = I(\mu_\phi).$$

Insert Figures: graph of T, graph of F [Pesin, p.219, 220]

Proof. 1. Since $\mathrm{P}(\phi) = 0$, the part (a) is an immediate consequence of Lemma 7.1.3 and its second and third equalities follow from Theorem 7.1.4. The first equality is also a special case of Lemma 7.8.3 with $\tau = \mu_\phi$.

2. We shall prove some statements of the part (b). The function $\phi_{q,t} = -t\log|f'| + q\phi$, from $I\!R^2$ to $C^\theta(X)$, where θ is a Hölder exponent of the function ϕ, is affine. Since by a result of Ruelle (see [Ru])??????? the topological pressure function $\mathrm{P} : C^\theta \to I\!R$ is real analytic, then the composition which we denote $\mathrm{P}(q,t)$ is real analytic. Hence the real analyticity of $T(q)$ follows immediately from the Implicit Function Theorem once we verify the non-degeneracy assumption. In fact C^2-smoothness of $\mathrm{P}(q,t)$ is sufficient to proceed further (here only C^1), which has been proved in Theorem 4.7.4. Indeed, due to Theorem 4.6.5 for every $(q_0,t_0) \in I\!\!R^2$

$$\frac{\partial \mathrm{P}(q,t))}{\partial t}|_{(q_0,t_0)} = -\int_X \log|f'|d\mu_{q_0,t_0} < 0, \tag{7.8.2}$$

where μ_{q_0,t_0} is the invariant Gibbs state of the function ϕ_{q_0,t_0}. Differentiating with respect to q the equality $\mathrm{P}(q,t)=0$ we obtain

$$0 = \frac{\partial \mathrm{P}(q,t)}{\partial t}|_{(q,T(q))} \cdot T'(q) + \frac{\partial \mathrm{P}(q,t))}{\partial q}|_{(q,T(q))} \tag{7.8.3}$$

hence we obtain the standard formula

$$T'(q) = -\frac{\partial \mathrm{P}(q,t))}{\partial q}|_{(q,T(q))} \Big/ \frac{\partial \mathrm{P}(q,t)}{\partial t}|_{(q,T(q))},$$

Again using (4.6.5') and $\mathrm{P}(\phi_{q,T(q)})=0$, we obtain

$$T'(q) = \frac{\int \phi d\mu_q}{\int \log|f'|d\mu_q} \le \frac{-\mathrm{h}_{\mu_q}(f)}{\int \log|f'|d\mu_q} < 0, \tag{7.8.4}$$

the latter true since the entropy of any invariant Gibbs measure for Hölder function is positive, see for example Theorem 4.2.7. The equality $T(0)=\mathrm{HD}(X)$ is just Corollary 7.1.2. $T(1)=0$ follows from the equality $\mathrm{P}(\phi)=0$.

3. The inequality $T''(q) \ge 0$ follows from the convexity of $\mathrm{P}(q,t)$, see Theorem 2.5.2. Indeed the assumption that the part of $I\!\!R^3$ above the graph of $\mathrm{P}(q,t)$ is convex implies that its intersection with the plane (q,t) is also convex. Since $\frac{\partial \mathrm{P}(q,t))}{\partial t}|_{(q_0,t_0)} < 0$, this is the part of the plane above the graph of T. Hence T is a convex function. We avoided in the above consideration an explicit computation of T''. However to discuss strict convexity (a part of (d)) it is necessary to compute it. Differentiating (7.8.3) with respect to q we obtain the standard formula

$$T''(q) = \frac{T'(q)^2\frac{\partial^2 \mathrm{P}(q,t)}{\partial t^2} + 2T'(q)\frac{\partial^2 \mathrm{P}(q,t)}{\partial q \partial t} + \frac{\partial^2 \mathrm{P}(q,t)}{\partial q^2}}{-\frac{\partial \mathrm{P}(q,t)}{\partial t}} \tag{7.8.5}$$

with the derivatives of P taken at $(q,T(q))$. The numerator is equal to

$$\Big(T'(q)\frac{\partial}{\partial t} + \frac{\partial}{\partial q}\Big)^2 \mathrm{P}(q,t) = \sigma^2_{\mu_q}(-T'(q)\log|f'| + \phi)$$

by Theorem 4.7.4, since this is the second order derivative of $P: C(X) \to I\!\!R$ in the direction of the function $-T'(q)\log|f'| + \phi$.

The inequality $\sigma^2 \geq 0$, true by definition, implies $T'' \geq 0$ since the denominator in (7.8.5) is positive by (7.8.2).

By Theorem 1.11.3 $\sigma^2_{\mu_q}(-T'(q)\log|f'| + \phi) = 0$ if and only if the function $-T'(q)\log|f'|+\phi$ is cohomologous to a constant, say to a. It follows then from the equality in (7.8.4) that $a = \int a\,d\mu_q = \int(-T'(q)\log|f'| + \phi)d\mu_q = 0$. Therefore $T'(q)\log|f'|$ is cohomologous to ϕ and, as $\mathrm{P}(\phi) = 0$, also $\mathrm{P}(T'(q)\log|f'|) = 0$. Thus, by Theorem 7.1.5 and Corollary 7.1.2, $T'(q) = -\mathrm{HD}(X)$ and consequently ϕ is cohomologous to the function $-\mathrm{HD}(X)\log|f'|$. This implies that $\mu_\phi = \mu_{-\mathrm{HD}(X)\log|f'|}$, the latter being the equilibrium (invariant Gibbs) state of the potential $-\mathrm{HD}(X)\log|f'|$. Therefore, in view of our formula for T'', if $\mu_\phi \neq \mu_{-\mathrm{HD}(X)\log|f'|}$, then $T''(q) > 0$ for all $q \in I\!R$.

4. We prove (c). By Lemma 7.8.3 applied to $\tau = \mu_q$, there exists a set $\tilde{X}_q \subset X$, of full measure μ_q, such that for every $x \in \tilde{X}_q$ there exists

$$d_{\mu_\phi}(x) = \lim_{r\to 0}\frac{\log\mu_\phi(B(x,r))}{\log r} = \frac{\int\phi d\mu_q}{-\int\log|f'|d\mu_q} = -T'(q).$$

the latter proved in (b). Hence $\tilde{X}_q \subset X_{-T'(q)}$. Therefore $\mu_q(X_{-T'(q)}) = 1$. By Lemma 7.8.2 for every $B = B(x,r)$

$$|\log\mu_q(B) - T(q)\log r - q\log\mu_\phi(B)| < C$$

for some constant $C \in I\!R$. Hence

$$\left|\frac{\log\mu_q(B)}{\log r} - T(q) - q\frac{\log\mu_\phi(B)}{\log r}\right| \to 0 \tag{7.8.6}$$

as $r \to 0$. Using (7.8.6), observe that for every $x \in X_{-T'(q)}$, in particular for every $x \in \tilde{X}_q$,

$$\lim_{r\to 0}\frac{\log\mu_q(B)}{\log r} = T(q) + q\lim_{r\to 0}\frac{\log\mu_\phi(B)}{\log r} = T(q) - qT'(q).$$

Although $\tilde{X}_q$ can be much smaller than $X_{-T'(q)}$, miraculously their Hausdorff dimensions coincide. Indeed the measure μ_q restricted to either $\tilde{X}_q$ or to $X_{-T'(q)}$ satisfies the assumptions of Theorem 6.6.3 with $\theta_1 = \theta_2 = T(q) - qT'(q)$. Therefore

$$\mathrm{HD}(\tilde{X}_q) = \mathrm{HD}(X_{-T'(q)}) = T(q) - qT'(q) \tag{7.8.7}$$

and consequently

$$F(-T'(q)) = T(q) - qT'(q).$$

Remarks. (a) If we take a set larger than $X_{-T'(q)}$, namely replacing in the definition of $X_\nu(\alpha)$ the dimension d_ν by the lower dimension $\underline{d}_\nu$ we still obtain the same Hausdorff dimension, again by Theorem 6.6.3.

(b) Some authors replace in the definition of $X_\nu(\alpha)$ the value $d_\nu(x)$ by $\underline{d}_\nu(x)$. Then there is no singular part. In view of a) the $F_\nu(\alpha)$ spectrum is the same for $\nu = \mu_\phi$.

(c) Notice that (7.8.7) means that $\mathrm{HD}(X_{-T'(q)})$ is the value where the straight line tangent to the graph of T at $(q, T(q))$ intersects the range axis.

In the next steps of the proof the following will be useful.

Claim. (Variational Principle fo T.) For any f-invariant ergodic probability measure τ on X, consider the following linear equation of variables q, t

$$\int \phi_{q,t} d\tau + \mathrm{h}_\tau(f) = 0$$

that is

$$t = t_\tau(q) = \frac{\mathrm{h}_\tau(f)}{\int \log|f'|d\tau} + q\frac{\int \phi d\tau}{\int \log|f'|d\tau}. \tag{7.8.8}$$

Then for every $q \in I\!\!R$

$$T(q) = \sup_\tau\{t_\tau(q)\} = t_{\mu_q}(q),$$

Where the supremum is taken over all f-invariant ergodic probability measures τ on X.

Proof of the Claim. Since $\int \phi_{q,t} d\tau + \mathrm{h}_\tau(f) \le \mathrm{P}(\phi_{q,t})$ and since $\frac{\partial \mathrm{P}(q,t)}{\partial t} < 0$ (compare the proof of convexity of T), we obtain

$$t_\tau(q) \le T(q).$$

On the other hand by (7.8.8), and using $\mathrm{P}(\phi_{q,T(q)} = 0$, we obtain

$$t_{\mu_q}(q) = \frac{\mathrm{h}_{\mu_q}(f) + q\int \phi d\mu_q}{\int \log|f'|d\mu_q} = \frac{T(q)\int \log|f'|d\mu_q}{\int \log|f'|d\mu_q} = T(q).$$

The Claim is proved. ♣

5. We continue Proof of Theorem 7.8.5. We shall prove the missing parts of (d). We have already proved ithat

$$F(-T'(q)) = \mathrm{HD}(X_{-T'(q)}) = \mathrm{HD}(\mu_q) = T(q) - qT'(q).$$

Note that $[-T'(\infty), -T'(-\infty)] \subset I\!\!R^+ \cup \{0, \infty\}$ since $T'(q) < 0$ for all q. Note finally that

$$-T'(-\infty) = \lim_{q\to-\infty} \frac{-\int \phi d\mu_q}{\int \log|f'|d\mu_q} \le \frac{\sup(-\phi)}{\inf \log|f'|} < \infty$$

and

$$-T'(\infty) = \lim_{q\to\infty} \frac{-\int \phi d\mu_q}{\int \log|f'|d\mu_q}.$$

The expressions under lim are positive, (see (7.8.4)). It is enough now to prove that they are bounded away from 0 as $q \to \infty$. To this end choose q_0 such that $T(q_0) < 0$. By our Claim (Variational Principle for T) $t_{\mu_q}(q_0) \le T(q_0)$. Since $t_{\mu_q}(0) \ge 0$, we get

$$-q_0 \frac{\int \phi d\mu_q}{\int \log|f'| d\mu_q} = t_{\mu_q}(0) - t_{\mu_q}(q_0) \ge |T(q_0)|.$$

Hence $\frac{-\int \phi d\mu_q}{\int \log|f'| d\mu_q} \ge |T(q_0)|/q_0 > 0$ for all q.

6. To end the proof of (d) we need to prove the formula for F at $-T'(\pm\infty)$ (in case T is not affine) and to prove that for $\alpha \notin [-T'(\infty), -T'(-\infty)]$ the sets $X_{\mu_\phi}(\alpha)$ are empty. First note the following.

6a. For any f-invariant ergodic probability measure τ on X, there exists $q \in I\!R \cup \{\pm\infty\}$ such that

$$\frac{\int \phi d\tau}{\int \log|f'| d\tau} = \frac{\int \phi d\mu_q}{\int \log|f'| d\mu_q}. \tag{7.8.9}$$

($\lim_{q\to\pm\infty}$ in the $\pm\infty$ case).

Indeed, by the Claim the graphs of the functions $t_\tau(q)$ and $T(q)$ do not intersect transversally (they can be only tangent) and hence the first graph which is a straight line, is parallel to a tangent to the graph of T at a point $(q_0, T(q_0)$, or one of its asymptots, at $-\infty$ or $+\infty$. Now (7.8.9) follows from the same formula (7.8.8) for $\tau = \mu_{q_0}$, since the graph of $t_{\mu_{q_0}}$ is tangent to the graph of T just at $(q_0, T(q_0))$. (Note that the latter sentence proves the formula $T'(q) = \frac{\int \phi d\mu_q}{\int \log|f'| d\mu_q}$ in a different way than in 2, namely *via* Variational Principle for T.).

6b. Proof that $X_\alpha = \emptyset$ for $\alpha \notin [-T'(\infty), -T'(-\infty)]$. Suppose there exists $x \in X$ with $\alpha := d_{\mu_\phi}(x) \notin [-T'(\infty), -T'(-\infty)]$. Consider any sequence of integers $n_k \to \infty$ and real numbers b_1, b_2 such that

$$\lim_{k\to\infty} \frac{1}{n_k} S_n\phi(x) = b_1, \quad \lim_{k\to\infty} \frac{1}{n_k}(-\log|(f^n)'(x)|) = b_2$$

and $b_1/b_2 = \alpha$. Let τ be any weak*-limit of the sequence of measures

$$\tau_{n_k} := \frac{1}{n_k} \sum_{j=0}^{n_k-1} \delta_{f^j(x)},$$

where $\delta_{f^j(x)}$ is the Dirac measure supported at $f^j(x)$, compare Remark 2.1.14a. Then $\int \phi d\tau = b_1$ and $\int (-\log|f'|) d\tau = b_2$. Due to Choquet Theorem (Section 2 1) (or due to the Decomposition into Ergodic Components Theorem, Theorem 1.8.8) we can assume that τ is ergodic. Indeed, τ is an "average" of ergodic measures. So among all ergodic measure ν involved in the average, there is ν_1 such that $\frac{\int \phi d\nu_1}{\int -\log|f'|d\nu_1} \le \alpha$ and ν_2 such that $\frac{\int \phi d\nu_2}{\int -\log|f'|d\nu_2} \ge \alpha$. If $\alpha < -T'(\infty)$ we consider ν_1 as an ergodic τ, if $\alpha > -T'(-\infty)$ we consider ν_2. For the ergodic τ found in this way, the limit α can be different than for the original τ, but it will not belong to $[-T'(\infty), -T'(-\infty)]$ and we shall use the same symbol α to denote it. By Birkhoff's Ergodic Theorem applied to the functions ϕ and $\log|f'|$, for τ-a.e. x we have $\lim_{n\to\infty} \frac{S_n(\phi)(x)}{-\log|(f^n)'(x)|} = \alpha$. Hence, applying Lemma 7.8.3, we get

$$\alpha = d_{\mu_\phi}(x) = \frac{\int \phi d\tau}{-\int \log|f'|d\tau}.$$

Finally notice that by (7.8.9) there exists $q \in I\!\!R$ such that $\alpha = \frac{\int \phi d\mu_q}{-\int \log|f'|d\mu_q}$, whence $\alpha \in [-T'(\infty), -T'(-\infty)]$. This contradiction finishes the proof. ♣

Remark. We have proved in fact that for all $x \in X$ any limit number of the quotien $\log \mu_\phi(B(x,r)/\log r$ as $r \to 0$ lies in $[-T'(\infty), -T'(-\infty)]$, the fact stronger than $d_{\mu_\phi}(x) \in [-T'(\infty), -T'(-\infty)]$ for all x in the regular part of X.

6c. $F(-T'(\pm\infty)) = \mathbf{HD}\big(X_{-T'(\pm\infty)}\big)$**.** Consider any τ being a weak*-limit of a subsequence of μ_q as q tends to, say, ∞. We shall try to proceed with τ similarly as we did with μ_q, though we shall meet some difficulties. We do not know whether τ is ergodic (and choosing an ergodic one from the ergodic decomposition we may loose the convergence $\mu_q \to \tau$). Nevertheless using Birkhoff Ergodic Theorem and proceeding as in the proof of Lemma 7.8.3, we get

$$\begin{aligned}\frac{\int \lim_{n\to\infty} \frac{1}{n} S_n\phi(x)\, d\tau(x)}{-\int \lim_{n\to\infty} \frac{1}{n} \log|(f^n)'(x)|\, d\tau(x)} &= \frac{\int \phi\, d\tau}{-\int \log|f'|\, d\tau} = \lim_{q\to\infty} \frac{\int \phi d\mu_q}{-\int \log|f'|d\mu_q} \\ &= \lim_{q\to\infty} (-T'(q)) = -T'(\infty)\end{aligned}$$

with the convergence over a subsequence of q's. Since we know already that

$$d_{\mu_\phi}(x) = \frac{\lim_{n\to\infty} \frac{1}{n} S_n\phi(x)}{-\lim_{n\to\infty} \frac{1}{n} \log|(f^n)'(x)|} \ge -T'(\infty),$$

we obtain for every x in a set $\tilde{X}_\tau$ of full measure τ that the limit $d_\tau(x) = -T'(\infty)$. We conclude with $\tilde{X}_\tau \subset X_{-T'(\infty)}$. Now we use the Volume Lemma for the measure τ. There is no reason for it to be Gibbs, neither ergodic, so we need to refer to the version of Volume Lemma coming from Theorem 7.1.10. We obtain

$$\begin{aligned}\mathrm{HD}(X_{-T'(\infty)}) \ge \mathrm{HD}^*(\tau) &\ge \frac{\mathrm{h}_\tau(f)}{\int \log|f'|\, d\tau} \ge \liminf_{q\to\infty} \frac{\mathrm{h}_{\mu_q}(f)}{\chi_{\mu_q}(f)} \\ &= \lim_{q\to\infty} T(q) - qT'(q) = F(-T'(\infty)).\end{aligned}$$

We have used here the upper semicontinuity of the entropy function $\nu \to \mathrm{h}_\nu(f)$ at τ due to the expanding property (see Theorem 2.4.6). It is only left to estimate $\mathrm{HD}(X_{-T'(\infty)})$ from above. As for μ_q, we have for every q and $x \in X_{-T'(\infty)}$ (see (7.8.6)) that

$$\lim_{r\to 0} \frac{\log \mu_q(B)}{\log r} = T(q) + q \lim_{r\to 0} \frac{\log \mu_\phi(B)}{\log r} = T(q) - qT'(\infty) \leq T(q) - qT'(q),$$

Hence $\mathrm{HD}(X_{-T'(\infty)}) \leq T(q) - qT'(q)$. Letting $q \to \infty$ we obtain $\mathrm{HD}(X_{-T'(\infty)}) \leq F(-T'(\infty))$.

7. HP and Rényi spectra. Recall that topological supports of μ_ϕ and μ_q are equal to X, since these measures as Gibbs states for Hölder functions, do not vanish on open subsets of X due to Proposition 4.2.5. For every $\mathcal{G}_r$ a finite or countable covering X by balls of radius r of multiplicity at most C we have

$$1 \leq \sum_{B\in\mathcal{G}_r} \mu_q(B) \leq C.$$

Hence, by Lemma 7.8.2

$$C^{-1} \leq r^{T(q)} \sum_{B\in\mathcal{G}_r} \mu_\phi(B)^q \leq C$$

with an appropriate another constant C. Taking logarithms and, for $q \neq 1$, dividing by $(1-q)\log r$ yields (e) for $q \neq 1$.

8. Information dimension. For $q = 1$ we have $\lim_{q\to 1, q\neq 1} \frac{T(q)}{1-q} = -T'(1)$ by the definition of derivative. It is equal to $\mathrm{HD}(\mu_\phi) = \mathrm{PD}(\mu_\phi)$ by (a) and (b) and equal to $I(\mu_\phi)$ by Exercise 7.8.5. ♣

Exercises

7.2.1. Prove the equalities of Rényi and Hentschel-Procaccia spectra.

7.2.2. Prove Proposition 7.8.4 about Legendre transform pairs and remarks preceding and following it.

7.2.3. Prove for $\alpha = -T'(1)$ that $F(\alpha) = \alpha$ and $F'(\alpha) = 1$ (see Fig.1.) and $F'(-T'(\pm\infty)) = \pm\infty$.

7.2.4. Prove that if ϕ is not cohomologous to $-\mathrm{HD}(X)\log|f'|$ then the singular part $\hat{X}$ of X is nonempty. Moreover $\mathrm{HD}(\hat{X}) = \mathrm{HD}(X)$.

Hint: Using the Shadowing Lemma from Chapter 3, find trajectories that have blocks close to blocks of trajectories typical for $\mu_{-\mathrm{HD}(X)\log|f'|}$ of length N interchanging with blocks close to blocks typical for μ_ϕ of length εN, for N arbitrarily large and $\varepsilon > 0$ arbitrarily small.

7.2.5. Define the *lower* and *upper information dimension* $\underline{I}(\nu)$ and $\overline{I}(\nu)$ replacing in the definition of $I(\nu)$ the limit $\lim_r$ by the lower and upper limits respectively. Prove that $\mathrm{HD}_*(\nu) \leq \underline{I}(\nu) \leq \overline{I}(\nu) \leq \mathrm{PD}^*(\nu)$, see (7.8.0).

Sketch of the proof. For an arbitrary $\varepsilon > 0$ there exist $C > 0$ and $A \subset X$, with $\nu(X \setminus A) \leq \varepsilon$ such that for all r small enough there exists a partition $\mathcal{F}_r$ of A, satisfying

$H_\nu(r)+\varepsilon \ \ge\ -\sum_{B\in\mathcal{F}_r}\nu(B)\log\nu(B)\ \ge\ \sum_{B\in\mathcal{F}_r}\nu(B)\mathrm{HD}_*(\nu)\log\frac{1}{C\mathrm{diam}B}\ \ge\ \mathrm{HD}_*(\nu)(1-\varepsilon)\log\frac{1}{Cr}$.

On the other hand for the partition $\mathcal{B}_r$ of X into intersections with boxes (cubes) of sides of length r (compare Proposition 6.4.6 and the partition involved in the definition of Rényi dimension, but consider here disjoint cubes, that is open from one side), we have

$$\overline{I}(\nu)=\limsup_{r\to 0}\frac{H_\nu(r)}{-\log r}\le\limsup_{r\to 0}\frac{-\sum_{B\in\mathcal{B}_r}\nu(B)\log\nu(B)}{-\log r}\le\limsup_{r\to 0}\frac{\int\log\nu(B_r(x))\,d\nu(x)}{\log r}$$
$$\le\int\Big(\limsup_{r\to 0}\frac{\log\nu(B_r(x))}{\log r}\Big)d\nu(x)\le\mathrm{PD}^*(\nu),$$

where $B_r(x)$ denotes the cube of side r containing x.
Prove that it has been eligible here to use cubes instead of balls standing in the definition of $\overline{d}_\nu(x)$. For this aim prove that for ν-a.e. $x\in X$, we have $\lim\frac{\log\nu(B_r(x))}{\log\nu(B(x,r))}=1$. Use Borel-Cantelli lemma.
Prove that we could use Fatou's lemma (changing the order of limsup and integral) indeed due to the existence of a ν-integrable function which bounds from above all the functions $\log\nu(B(x,r))/\log r$ (or $\log\nu(B_r(x))/\log r$). Use again Borel-Cantelli lemma, for, say, $r=2^{-k}$.

7.2.6. Let $\mu=\mu_\phi$ be a measure of maximal entropy on a topologically exact conformal expanding repeller X such that every point $x\in X$ has exactly d preimages (so $\phi=-\log d$). Prove (deduce from Theorem 7.8.5) that $F(\alpha)=\sup_{t\in\mathbb{R}}\Big(t+\frac{\alpha\mathrm{P}(t)}{\log d}\Big)$, more concretely $F(\alpha)=T+\frac{\alpha\mathrm{P}(T)}{\log d}$, where $\alpha=-\frac{\log d}{P'(T)}$.

7.2.7 Let $\phi_i:X\to\mathbb{R}$ be Hölder continuous functions for $i=1,...,k$ and μ_{ϕ_i} their Gibbs measures. Define $X_{\alpha_1,\dots,\alpha_k}=\{x\in X: d_{\mu_i}(x)=\alpha_i \text{ for all } i=1,...,k\}$. Define $\phi_{q_1,\dots,q_k,t}=-t\log|f'|+\sum_i q_i\phi_i$ and $T(q_1,...,q_k)$ as the only zero of the function $t\mapsto\mathrm{P}(\phi_{q_1,\dots,q_k,t})$. Prove the same properties of T as in Theorem 7.8.5, in particular that

$$\mathrm{HD}(X_{\alpha_1,\dots,\alpha_k})=\inf_{(q_1,\dots,q_k)\in\mathbb{R}^k}\sum_i q_i\alpha_1+T(q_1,...,q_k)$$

wherever the infimum is finite.

Historical and bibliographical notes

The section on multifractal analysis relies mainly on the monographs by Y. Pesin [P] and K. Falconer [F3] (though details are modified, for example we do not use Markov partition). The reader can find there comprehensive expositions and further references. The development of this theory has been stimulated by physicists, the paper often quoted is [HJKPS].

Bibliography

[F3] K. Falconer "Techniques in Fractal Geometry". John Wiley and Sons, Chichester, 1997.

[HJKPS] T.C. Hasley, M. Jensen, L. Kadanoff, I. Procaccia, B. Shraiman "Fractal measures and their singularities: the characterization of strange sets". Phys. Rev. A 33.2 (1986), 1141-1151.

[P] Y.B. Pesin "Dimension Theory in Dynamical Systems". The University of Chicago Press, Chicago and London, 1997.

Section 2. Fluctuations for Gibbs measures.

Theorem 7.2.1. Let $f : X \to X$ be a holomorphic expanding repeller. Let ϕ be a Hölder continuous function and let μ_ϕ be it's Gibbs measure. Then, with $\kappa = \mathrm{HD}(\mu_\phi)$, either

(a) $\kappa = \mathrm{HD}(X)$ and $\mu_\phi \asymp \mathrm{H}_\kappa$ on X (if $\psi = \phi + \kappa \log |f'| - \mathrm{P}(\phi)$ is a coboundary) or

(b) $\mu_\phi \perp \mathrm{H}_\kappa$ (if $\psi = \phi + \kappa \log |f'| - \mathrm{P}(\phi)$ is not a coboundary) and moreover, there exists $c_0 > 0$, ($c_0 = \sqrt{2\sigma^2(\psi)/\chi_{\mu_\phi}(f)}$), such that with the gauge function $\phi_c(r) = r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})$, where $\log_3$ means the iteration of the log function 3 times:

(c) $\mu_\phi \perp \mathrm{H}_{\phi_c}$ for all $0 < c < c_0$, and

(d) $\mu_\phi \ll \mathrm{H}_{\phi_c}$ for all $c > c_0$.

Proof. If ψ is a coboundary, then it follows from equality $\phi - \psi = -\kappa \log |f'| + \mathrm{P}(\phi)$ that $\mathrm{P}(\phi) = \mathrm{P}(\phi - \psi) = \mathrm{P}(\phi) + \mathrm{P}(-\kappa \log |f'|)$. Thus $\mathrm{P}(-\kappa \log |f'|) = 0$ and the part (a) follows immediately from Theorem 7.1.1 and the observation saying that the potential cohomological up to an additive constant have the same Gibbs states.

Suppose now that ψ is not a coboundary. As in the previous section let $\lambda = \inf |f'| > 1$. Recall that then there exists $\tau > 0$ small enough that firstly, $f\ |_A$ is one - to - one for all sets $A \subset X$ with $\mathrm{diam}(A) \le \tau$ and secondly, $d(f(y), f(z)) \ge \lambda d(y,z)$ if $d(y,z) < \tau$. Fix $x \in X$ and $r > 0$ such that $r < \tau$ Define $n = n(x,r)$ to be the least number such that $\mathrm{diam}(f^n(B(x,r))) \ge \tau$. Since $f : X \to X$ is topologically exact, $n(x,r)$ is finite. By definition of n, $\mathrm{diam}(f^j(B(x,r))) < \tau$ for all $j = 0, 1, ...n-1$. Therefore $f^n\ |_{(B(x,r))}$ is one - to - one and

$$m_\phi(f^n(B(x,r))) = \int_{B(x,r)} \exp(P(\phi)n - S_n\phi(z)) dm_\phi(z),$$

where m_ϕ is the fixed point of the dual operator of $\mathcal{L}_\phi$, the transfer operator associated with the function $\phi - \mathrm{P}(\phi)$. Since the function ϕ is Hölder continuous, it follows from ??? that there exists a constant $K > 0$ such that $|S_n\phi(z) - S_n\phi(x)| \le K$ for all $x \in X$, $z \in B(x,r)$, and $n = n(x,r)$. Hence, we get

$$\begin{aligned} e^{-K} \exp(P(\phi)n - S_n\phi(x)) m_\phi(B(x,r)) &\le m_\phi(f^n(B(x,r))) \\ &\le e^K \exp(P(\phi)n - S_n\phi(x)) m_\phi(B(x,r)) \end{aligned}$$

or equivalently

$$\begin{aligned} e^{-K} \exp(S_n\phi(x) - P(\phi)n) m_\phi(f^n(B(x,r))) &\le m_\phi(B(x,r)) \\ &\le e^K \exp(S_n\phi(x) - P(\phi)n) m_\phi(f^n(B(x,r))). \end{aligned}$$

Similarly, since $\log|f'(z)|$ is a Hölder continuous function, there exists a constant $K_1 > 0$ such that

$$|S_n(\log|f'(z)|) - S_n(\log|f'(x)|)| \le K_1$$

for all $x \in X$, $z \in B(x,r)$, and $n = n(x,r)$. Then

$$\frac{|(f^n)'(z)|}{|(f^n)'(x)|} \le e^{K_1}$$

Therefore

$$\begin{aligned}\operatorname{diam}(B(x,r))e^{-K_1}|(f^n)'(x)| &\le \operatorname{diam}(f^n(B(x,r))) \\ &\le \operatorname{diam}(B(x,r))|(f^n)'(x)|e^{K_1}.\end{aligned}$$

Hence, $\operatorname{diam}(B(x,r))e^{-K_1}|(f^n)'(x)| \le \tau\|f'\|$, or equivalently $r \le \frac{1}{2}e^{K_1}|(f^n)'(x)|^{-1}\tau\|f'\|$. Also, since $\operatorname{diam}(f^n(B(x,r))) \ge \tau$ we get similarly $r \ge \frac{1}{2}e^{-K_1}|(f^n)'(x)|^{-1}\tau$. In conclusion,

(7.2.1) $$\frac{1}{2}e^{-K_1}|(f^n)'(x)|^{-1}\tau \le r \le \frac{1}{2}e^{K_1}|(f^n)'(x)|^{-1}\tau\|f'\|.$$

Hence, denoting $\sqrt{\log(|(f^n)'(x)|)\log_3(|(f^n)'(x)|)}$ by $g_n(x)$, we get

$$\begin{aligned}\frac{m_\phi(B(x,r))}{r^\kappa\exp(c\sqrt{\log 1/r\log_3 1/r})} &\le \frac{e^K\exp(S_n\phi(x) - P(\phi)n)m_\phi(f^n(B(x,r)))}{r^\kappa\exp(c\sqrt{\log 1/r\log_3 1/r})} \\ &\le \frac{e^K\exp(S_n\phi(x) - P(\phi)n)}{(\frac{1}{2}\frac{1}{|(f^n)'(x)|e^{K_1}}\tau)^\kappa\exp(c\sqrt{\log(2e^{-K_1}\frac{|(f^n)'(x)|}{\tau\|f'\|})\log_3(2e^{-K_1}\frac{|(f^n)'(x)|}{\tau\|f'\|})})} \\ &\le \frac{Q\exp(S_n\phi(x) - P(\phi)n)}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))},\end{aligned}$$

where Q is a large enough constant. Similarly,

$$\begin{aligned}&\frac{m_\phi(B(x,r))}{r^\kappa\exp(c\sqrt{\log 1/r\log_3 1/r})} \\ &\ge \frac{e^{-K}\exp(S_n\phi(x) - P(\phi)n)m_\phi(f^n(B(x,r)))}{r^\kappa\exp(c\sqrt{\log 1/r\log_3 1/r})} \\ &\ge \frac{e^{-K}\exp(S_n\phi(x) - P(\phi)n)m_\phi(f^n(B(x,r)))}{(\frac{1}{2}|(f^n)'(x)|^{-1}e^{K_1}\tau\|f'\|)^\kappa\exp(c\sqrt{\log(2e^{K_1}|(f^n)'(x)|\tau^{-1})\log_3(2e^{K_1}|(f^n)'(x)|\tau^{-1})})} \\ &\ge \frac{Q_1\exp(S_n\phi(x) - P(\phi)n)m_\phi(f^n(B(x,r)))}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))} \quad (Q_1 \text{ large enough}) \\ &\ge \frac{Q_2\exp(S_n\phi(x) - p(\phi)n)}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))},\end{aligned}$$

where $Q_2 = Q_1\min\{m_\phi(f^n(B(x,r))) : x \in X, r > 0\}$. Note that, since the topological support of μ_ϕ and m_ϕ is equal to X, using the the Bounded Distortion Theorem we get $Q_2 > 0$ because $f^n(B(x,r)) \supset B(f^n(x), r|(f^n)'(x)|K^{-1}) \supset B(f^n(x), R)$, where $R = K^{-1}\tau e^{-K_1}$. Finally, we get

$$0 < \frac{Q_2 \exp(S_n\phi(x) - P(\phi)n)}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))} \leq \frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})} \leq \frac{Q \exp(S_n\phi(x) - P(\phi)n)}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))}$$

Hence,

$$\begin{aligned}\left[\frac{Q_2 \exp(S_n\phi(x) - P(\phi)n)}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))}\right] &\leq \log\left[\frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})}\right]. \\ &\leq \log\left[\frac{Q \exp(S_n\phi(x) - P(\phi)n)}{|(f^n)'(x)|^{-\kappa}\exp(cg_n(x))}\right]\end{aligned}$$

Therefore,

$$\begin{aligned}&\log Q_2 + S_n\phi(x) - P(\phi)n + \kappa \log|(f^n)'(x)| - cg_n(x) \\ &\quad\leq \log\left[\frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})}\right] \\ &\quad\leq \log Q + S_n\phi(x) - P(\phi)n + \kappa \log|(f^n)'(x)| - cg_n(x),\end{aligned}$$

or equivalently,

$$\begin{aligned}&\log Q_2 + g_n(x)\left[\frac{S_n\phi(x) - P(\phi)n + \kappa\log|(f^n)'(x)|}{g_n(x)} - c\right] \\ &\quad\leq \log\left[\frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})}\right] \\ &\quad\leq \log Q + g_n(x)\left[\frac{S_n\phi(x) - P(\phi)n + \kappa\log|(f^n)'(x)|}{g_n(x)} - c\right]. \qquad (7.2.1)\end{aligned}$$

So, in order to proceed further, we are tempted to evaluate the following upper limit.

$$\limsup_{n\to\infty} \frac{S_n\psi(x)}{\sqrt{\log|(f^n)'(x)| \log_3|(f^n)'(x)|}}$$

First, by the Birkhoff ergodic theorem, for every $\epsilon > 0$ there exists $X_1 \subset X$ such that $m_\phi(X_1) = 1$ and forall $x \in X_1$ there exists $N > 0$ such that $\forall n \geq N$

$$n(\chi_\mu - \epsilon) \leq \log|(f^n)'(x)| \leq n(\chi_\mu + \epsilon).$$

Second, since ψ is not a coboundary, by the Law of Iterated Logarithms (see ???), there exists $X_2 \subset X$ such that $m_\phi(X_2) = 1$, and forall $x \in X_2$

$$\limsup_{n\to\infty} \frac{S_n\psi(x)}{\sqrt{n\log_2(n)}} = \sqrt{2\sigma^2}.$$

Therefore, for all $x \in X_1 \bigcap X_2$

$$\begin{aligned}\limsup_{n\to\infty} \frac{S_n\psi(x)}{g_n(x)} &\le \limsup_{n\to\infty} \frac{S_n\psi(x)}{\sqrt{n(\chi_\mu - \epsilon)\log_2(n(\chi_\mu - \epsilon))}} \\ &\le \limsup_{n\to\infty} \frac{S_n\psi(x)}{\sqrt{(\chi_\mu - 2\epsilon)}\sqrt{n\log_2(n)})} \\ &= \frac{\sqrt{2\sigma^2}}{\sqrt{(\chi_\mu - 2\epsilon)}}.\end{aligned}$$

and

$$\begin{aligned}\limsup_{n\to\infty} \frac{S_n\psi(x)}{g_n(x)} &\ge \limsup_{n\to\infty} \frac{S_n\psi(x)}{\sqrt{n(\chi_\mu + \epsilon)\log_2(n(\chi_\mu + \epsilon))}} \\ &\ge \limsup_{n\to\infty} \frac{S_n\psi(x)}{\sqrt{(\chi_\mu + 2\epsilon)}\sqrt{n\log_2(n)})} \\ &= \frac{\sqrt{2\sigma^2}}{\sqrt{(\chi_\mu + 2\epsilon)}}.\end{aligned}$$

In conclusion, forall $x \in X_1 \bigcap X_2$, we have

$$\frac{\sqrt{2\sigma^2}}{\sqrt{(\chi_\mu + 2\epsilon)}} \le \limsup_{n\to\infty} \frac{S_n\psi(x)}{g_n(x)} \le \frac{\sqrt{2\sigma^2}}{\sqrt{(\chi_\mu - 2\epsilon)}}$$

Hence, letting $\epsilon \searrow 0$,

$$\limsup_{n\to\infty} \frac{S_n\psi(x)}{g_n(x)} = \sqrt{\frac{2\sigma^2}{\chi_\mu}}.$$

Thus, with

$$\beta_n(c) = g_n(x)\left[\frac{S_n\phi(x) - P(\phi)n + \kappa\log|(f^n)'(x)|}{g_n(x)} - c\right],$$

we obtain

$$\limsup_{n\to\infty} \beta_n(c) = \begin{cases} \infty & \text{if } \sqrt{\frac{2\sigma^2}{\chi_\mu}} > c \\ -\infty & \text{if } \sqrt{\frac{2\sigma^2}{\chi_\mu}} < c \end{cases} \tag{7.2.3}$$

since, $\lim_{n\to\infty} g_n(x) = \infty$, and this is because f is expanding. By (7.2.1)

$$\begin{aligned}\limsup_{n\to\infty}[\log Q_2 + \beta_n(c)] &\le \limsup_{r\to 0} \log\left[\frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})}\right] \\ &\le \limsup_{n\to\infty}[\log Q + \beta_n(c)].\end{aligned}$$

By (7.2.3), it gives

$$\limsup_{r\to 0} \log\left[\frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})}\right] = \begin{cases} \infty & \text{if } \sqrt{\frac{2\sigma^2}{\chi_\mu}} > c \\ -\infty & \text{if } \sqrt{\frac{2\sigma^2}{\chi_\mu}} < c \end{cases}.$$

In other words,

$$\limsup_{r\to 0} \frac{m_\phi(B(x,r))}{r^\kappa \exp(c\sqrt{\log 1/r \log_3 1/r})} = \begin{cases} \infty & \text{if } \sqrt{\frac{2\sigma^2}{\chi_\mu}} > c \\ 0 & \text{if } \sqrt{\frac{2\sigma^2}{\chi_\mu}} < c \end{cases}.$$

Therefore, by Theorem 5.??, $\mu_\phi \perp \mathrm{H}_{\phi_c}$ for all $c < \sqrt{\frac{2\sigma^2}{\chi_\mu}}$ and $\mu_\phi \ll \mathrm{H}_{g_c}$ for all $c > \sqrt{\frac{2\sigma^2}{\chi_\mu}}$. The proof is finished. ♣

Note that this proof is done without the use of Markov partitions. Note also that the last display in the proof of Theorem 7.2.1 is known as a refined volume lemma.

Section 7.3. Radial behaviour of the Riemann map,I.

In this section $f: X \to X$ continues to be a conformal expanding repeller and we assume additionally that X is a Jordan curve. We then call f a boundary expending conformal repeller. Let $\Omega \subset \overline{\mathbb{C}}$ be one of the components of $\overline{\mathbb{C}} \setminus X$ and fix $z_0 \in \Omega$. Let $D^1 = \{z : |z| < 1\}$. In view of Caratheodory's theorem, any Riemann map $R : D^1 \to \Omega$ (conformal homeomorphism) sending zero to z_0 (which is unique up to rotation) extends homeomorphically to $\overline{D^1}$. For more information about a Riemann map we refer the reader to [?]. We also assume that there exists an open topological annulus $A \subset \overline{\mathbb{C}}$ surronding $\partial\Omega$ such that $f(A \cap \Omega) \subset \Omega$. Since f is a local homeomorphism, it is easy to see that $f^{-1}(A\cap\Omega) \subset \Omega$. With all these assumptions we speak about the expanding map $f : \partial\Omega \to \partial\Omega$ as a conformal boundary repeller with the Jordan domain Ω. This enables us to lift the map f to the closed topological annulus $B = R^{-1}(A\cap\Omega) \subset \overline{D^1}$ one of whose boundary components is the circle S^1 by setting

$$g = R^{-1} \circ f \circ R : \overline{B} \to \overline{D^1}.$$

Obviously $g(B) \subset \overline{D^1}$. Denoting by $I : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ the inversion with respect to the unit circle S^1 and applying the Schwarz reflection principle we see see that g extends analytically to the topological annulus $F = B \cup I(B)$. Our first aim is to show that g is a boundary expending conformal repeller. We begin with the following.

Lemma 7.3.1. There exists $\delta > 0$ such that for every $z \in S^1$ there exists a unique holomorphic inverse branch $g_z^{-n} : B(g^n(z), 2\delta) \to F$ sending $g^n(z)$ to z. If n is large enough (independent of z), then $|(g_z^{-n})'(w)| \le 1/2$ for all $w \in B(g^n(z), \delta)$. In particular the map $g : S^1 \to S^1$ is expanding.

Proof. Fix $\delta > 0$ so small that $B(S^1, 2\delta) \subset F$. Fix $z \in S^1$ and consider the ball $B(g^n(z), 2\delta)$. Define the inverse branch $g_z^{-n} : \overline{D^1} \cap B(g^n(z), 2\delta) \to \mathbb{C}$ of g^n by putting $g_z^{-n}(w) = R^{-1} \circ f_{R(z)}^{-n} \circ R(w)$, where $f_{R(z)}^{-n}$ is a local holomorphic inverse branch of f^n defined throughout a neighbourhood of $f^n(R(z))$ sending $f^n(R(z))$ to $R(z)$. Notice that $g_z^{-n}(g^n(z)) = z$, g_z^{-n} is continuous and restricted to the set $\underline{D^1 \cap B(g^n(z), 2\delta)}$ is analytic. Since $f^{-n}(A \cap \overline{\Omega}) \subset A \cap \Omega$, we conclude that $g_z^{-n}(\overline{D^1} \cap B(g^n(z), 2\delta) \subset F \cap \overline{D^1}$. Since in addition $I(B(g^n(z), 2\delta)) = B(g^n(z), 2\delta)$ and $I(F) = F$, applying the Schwarz reflection principle again, we conclude that g_z^{-n} extends to an analytic map $g_z^{-n} : B(g^n(z), 2\delta) \to F$. Since $g^n \circ g_z^{-n}$ is an identity map on $D^1 \cap B(g^n(z), 2\delta)$ and g_z^{-n} is analytic on this intersection, we conclude that $g^n \circ g_z^{-n}$ is an identity map on the entire ball $B(g^n(z), 2\delta)$. This means that $g_z^{-n} : B(g^n(z), 2\delta) \to F$ is a holomorphic inverse branch of g^n sending $g^n(z)$ to z. Since

$$\lim_{n\to\infty} \sup_{z\in S^1} \left\{ \operatorname{diam} \left(f_{R(z)}^{-n} \left(R(B(g^n(z), 2\delta)) \right) \right\} = 0,$$

since

$$g_z^{-n} \left(\overline{D^1} \cap B(g^n(z), 2\delta) \right) = R^{-1} \left(f_{R(z)}^{-n} \left(R(B(g^n(z), 2\delta)) \right) \right)$$

and since $R^{-1} : \overline{D^1} \to \overline{\Omega}$ is a uniformly continuous funtion, we get

$$\lim_{n\to\infty} \sup_{z\in S^1} \left\{ \operatorname{diam} \left(g_z^{-n} (\overline{D^1} \cap B(g^n(z), 2\delta) \right) \right\} = 0.$$

Since

$$g_z^{-n} \big(B(g^n(z), 2\delta) \cap (\mathbb{C} \setminus D^1) \big) = I\big(g_z^{-n} \big(B(g^n(z), 2\delta) \cap (\mathbb{C} \setminus D^1) \big) \big),$$

we thefore see that

$$\lim_{n\to\infty} \sup_{z\in S^1} \left\{ \operatorname{diam} \left(g_z^{-n} (B(g^n(z), 2\delta) \right) \right\} = 0.$$

Hence, applying Koebe's distortion theorem, we get that

$$\lim_{n\to\infty} \sup_{z\in S^1} \{ |(g_z^{-n})'(w)| : w \in B(g^n(z), \delta) \} = 0.$$

There thus exists $n \geq 1$ such that

$$\sup_{z\in S^1} \{ |(g_z^{-n})'(w)| : w \in B(g^n(z), \delta) \} \leq \frac{1}{2}.$$

Hence, for every $z \in S^1$, $|(g^n)'(z)| = |(g_z^{-n})'(g^n(z))|^{-1} \geq 2$ and the proof is complete. ♣

Lemma 7.3.2. There exists $\delta > 0$ such that if $\{x, g(x), g^2(x), \ldots, g^n(x)\} \subset B(S^1, \delta)$, then there exists a unique holomorphic inverse branch g_x^{-n} of g^n defined on $B(g^n(x), 4\delta)$ and sending $g^n(x)$ to x. In addition, there exist $\lambda > 1$ and $C > 0$ such that $|(g_x^{-n})'(w)| \leq C\lambda^{-n}$ for all $w \in B(g^n(x), 4\delta)$.

Proof. It follows from Lemma 7.3.1 that there exist $c > 0$ and $\eta > 1$ such that for every $z \in S^1$ and all $w \in B(g^n(z), \delta)$,

$$|(g_z^{-n})'(w)| \le C\eta^{-n}. \tag{7.3.1}$$

Fix $q \ge 1$ so large that $C\eta^{-q} \le 1/6$. There obviously exists $0 < \theta < \delta/6$ such that for every $1 \le j \le q$ and every $x \in \mathbb{C}$ such that $\{x, g(x), g^2(x), \ldots, g^j(x)\} \subset B(S^1, \theta)$, there exists a unique holomorphic branch $g_x^{-j} : B(g^j(x), 4\theta) \to F$ of g^j sending $g^j(x)$ to x. Suppose now that $j = q$. Since $B(g^q(x), 2\theta) \cap S^1 \ne \emptyset$, we can take a point z lying in this intersection. Since $g^{-q}(S^1) = S^1$, we have $g_x^{-q}(z) \in S^1$. Since in addition $B(g^q(x), 4\theta) \subset B(z, 6\theta) \subset B(z, \delta)$, it follows (7.3.1) that

$$|(g_x^{-q})'(w)| \le \frac{1}{6} \tag{7.3.2}$$

for all $w \in B(g^q(x), 4\theta)$. Recalling the choice of w we therefore obtain that

$$g_x^{-q}(B(g^q(x), 4\theta)) \subset B(S^1, \theta) \cap B(x, \theta) \tag{7.3.3}$$

Thus given a piece $\{x, g(x), \ldots, g^n(x)\}$ of the forward trajectory of x, we can split it into blocks

$$g^n(x), g^{n-1}(x), \ldots, g^{n-q}(x);\ g^{n-q}(x), g^{n-q-1}(x), \ldots, g^{n-2q}(x);\ \ldots, x$$

all of them of length q except the last one of length $i \le q$. In view of (7.3.3) we can inductively form the composition

$$g_x^{-i} \circ \ldots \circ g^{-q}_{g^{n-3q}(x)} \circ g^{-q}_{g^{n-2q}(x)} \circ g^{-q}_{g^{n-q}(x)}$$

which is an inverse branch g_x^{-n} of g^n defined on $B(g^n(x), 4\theta)$ and sending $g^n(x)$ to x. Writing $n = pq + i$, $0 \le i \le q-1$ we see from (7.3.2) that for every $w \in B(g^n(x), 4\theta)$, we have

$$|(g_x^{-n})'(w)| \le M\left(\frac{1}{6}\right)^p \le M\left(\frac{1}{6}\right)^{\frac{n}{q}-1},$$

where $M = (\min_{j \le q-1} \inf_{z \in F}\{|(g^j)'(z)|\})^{-1}$. We are done redefining δ to be θ. ♣

Proposition 7.3.3 With $U = B(S^1\delta)$, where δ comes from Lemma 7.3.2,the map $F : S^1 \to S^1$ is a boundary expanding repeller.

Proof. Since $R|_{S^1}$ establishes a topological conjugacy between $g|_{S^1}$ and $f|_{\partial\Omega}$, the map $g : S^1 \to S^1$ is topologically transitive. Since, by Lemma 7.3.1, $g : S^1 \to S^1$ is expanding, we only need to check that there exists an open set $U \subset F$ containing S^1 such that $\bigcap_{n \ge 0} g^{-n}(U) = S^1$. And indeed, suppose that $\{g^n(x) : n \ge 0\} \subset B(S^1, \delta)$, where δ is taken from Lemma 7.3.2. It follows from this lemma that for every $n \ge 0$ there exists a unique holomorphic inverse branch g_x^{-n} of g^n defined on $B(g^n(x), 4\delta)$ and sending $g^n(x)$ to x. In addition

$$\operatorname{diam}\bigl(g_x^{-n}(B(g^n(x), 3\delta)\bigr) \le 6\delta c\lambda^{-n}.$$

Since $B(g^n(x),3\delta)\cap S^1 \neq \emptyset$, we can choose a point w in this intersection. Then $g_x^{-n}(w) \in S^1$ and therefore

$$\operatorname{dist}(x,S^1) \leq |g_x^{-n}(g^n(x)) - g_x^{-n}(w)| \leq \operatorname{diam}\bigl(g_x^{-n}(B(g^n(x),3\delta)\bigr) \leq 6\delta c\lambda^{-n}.$$

Letting $n \to \infty$, we therefore conclude that $x \in S^1$. Thus g becomes a boundary expanding repeller with $U = B(S^1,\delta)$. ♣

Our next result is the following the following.

Theorem 7.3.2. If μ is a g-invariant ergodic probability measure of of positive entropy on S^1, then the non-tangent limit

$$\lim_{x\to z} \frac{-\log|R'(x)|}{\log(1-|x|)}$$

exists for μ-almost every point $z \in S^1$, is constant almost everywhere (denote it by $\chi_\mu(R)$), and

$$\chi_\mu(R) = 1 - \frac{\chi_{\mu\circ R^{-1}}(f)}{\chi_\mu(g)}.$$

Proof. Given $z \in S^1$, $0 < \alpha < \pi/2$, and $0 < r < 1$ let

$$S_{\alpha,r}(z) = z\cdot(1 + \{x \in \mathbb{C}\setminus\{0\} : 0 < |x| \leq r, \pi - \alpha \leq \operatorname{Arg}(x) \leq \pi + \alpha\}).$$

A straightforward trigonometrical argment shoows that for all $0 < \alpha < \pi/2$, and $0 < r < 1$ there exists $0 < \kappa < 1/2$ such that for all $z \in S^1$ and all $x \in S_{\alpha,r}(z)$ we have

$$\kappa|x - z| \leq 1 - |x|. \tag{7.3.4}$$

Fix now $z \in S^1$ and $x \in S_{\alpha,r}(z)$. It follows from Proposition 7.3.3 that there exists $k \geq 1$ such that $g^j(x) \in B(S^1,\delta)$ for all $0 \leq j \leq k-1$ and $g^k(x) \notin B(S^1,\delta)$ which means that $1 - |g^k(x)| \geq \delta$. Therefore there exists the least $n = n(x,z) \geq 0$ such that

$$1 - |g^{n+1}(x)| \geq \delta(8K\kappa^{-1})^{-1}, \tag{7.3.5}$$

where $K \geq 1$ is the Koebe's constant assiciated with the scale $1/2$. Hence

$$1 - |g^j(x)| < \delta(4K\kappa^{-1})^{-1} < \delta \tag{7.3.6}$$

for all $0 \leq j \leq n$. It therefore follows from Lemma 7.3.2 that thaere exists there a unique holomorphic inverse branch g_x^{-n} of g^n defined on $B(g^n(x),4\delta)$ and sending $g^n(x)$ to x. Let $w = g^n(x)/|g^n(x)|$. Then $w \in \overline{B}(g^n(x), 1 - |g^n(x)|) \cap S^1$ and using (7.3.4) along with Koebe's Distortion Theorem, we get

$$\begin{aligned}
\kappa|x - z| &\leq 1 - |x| \\
&\leq |x - g_x^{-n}(w)| = |g_x^{-n}(g^n(x)) - g_x^{-n}(w)| \leq K|g^n(x) - g_x^{-n}(w)||(g^n)'(x)|^{-1} \\
&= K(1 - |g^n(x)|)|(g^n)'(x)|^{-1}.
\end{aligned}$$

Hence, applying Koebe's $\frac{1}{4}$-Theorem, we obtain

$$
\begin{aligned}
g_x^{-n}\big(B(g^n(x), 8K\kappa^{-1}(1-|g^n(x)|)\big) &\supset B\big(x, 2K\kappa^{-1}(1-|g^n(x)|)|(g^n)'(x)|^{-1}\big) \\
&\supset B(x, 2|z-x|) \ni z.
\end{aligned} \tag{7.3.7}
$$

It therefore follows from Koebe's Distortion Theorem that

$$
K^{-1} \le \frac{|(g^n)'(x)|}{|(g^n)'(z)|} \le K, \tag{7.3.8}
$$

that (using (7.3.4))

$$
\begin{aligned}
1-|g^n(x)| &\le |g^n(z)-g^n(x)| \le K|(g^n)'(x)|\cdot|z-x| \le K|(g^n)'(x)|\kappa^{-1}(1-|x|) \\
&= K\kappa^{-1}|(g^n)'(x)|(1-|x|),
\end{aligned} \tag{7.3.9}
$$

and (using (7.3.7)) that

$$
\begin{aligned}
1-|x| &\le |z-x| = |g_x^{-n}(g^n(z)) - g_x^{-n}(g^n(x))| \le K|(g^n)'(x)|^{-1}|g^n(z)-g^n(x)| \\
&\le K|(g^n)'(x)|^{-1} 8K\kappa^{-1}(1-|g^n(x)|) = 8K^2\kappa^{-1}(1-|g^n(x)|)|(g^n)'(x)|^{-1}.
\end{aligned} \tag{7.3.10}
$$

Since the Riemann map $: D^1 \to \Omega$ is uniformly continuous, $R(g^n(x))$ lies close to $R(g^n(z))$. Let $f_{R(z)}^{-n}$ be a holomorphic inverse branch of f^n defined on some small neighbourhood of $R(g^n(z))$, containing $R(g^n(x))$ and sending $R(g^n(z)) = f^n(R(z))$ to $R(z)$. Then $f_{R(z)}^{-n}(R(g^n(x)) = x$ and applying Koebe's Distortion Theorem, we obtain

$$
\hat{K}^{-1} \le \frac{|(f^n)'(R(x))|}{|(f^n)'(R(z))|} \le \hat{K}, \tag{7.3.11}
$$

for some constant $\hat{K}$ independent of z, x and n. Since

$$
1-|g^{n+1}(x)| \le ||g'||(1-|g^n(x)|), \tag{7.3.12}
$$

Combining (7.3.5), (7.3.6), (7.3.9), (7.3.10) and (7.3.12), we get

$$
(8K||g'||)^{-1}\kappa^2\delta \le (1-|x|)|(g^n)'(x)| \le K\delta. \tag{7.3.13}
$$

By Birkhoff's Ergodic Theorem there exists a Borel set $Y \in S^1$ such that $\mu(Y)=1$ and

$$
\lim_{k\to\infty}\frac{1}{k}\log|(g^k)'(z)| = \chi_\mu(g) \quad\text{and}\quad \lim_{n\to\infty}\frac{1}{k}\log|(f^k)'(R(z))| = \chi_{\mu\circ R^{-1}}(f).
$$

for all $z \in Y$. Suppose that $z \in Y$. Fix also $\varepsilon > 0$. Then for all k sufficiently large

$$
\exp(k(\chi_\mu(g)-\varepsilon)) \le |(g^k)'(z)| \le \exp(k((\chi_\mu(g)+\varepsilon)), \tag{7.3.14}
$$

and

$$\exp(k(\chi_{\mu\circ R^{-1}}(f)-\varepsilon)) \le |(f^k)'(R(z))| \le \exp(k((\chi_{\mu\circ R^{-1}}(f)+\varepsilon)).$$

Combining these two formulas along with (7.3.8) and (7.3.11), we get for all n large enough (which can be assured by taking x sufficiently close to z),

$$\exp(n(\chi_\mu(g)-2\varepsilon)) \le |(g^n)'(x)| \le \exp(n(\chi_\mu(g)+2\varepsilon)), \tag{7.3.15}$$

and

$$\exp(n(\chi_{\mu\circ R^{-1}}(f)-2\varepsilon)) \le |(f^n)'(R(x))| \le \exp(n(\chi_{\mu\circ R^{-1}}(f)+2\varepsilon)) \tag{7.3.16.}$$

Combining (7.3.13) and (7.3.14), we get

$$\log(4||g'||)^{-1}\delta) - n(\chi_\mu(g)+2\varepsilon) \le \log(1-|x|) \le \log(K\delta) - n(\chi_\mu(g)-2\varepsilon).$$

Thus, for all n large enough

$$\frac{-\log(1-|x|)}{\chi_\mu(g)+3\varepsilon} \le n \le \frac{-\log(1-|x|)}{\chi_\mu(g)-3\varepsilon}. \tag{7.3.17}$$

In view of Koebe's Distortion Theorem, (7.3.12) and (7.3.5) there exists a constant C depending only on K, κ and $||g'||$ such that

$$C^{-1}|R'(0)| \le |R'(g^n(x))| \le C|R'(0)|$$

Therefore, differentiating the equality $R\circ g^n = f^n\circ R$, we get

$$|R'(x)| = |(g^n)'(x)|\frac{|R'(g^(x))|}{|(f^n)'(R(x))|} \in \left[C^{-1}|R'(0)|\frac{|(g^n)'(x)|}{|(f^n)'(R(x))|}, C|R'(0)|\frac{|(g^n)'(x)|}{|(f^n)'(R(x))|}\right].$$

Hence, using (7.3.15), (7.3.16) and (7.3.17)

$$\begin{aligned}|R'(x)| &\le C|R'(0)||(g^n)'(x)|\cdot|(f^n)'(R(x))|^{-1}\\ &\le C|R'(0)|\exp(n(\chi_\mu(g)+2\varepsilon))\exp(-n(\chi_{\mu\circ R^{-1}}(f)-2\varepsilon))\\ &\le \exp\left(\frac{-\log(1-|x|)}{\chi_\mu(g)-3\varepsilon}(\chi_\mu(g)-\chi_{\mu\circ R^{-1}}(f)+4\varepsilon)\right).\end{aligned}$$

Thus

$$\frac{-\log|R'(x)|}{\log(1-|x|)} \preceq \frac{\chi_\mu(g)-\chi_{\mu\circ R^{-1}}(f)+4\varepsilon}{\chi_\mu(g)-3\varepsilon}$$

and we obtain similarly

$$\frac{-\log|R'(x)|}{\log(1-|x|} \succeq \frac{\chi_\mu(g)-\chi_{\mu\circ R^{-1}}(f)-4\varepsilon}{\chi_\mu(g)+3\varepsilon}$$

So, letting $\varepsilon \to 0$ (which forces us to let $n(x,z) \to \infty$ and which in turn forces us to let $x \to z$) finishes the proof. ♣

In order to make use of this result we shall provide a simple proof of the following Proposition.

Section 7.4. Harmonic measure.

In this section we keep the notation of the previous one. The measure $l \circ R^{-1}$, the image of the image of the Lebesgue measure l on S^1 under the Riemann map is said to be the harmonic measure of $\partial\Omega$ with respect to (or viewed from) the point z_0. Since all the Riemann maps differ by compositions with M̈bius transformations preserving t he unit circle, all the harmonic measures are strongly equivalent and corresponding Radon-Nikodyn derivatives are bounded away from zero and infinity. In particular all the harmonic measures induce the same strong measure class, which as long as we are only interested in metric properties of this class, enables us to speak generally about a harmonic measure without specifying the point z_0. Writing ω we will actually mean the class of all measures equivalent with the harmonic measure with Radon-Nikodym derivatives bounded uniformly from above and below. For more information about harmonic measure we refer the reader to [?]. Our first aim is to represent harmonic measure as a Gibbs measure and then to apply the results of the previous section. Since the Hausdorff dimension of the circle S^1 is equal to 1, it follows from Theorem 7.1.1 that $P(g,\phi)=0$, where $\phi = -\log|g'|$. Of course the Lebesgue measure l on S^1 is equivalent with H_1. Since R is a topological conjugacy between g and f on S^1, $P(f,\phi\circ R^{-1}) = P(g,\phi) = 0$. Since moreover R^{-1} is a metric conjugacy between metric dynamical systems $(f,\mu_\phi\circ R^{-1})$ and (g,μ_ϕ), we therefore have

$$\mathrm{h}_{\mu_\phi\circ R^{-1}}(f) + \int \phi\circ R^{-1}\, d\mu_\phi\circ R^{-1} = \mathrm{h}_{\mu_\phi}(g) + \int \phi\, d\mu_\phi = 0.$$

Since in addition $\mu_\phi \circ R^{-1} \in \omega$, the uniqueness of an equilibrium state for $\phi\circ R^{-1}$ results in the following.

Theorem 7.4.2. The harmonic measure ω coincides with the class of the Gibbs states of the map $f: X \to X$ and the Hölder continuous potential $-\log|g'|\circ R^{-1}$.

We now want to argue that $\mathrm{HD}(\omega)=1$. This is a general result due to Makarov (see ???) true for any simply connected domain Ω with no dynamics involved. We shall however provide here a proof in the dynamical context only which is shorter and simpler than the general one.

Proposition 7.4.3. If ω is a harmonic measure on the boundary of a Jordan domain Ω, then $\chi_l(R)(z) = 0$ for l-a.e. $z \in S^1$, where l denotes the normalized Lebesgue measure on S^1.

Proof. Fix $z \in S^1$ and $0 < r < 1$. Then by the Koebe distortion theorem $2^{-1}(1-r)^{-3} \leq |R'(z)| \leq (1-r)^{-3}$ which implies that

$$2 \leq \frac{-\log |R'(rz)|}{\log(1-r)} \leq 4$$

for all r sufficiently close to 1. Since the Gibbs state of the function $-\log |g'|$ is a g-invariant probability measure absolutely continuous with respect to the Lebesgue measure it follows from Theorem 7.4.2 that $\chi_l(R)(z)$ is constant for l-a.e. $z \in S^1$. Hence

$$\begin{aligned}\chi_l(R) &= \int \chi_l(R)\, dl = \int \lim_{r \to 1} \frac{\log |R'(rz)|}{-\log(1-r)}\, dl(z) \\ &= \lim_{r \to 1} \frac{1}{-\log(1-r)} \int \log |R'(rz)|\, dl(z).\end{aligned}$$

Since $\log |R'|$ is a harmonic function we can continue the above chain of equalities writing

$$\chi_l(R) = \lim_{r \to 1} \frac{r \log |R'(0)|}{-\log(1-r)} = 0.$$

The proof is finished. ♣

Corollary 7.4.4. Suppose that an expanding map $f : \partial\Omega \to \partial\Omega$ is a conformal boundary repeller, and Ω is a Jordan domain. Then $\mathrm{HD}(\omega) = 1$, for ω, the harmonic measure viewed from Ω.
proof. Let μ be the Gibbs state on S^1 corresponding to the potential $-\log |g'|$. Simultaneously μ is the unique probability g-invariant measure equivalent with the Lebesgue measure l on S^1. Then $\mu \circ R^{-1}$ is a probability f-invariant measure equivalent to ω. In view of Theorem 7.4.2 and Lemma 7.4.3, $\chi_{\mu \circ f^{-1}}(f) = \chi_\mu(g) > 0$. Since $R : S^1 \to \partial\Omega$ is a topological conjugacy between $g : S^1 \to S^1$ and $f : \partial\Omega \to \partial\Omega$, $\mathrm{h}_{\mu \circ f^{-1}}(f) = \mathrm{h}_\mu(g)$. Thus

$$\frac{\mathrm{h}_{\mu \circ f^{-1}}(f)}{\chi_{\mu \circ f^{-1}(f)}} = \frac{\mathrm{h}_\mu(g)}{\chi_\mu(g)}.$$

Since $\mathrm{HD}(\mu) = 1$, an immediate application of Lemma 7.1.4 (Volume Lemma) finishes the proof. ♣

Theorem 7.4.5. Let $f : \partial\Omega \to \partial\Omega$ be a conformal boundary repeller, where Ω is a Jordan domain. Then either
(a) $\omega \asymp \mathrm{H}_1$ on $\partial\Omega$ (if $\log |g'|$ and $\log |f' \circ R|$ are cohomologous) or
(b) $\omega \perp \mathrm{H}_1$ (if $\log |g'|$ and $\log |f' \circ R|$ are not cohomologous), and then there exists $c_0 > 0$ such that with the gauge function $\phi_c(t) = t \exp(c\sqrt{\log(1/t) \log_3(1/t)})$,

$$\omega \perp \mathrm{H}_{\phi_c} \quad \text{for all} \quad 0 \leq c < c_0$$

and

$$\omega \ll \mathrm{H}_{\phi_c} \quad \text{for all } c > c_0.$$

Proof. If $\log|g'|$ and $\log|f' \circ R|$ are cohomologous, then also the functions $-\log|g' \circ R^{-1}|$ and $-\log|f'|$ are cohomologous (with respect to the map $f : \partial\Omega \to \partial\Omega$). By Theorem 7.4.2 ω is a Gibbs state of the potential $-\log|g' \circ R^{-1}|$. Since $\mathrm{P}(f, -\log|f'|) = \mathrm{P}(f, -\log|g' \circ R^{-1}|) = \mathrm{P}(g, -\log|g'|) = 0$, it follows from Theorem 7.1.1 that $\mathrm{HD}(\partial\Omega) = 1$, and the Gibbs states of the potential $-\log|f'|$ are equivalent to the 1-dimensional Hausdorff measure on $\partial\Omega$. So, the part (a) of Theorem 7.4.5 is proved.
Suppose now that $\log|g'|$ and $\log|f' \circ R|$ are not cohomologous. Then $-\log|g' \circ R^{-1}|$ and $-\log|f'|$ are not cohomologous. Let μ be the invariant Gibbs state of $-\log|g' \circ R^{-1}|$. By Corollary 7.4.4 and Theorem 7.a.2, $\mathrm{HD}(\mu) = 1$. Hence $\psi = -\log|g' \circ R^{-1}| + \mathrm{HD}(\mu)\log|f'| - \mathrm{P}(f, -\log|g' \circ R^{-1}|)$ is not a coboundary and, since by Theorem 7.4.2 and Lemma 7.4.3, $\int \log|f'|\, d\mu = \int \log|g' \circ R^{-1}|\, d\mu$, the potential ψ is not cohomologous to a constant. The second part of Theorem 7.4.5 is now an immediate consequence of Theorem 7.2.1(b). ♣

Theorem 7.4.6. If $f : \partial\Omega \to \partial\Omega$ is a conformal boundary repeller, Ω is a Jordan domain and the functions $\log|g'|$ and $\log|f'| \circ R$ are cohomologous, (this is exactly the case of Theorem 7.4.5(b)), then Ω is an analytic curve. If additionally f extends analytically onto $\overline{\mathbb{C}}$, then f is analytically conjugate with a Möbius transformation and $\partial\Omega$ is a geometric circle.

Proof. If the functions $\log|g'|$ and $\log|f'| \circ R$ are cohomologous, then according to Theorem 7.4.5(a) the boundary $\partial\Omega$ is a rectifiable Jordan curve. So, in view of Riesz's theorem, the map $R : S^1 \to \partial\Omega$ transports the measure class of the Lebesgue measure on S^1 onto the measure class of H_1 on $\partial\Omega$. Let now $R_* : S^1 \to \partial\Omega$ be the restriction to the unit circle of the Riemann map induced by the second component of the complement of $\partial\Omega$. For technical reason, which will be clear at the end of this proof we assume here that the Riemann map R_* is defined on $D^1_* = \{z : |z| > 1\}$, the complement of the closure of the unit disk D^1. Since $\partial\Omega$ is a Jordan curve and since f has no critical points on $\partial\Omega$, there exists an open neighborhood A_* of $\partial\Omega$ such that $f(A_* \cap (\overline{\mathbb{C}} \setminus \partial\Omega))$ is well-defined, and moreover $f(A_* \cap \overline{\mathbb{C}} \setminus \partial\Omega) \subset \overline{\mathbb{C}} \setminus \partial\Omega)$. Therefore, we can define $g_* = R_*^{-1} \circ f \circ R_*$, the lift of f via the Riemann map R_* on the set D^1_* intersected with a sufficiently thin open annulus surrounding S^1. Set

$$h = R_*^{-1} \circ R|_{S^1} : S^1 \to S^1.$$

Composing, if necessary, the Riemann maps R and R_* with appropriate rotations, we may assume that 1 is a fixed point of g and g_* and $h(1) = 1$. Our first objective is to demonstrate that h is real-analytic. Indeed, Let $\mu_1 = \phi_1 l$ and $\mu_2 = \phi_2 l$ be the (unique) probability measures equivalent with the Lebesgue measure on the circle, respectively invariant under the action of g and g_*. In view of Theorem 7.9.2 ϕ_1 and ϕ_2 are both Real-analytic. Since, also by Riesz's theorem, the map $R_* : S^1 \to \partial\Omega$ transports the measure class of the Lebesgue measure on S^1 onto the measure class of H_1 on $\partial\Omega$, the homeomorphism $h : S^1 \to S^1$ sends the measure class of the Lebesgue measure on S^1 onto itself. Since h establishes conjugacy between g and g_*, it therefore maps the invariant measure μ_1 onto some probability g_*-ivariant measure equivalent with the Lebesgue measure. Since such a measure is unique, it must be equal to μ_2. Symbolically, $\mu_1 \circ h^{-1} = \mu_2$. Define now two functions $M, N : S^1 \to [0,1]$ by setting

$$M(z) = \mu_1([1,z]) = \int_1^z \phi_1 \, dl$$

and

$$N(z) = \mu_2([1,z]) = \int_1^z \phi_2 \, dl.$$

Since $\mu_1 \circ h^{-1} = \mu_2$ and $h(1) = 1$, the functions M and N are related by the formula $N(h(z)) = M(z)$. Since M and N are strictly increasing, we may solve the last equation for h to get $h = M \circ N^{-1}$. We are now done because the real analyticity of ϕ_1 and ϕ_2 (see ???) implies that the functions M and N are real-analytic and this in turn results in real analyticity of $h = M \circ N^{-1}$. Hence h extends to an analytic map $\tilde{h}$ defined on an open neighborhood H of S^1 in $\mathbb{C}$. Since $h : S^1 \to S^1$ preserves orientation, decreasing H if necessary, we get $\tilde{h}(H \cap \overline{D^1_*}) \subset \overline{D^1_*}$. Thus we produced two continuous maps $R : \overline{D^1} \to \mathbb{C}$ and $R_* \circ \tilde{h} : H \cap \overline{D^1_*} \to \mathbb{C}$ which are analytic on D^1 and $H \cap D^1_*$ respectively and which coincide on their common boundary, the unit circle S^1. Thus R and $R_* \circ \tilde{h}$ glue together to an analytic map $S : D^1 \cup H \to \overline{\mathbb{C}}$. So, since $S(S^1) = R(S^1) = \partial\Omega$, the proof of the first part is finished.

If now f extends analytically to $\overline{\mathbb{C}}$, that is if f is a rational function, then by the maximum principle $f(\Omega) \subset \Omega$, $f(\overline{\mathbb{C}} \setminus \overline{\Omega}) \subset \overline{\mathbb{C}} \setminus \overline{\Omega}$, and both maps g and g_* are defined on D^1 and D^1_* respectively. Since these maps are surjective, since they extend continuously to $\overline{D^1}$ and $\overline{D^1_*}$ respectively, and since they preserve the unit circle S^1, it follows from the Schwarz reflection principle that they extend analytically to $\overline{\mathbb{C}}$. Since g and g^* preserve S^1 (we use this fact second time) they must be finite Blaschke products. Since $f(\Omega) \subset \Omega$ and $f(\overline{\mathbb{C}} \setminus \overline{\Omega}) \subset \overline{\mathbb{C}} \setminus \overline{\Omega}$, by the Montel theorem both Ω and $\overline{\mathbb{C}} \setminus \Omega$ are components of the Fatou set of f. And since $f|_{\partial\Omega}$ is expanding, both Ω and $\overline{\mathbb{C}} \setminus \overline{\Omega}$ are basins of immediate attraction to stable fixed points. Conjugating f if necessary with a Möbius transformation we may assume that this fixed point in D^1 is 0 and that one in D^1_* coincides with ∞. But every Blaschke product B preserving S^1 and having 0 as its fixed point, preserves the Lebesgue measure on S^1. In order to see it consider an arbitrary continuous function $\phi : S^1 \to \mathbb{R}$ and then its harmonic extension $\tilde{\phi} : D^1 \to \mathbb{R}$. Since $\phi \circ B$ is also harmonic, we have

$$\int_{S^1} \phi \circ B \, dl = \phi \circ B(0) = \phi(0) = \int_{S^1} \phi \, dl.$$

which means that B preserves the Lebesgue measure l. Thus, $\mu_1 = \mu_2 = l$ and consequently $M = N$ and h is the identity map on S^1. In conclusion $\tilde{h}$ is an identity map and R_* and $R = R_* \tilde{h}$ coincide on a neighborhood of S^1. Thus R and R_* glue together to an analytic map $\tilde{R} : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$. Since $\tilde{R}$ is injective it must be a Möbius transformation and $\tilde{R}^{-1} \circ f \circ R^{-1}$ is a finite Blaschke product. The proof is finished. ♣

Section 7.5. Radial behaviour of the Riemann map,II.

Keeping notation from the previous sections we shall prove here the following .

Theorem 7.5.1. Depending on whether $c(\omega) = 0$ or $c(\omega) \neq 0$, either $\partial\Omega$ is real-analytic and the Riemann map $R : D^1 \to \Omega$ and its derivative R' extend holomorphically beyond ∂D^1 or for almost every $z \in \partial D^1$

$$\limsup_{r\to 1} |R'(rz)| \exp c\sqrt{\log(1/1-r)\log_3(1/1-r)} = \begin{cases} \infty & \text{if } c \leq c(\omega) \\ 0 & \text{if } c > c(\omega) \end{cases} \tag{7.5.1}$$

and

$$\limsup_{r\to 1} \big(|R'(rz)| \exp c\sqrt{\log(1/1-r)\log_3(1/1-r)}\big)^{-1} = \begin{cases} \infty & \text{if } c \leq c(\omega) \\ 0 & \text{if } c > c(\omega) \end{cases} \tag{7.5.2}$$

Moreover the radial limsup can be replaced by the nontangential one.

Proof. Let $n > 0$ be the least integer for which $g^n(rz) \in B(0, r_0)$ for some fixed $r_0 < 1$. We have $R'(rz) = ((f^n)'(R(rz)))^{-1} \cdot R'(g^n(rz)) \cdot (g^n)'(rz)$. Hence, for some constant $K > 0$ independent of r and z

$$K^{-1} \leq \frac{|R'(rz)|}{|((f^n)'(R(rz)))^{-1}| \cdot |R'(g^n(rz))|} \leq K.$$

By the bounded distortion theorem the rz in the denominator can be replaced z and n depends on r as described by (7.2.1) with r replaced by $1-r$.Now we proceed as in the proof of Lemma 7.2.1 replacing deviations of $S_n(\phi) - P(\phi)n + \kappa \log|(f^n)'(x)|$ by the deviations of $\log|(g^n)'(x)| - \log|(f^n)'(x)|$. The proof is finished. ♣

Section 7.6.. Pressure versus integral means of the Riemann map

In this section we establish a close relation between integral means of derivatives of the Riemann map and topological pressure of the function $-t\log|f'|$. Given $t \in I\!R$ define

$$\beta(t) = \limsup_{r\to 1} \frac{\log \int_{S^1} |R'(rz)|^t \, dl(z)}{-\log(1-r)}$$

We shall prove the following.

Theorem 7.6.1. If the lifted map $g : S^1 \to S^1$ is of the form $z \mapsto z^d$, $d \geq 2$, then

$$\beta(t) = t - 1 + \frac{\mathrm{P}(f, -t\log|f'|)}{\log d}.$$

Proof. Fix $0 < r < 1$ and divide the circle S^1 into $[2\pi/(1-r)]$ arcs of length $1-r$ and one arc of length $\leq 1-r$. Denote these arcs by $I_1, I_2, \ldots, I_k$ and I_{k+1} respectively, where $k = [2\pi/(1-r)]$. Then

$$\int_{S^1} |R'(rz)|^t dl(z) = \sum_{j=1}^{k+1} \int_{I_j} |R'(rz)|^t dl(z) = \sum_{j=1}^{k+1} \int_{I_j} |R'(g^n(rz))|^t \frac{(|g^n)'(rz)|^t}{|(f^n)'(R(rz))|^t} dl(z).$$

Fix now $n = n(r)$ to be the first integer for which $|g^n(z)| < 1/2$. Note that n is independent of z and that there exists a constant $A \geq 1$ such that $A^{-1} \leq |R'(w)| \leq A$ for all $w \in B(0, 1/2)$. Hence

$$\int_{S^1} |R'(rz)|^t dl(z) \asymp \sum_{j=1}^{k+1} \int_{I_j} \frac{(|g^n)'(rz)|^t}{|(f^n)'(R(rz))|^t} dl(z). = \sum_{j=1}^{k+1} \int_{I_j} \frac{(d^{nt}}{|(f^n)'(R(rz))|^t} dl(z).$$

Now, by the Mean Value Theorem for every $j = 1, 2, \ldots, k, k+1$ there exists $z_j \in I_j$ such that

$$\int_{I_j} \frac{(d^{nt}}{|(f^n)'(R(rz))|^t} dl(z) = l(I_j) \frac{(d^{nt}}{|(f^n)'(R(rz_j))|^t}.$$

Hence, as $l(I_j) = 1-r$ for all $j = 1, \ldots, k$

$$\begin{aligned}
&\log \int_{S^1} |R'(rz)|^t dl(z) \\
&= \log(1-r) + \log\left(\sum_{j=1}^{k} \frac{(d^{nt}}{|(f^n)'(R(rz_j))|^t} + \frac{(d^{nt}}{|(f^n)'(R(rz_{k+1}))|^t} \cdot \frac{l(I_{k+1})^t}{(1-r)^t}\right) + O(1) \\
&= \log(1-r) + nt\log d + \\
&+ \log\left(\sum_{j=1}^{k} \exp \sum_{u=0}^{n-1} -t\log|f'(R(g^u(rz_j)))| + |(f^n)'(R(rz_j))|^t \frac{l(I_{k+1})^t}{(1-r)^t}\right) + O(1)
\end{aligned}$$

By our definition of n, $(1/2)^d \leq r^{d^n} \leq 1/2$; hence $d\log(1/2) \leq d^n \log r \leq \log(1/2)$. Since there exists a constant $B \geq 1$ such that $B^{-1}(1-r) \leq -\log r \leq B(1-r)$ for all r sufficiently close to 1, we get $B^{-1}\log 2 \leq d^n(1-r) \leq Bd\log 2$. Therefore $-\log B + \log\log 2 \leq n\log d + \log(1-r) \leq \log B + \log\log 2 + \log d$. Hence $n\log d - C \leq -\log(1-r) \leq n\log d + C$ for some universal constant C. Thus

$$\begin{aligned}
&\frac{\log \int_{S^1} |R'(rz)|^t dl(z)}{-\log(1-r)} = -1 + t + \\
&\frac{1}{n\log d} \log\left(\sum_{j=1}^{k} \exp \sum_{u=0}^{n-1} -t\log|f'(R(g^u(rz_j)))| + \exp \sum_{u=0}^{n-1} -t\log|f'(R(g^u(rz_{k+1})))| \frac{l(I_{k+1})^t}{(1-r)^t}\right) \\
&+ o(1).
\end{aligned}$$

Now, using the fact that $0 \leq l(I_{k+1})/(1-r) \leq 1$, it follows from the definition of pressure and the bounded distortion property that

$$\limsup_{r\to 1} \frac{\log \int_{S^1} |R'(rz)|^t dl(z)}{-\log(1-r)} = -1+t+\frac{1}{\log d}\mathrm{P}(g, -t\log|f'\circ R|) = -1+t+\frac{\mathrm{P}(f, -t\log|f'|)}{\log d}.$$

The proof is finished. ♣

Section 7.7. Geometric examples.

This last section of this chapter is devoted to explore applications of previous sections to geometric examples like the Koch's snowflake and Carleson's example. Following the idea of the proof of Theorem 7.2.1 and copeing with a biggere number of technicalities one can prove the following.

Theorem 7.7.1. Let Ω be a simply connected domain in $\mathbb{C}$ with $\partial\Omega$ a Jordan curve . Let ∂_j, $j = 1, 2 \ldots, k$ be a finite family of compact arcs in $\partial\Omega$ with pairwise disjoint interriors. Denote $\bigcup \partial_j$ by ∂ (we do not assume that this curve is connected. Assume that there exists a family of conformal maps f_j, $j = 1, \ldots, k$, (which may reverse the orientation on $\mathbb{C}$) on neighbourhoods U_j of ∂_j. For every j assume that $f_j(\Omega \cap U - J) \subset \Omega$, $|f'| > 1$ on U_j and

$$f_j(\partial\Omega \cap U_j) \subset \partial\Omega. \tag{7.7.1}$$

Assume also the Markov partition property: for every $j = 1, \ldots k$, $f_j(\partial_j) = \bigcup_{s\in I_j} \partial_s$ for some subset $I_j \subset \{1, 2, \ldots, k\}$. Consider the $k \times k$ matrix $A = A_{jk}$ where $A_{jk} = 1$ if $k \in I_j$ and $A_{jk} = 0$ if $k \notin I_j$. Then there exists a transition parameter $c(\omega, \partial)$ such that the claims of tTheorem 7.6.5 and 7.6.6 hold.

Example 1(*the snowflake*). To every side of an equilateral triangle, in the middle we glue from outside as small as three times. To every side of the resulting polygon we we glue again an equilateral triangle as small as three times and so on infinitely many times. The triangles do not overlap in this construction and the boundary of the resulting domain Ω is a Jordan curve.This Ω is called the Koch's snowflake. It was first describe by Helge Koch in 1904. Denote the curve in $\partial\Omega$ joining a point $x \in \partial\Omega$ to $y \in \partial\Omega$ in the clockwise direction just by xy. For every $\partial A_i A_{i+1 (mod 12)} \subset \partial\Omega$, $i = 0, 1, \ldots, 11$, we consider its civering by the curves 12, 23, 45, 56 in Ω (see Fig.2). This covering together with the affine maps

$$\begin{aligned}
&12,\ 34 \to 12 \text{ (preserving orientation on } \partial\Omega) \\
&23 \to 61 \text{ (reversing orientation)} \\
&56 \to 36 \text{ (preserving orientation)} \\
&45 \to 63 \text{ (reversing orientation)}
\end{aligned}$$

gives a Markov partition of ∂_i satisfying the assumptions of Theorem 6.7.1. Since $\partial\Omega$ (and every its subcurve) is definitely not real-analytic $(\mathrm{HD}(\partial\Omega) = \log 4/\log 3)$, the assertion of Theorem 6.7.1 is valid with $c(\omega, \partial_i) > 0$. We may denote $c(\omega, \partial_i)$ by $c(\omega)$ since it is independent of ∂_i by symmetry.

Example 2(*Carleson's domain*). We recall Carleson's construction from [Ca2]. We fix a broken line γ with the first and last segment lying in the same straight line in $I\!R^2$, with no other segments intersecting the segment $\overline{1, d-1}$ (see Fig. 3). Then we take a regular polygon Ω^1 with vertices $T_0, T_1, \dots, T_n$ and glue to every side of it, from outside, the rescaled, not mirror reflected, curve γ so that the ends of the glued curve coincide with the ends of the side. The resulting curve bounds a second polygon Ω^2. Denote its vertices by $A_0, A_1, \dots$ (Fig. 4). Then we glue again the rescaled γ to all sides of Ω^2 and a third order polygon Ω^3 with vertices $B_0, B_1, \dots$. Then we bild Ω^4 with vertices $C_0, C_1, \dots \Omega^5$ with $D_0, D_1, \dots$ etc. Assume that there is no self-intersecting of the curves $\partial\Omega^n$ in this construction. Moreover assume that in the limit we obtain a Jordan curve $\mathcal{L} = \mathcal{L}(\Omega^1, \gamma) = \partial\Omega$. The natural Markov partition of each curve T_iT_{i+1}in $\mathcal{L}$ into curves A_jA_{j+1} with $f(A_jA_{j+1}) = T_iT_{i+1}$, considered by Carleson does not satisfy the property (6.7.1) so we cannot succeed with it. Instead we proceed as follows: Define in an affine fashion

$$f(B_{d(j-1)+1}B_{dj-1}) = A_1A_{d-1}$$

for every $j = 1, 2 \dots, d$. Divide now every arc $B_{dj-1}A_j$ for $j = 1, 2 \dots, d$ and A_jB_{dj+1}, $j = 1, 2 \dots, d$ into curves with ends in the vertices of the polygon $\Omega^4 : C^j \in B_{dj_1}A_j$, $\tilde{C}^j \in A_jB_{dj+1}$ respectively, the closest to $A_j (\neq A_j)$. Let for $j = 1, 2, \dots, d-1$,

$$f(C^jA_j) = B_{dj-1}A_j, f(B_{dj-1}C^j) = A_{d-1}B_{d^2-1},$$
$$f(A_j\tilde{C}^j) = A_jB_{dj+1}, f(\tilde{C}^jB_{dj+1}) = B_1A_1.$$

This gives a transitive aperiodic Markov partition of $B_1B_{d^2-1}$. We can consider instead of the broken line γ in the construction of Ω, the line $\gamma^{(2)}$, consisting of d^2 segments, which arises by glueing to every side of γ a rescaled γ. Consecutive gluing of the rescaled $\gamma^{(2)}$ to the polygon $\Omega^{(1)}$ gives consecutively Ω^3, Ω_5 etc. The same construction as above gives a Markov partition of $D_1D_{d^2-1}$ in T_iT_{i+1}. By continuing this procedure we approximate T_iT_{i+1}, so from Theorem 6.7.1 and from symmetry we deduce that there exists a transition parameter $c(\omega)$ such that the assertion of Theorem 6.6.5(b) is satisfied. Observe that Carleson's assumption that the broken line $1, 2, \dots, d-1$ has no self-intersections has not been needed in these considerations. Also the assumption that $\Omega^{(1)}$ is a regular polygon can be omitted; one can prove that $c(\omega)$ doesnot depend on T_iT_{i+1} by considering a transitive, aperiodic Markov partition which involves all the sides of Ω^1 simultaneously.

Section 7.9. Real analyticity of the density functions. In this section we consider potentials of the form $-t\log|f'|$, fixed points of the corresponding conjugate transfer operators m_t and invariant Gibbs states μ_t. Our aim is to show that the Radon-Nikodym derivative $\frac{d\mu_t}{dm_t}$ has a real-analytic extension. We begin with the following.

Definition 7.9.1. A conformal expanding repeller $f : X \to X$ is said to be real-analytic if it is contained in a finite union of pairwise disjoint real-analytic curves which will be denoted by $\Gamma = \Gamma_f$. Frequently in such a context we will alternatively speak about real analyticity of the set X.

The main (and only) result of this section is the following.

Theorem 7.9.2. If $f : X \to X$ is an orientation preserving conformal expanding repeller, $X \subset C$, then the Radon-Nikodym derivative $\rho = d\mu_t/dm_t$ has a real-analytic real-valued extension on a neighbourhood of X in $\mathbb{C}$. If f is real-analytic, then ρ has a real-analytic extension on a neighbourhood of X in Γ.

Proof. Observe that since f is conformal and orientation preserving, f is holomorphic on a neighbourhood of X in $\mathbb{C}$. Take $r > 0$ so small that for every $x \in X$, every $n \geq 1$ and every $y \in f^{-n}(x)$ the holomorphic inverse branch $f_y^{-n} : B(x, 2r) \to \mathbb{C}$ sending x to y is well-defined. Suppose first that f is real-analytic. We need to show that there exists a holomorphic complex-valued extension of ρ on a neighbourhood of X in $\mathbb{C}$. Taking an appropriate atlas we may assume that X is contained in a real axis (if a closed curve is a component of Γ we can use Arg). For all $k \geq 1$ and all $y \in f^{-k}(x)$ let $\nu(k, y) = 1$ or -1 depending as f_y^{-k} preserves or reverses the orientation on Γ. So

$$|(f_y^{-k})'(z)| = \nu(k, y)((f_y^{-k})'(z))$$

for all $z \in J(f) \cap B(x, r)$. Consider the following sequence of complex analytic functions on $z \in B(x, r)$

$$g_n(z) = \sum_{y \in f^{-n}(x)} \Big(\nu(n, y)((f_y^{-n})'(z))\Big)^t \exp(-nP(t)).$$

There is no problem here with raising to the t-th power since $B(x, r)$, the domain of all $\nu(n, y)(f_y^{-n})'$ is simply connected. Since the latter functions are positive in $\mathbb{R}$, we can choose the branches of the t-th powers to be also positive in $\mathbb{R}$. By Koebe's Distortion Theorem for every $z \in B(x, r/2)$, every $n \geq 1$ and every $y \in f^{-n}(x)$ we have $|(f_y^{-n})'(z)| \leq K|(f_y^{-n})'(x)|$. Hence $|g_n(z)| \leq K g_n(x)$. Since, by (3.4.2) with $u = 1$ and $c = p(t)$, the sequence $g_n(x)$ converges, we see that the functions $\{g_n|_{B(x,r/2)}\}_{n \geq 1}$ are uniformly bounded. So they form a normal family in the sense of Montel. Since $g_n(z)$ converges for all $z \in X \cap B(x, r/2)$, it follows that g_n converges to an analytic function g on $B(x, r/2)$ whose restriction to Γ is by our construction an extension of ρ.

Let us pass now to the proof of the first part of this proposition. That is, we relax the Julia real analyticity assumption and we want to construct a real-analytic real-valued extension of ρ to a neighbourhood of X in $\mathbb{C}$. Our strategy is to work in $\mathbb{C}^2$, to use an appropriate version of Montel's theorem and, in general, to proceed similarly as in the first part of the proof. So, fix $v \in X$. Identify now $\mathbb{C}$, where our f acts, to $\mathbb{R}^2$ with coordinates x, y, the real and complex part of z. Embed this into $\mathbb{C}^2$ with x, y complex. Denote the above

$\mathbb{C} = \mathbb{R}^2$ by $\mathbb{C}_0$. We may assume that $v = 0$ in $\mathbb{C}_0$. Given $k \ge 0$ and $v_k \in f^{-k}(v)$ define the function $\rho_{v_k} : B_{\mathbb{C}_0}(0, 2r) \to \mathbb{C}$ (the ball in $\mathbb{C}_0$) by setting

$$\rho_{v_k}(z) = \frac{(f_{v_k}^{-k})'(z)}{(f_{v_k}^{-k})'(0)},$$

Since $B_{\mathbb{C}_0}(0, 2r) \subset \mathbb{C}_0$ is simply connected and ρ_{v_n} nowhere vanishes, all the branches of logarithm $\log \rho_{v_n}$ are well defined on $B_{\mathbb{C}_0}(0, 2r)$. Choose this branch that maps 0 to 0 and denote it also by $\log \rho_{v_n}$. By Koebe's Distortion Theorem $|\rho_{v_k}|$ and $|\mathrm{Arg}\rho_{v_k}|$ are bouned on $B(0, r)$ by universal constants K_1, K_2 respectively. Hence $|\log \rho_{v_k}| \le K = (\log K_1) + K_2$. We write

$$\log \rho_{v_k} = \sum_{m=0}^{\infty} a_m z^m$$

and note that by Cauchy's inequalities

$$|a_m| \le K/r^m. \tag{7.9.1}$$

We can write for $z = x + iy$ in $\mathbb{C}_0$

$$\mathrm{Re} \log \rho_{v_k} = \mathrm{Re} \sum_{m=0}^{\infty} a_m (x + iy)^m = \sum_{p,q=0}^{\infty} \mathrm{Re}\Big(a_{p+q} \binom{p+q}{q} i^q\Big) x^p y^q := \sum c_{p,q} x^p y^q.$$

In view of (2.1), we can estimate $|c_{p,q}| \le |a_{p+q}| 2^{p+q} \le K r^{-(p+q)} 2^{p+q}$. Hence $\mathrm{Re} \log \rho_{v_k}$ extends, by the same power series expansion $\sum c_{p,q} x^p y^q$, to the polydisc $D_{\mathbb{C}^2}(0, r/2)$ and its absolute value is bounded there from above by K. Now for every $k \ge 0$ consider a real-analytic function b_k on $B_{\mathbb{C}_0}(0, 2r)$ by setting

$$b_k(z) = \sum_{v_k \in f^{-k}(0)} |(f_{v_k}^{-k})'(z)|^t \exp(-kP(t)).$$

By (3.4.2) the sequence $b_k(0)$ is bounded from above by a constant L. Each function b_k extends to the function

$$B_k(z) = \sum_{v_k \in f^{-k}(0)} |(f_{v_k}^{-k})'(0)|^t \mathrm{e}^{t \mathrm{Re} \log \rho_{v_k}(z)} \exp(-kP(t)).$$

whose domain, similarly as the domains of the functions $\mathrm{Re} \log \rho_{v_k}$, contains the polydisc $D_{\mathbb{C}^2}(0, r/2)$. Finally we get for all $k \ge 0$ and all $z \in D_{\mathbb{C}^2}(0, r/4)$

$$\begin{aligned}
|B_k(z)| &= \sum_{v_k \in f^{-k}(0)} |(f_{v_k}^{-k})'(0)|^t \mathrm{e}^{\mathrm{Re}(t \mathrm{Re} \log \rho_{v_k}(z))} \exp(-kP(t)) \\
&\le \sum_{v_k \in f^{-k}(0)} |(f_{v_k}^{-k})'(0)|^t \mathrm{e}^{t |\mathrm{Re} \log \rho_{v_k}(z)|} \exp(-kP(t)) \\
&\le \mathrm{e}^{Kt} \sum_{v_k \in f^{-k}(0)} |(f_{v_k}^{-k})'(0)|^t \exp(-kP(t)) \le \mathrm{e}^{Kt} L.
\end{aligned}$$

Now by Cauchy's integral formula (in $D_{\mathbb{C}^2}(0, r/4)$) for the second derivatives we prove that the family B_n is equicontinuous on, say, $D_{\mathbb{C}^2}(0, r/5)$. Hence we can choose a uniformly convergent subsequence and the limit function G is complex analytic and extends ρ on $X \cap B(0, r/5)$, by (3.4.2). Thus we have proved that ρ extends to a complex analytic function in a neighbourhood of every $v \in X$ in $\mathbb{C}^2$, i.e. real analytic in $\mathbb{C}_0$. These extensions coincide on the intersections of the neighbourhoods, otherwise X is real analytic and we are in the case considered at the beginning of the proof. ♣

CHAPTER 7. SULLIVAN'S CLASSIFICATION OF CONFORMAL EXPANDING REPELLERS.

Dijon, June, 1991

(This is a very preliminary version of one of chapters of a book by Przytycki and Urbański, in preparation, on conformal fractals. It relies on ideas of the proof of the rigidity theorem drafted by D. Sullivan in Proceedings of Berkeley's ICM in 1986.)

In Chapter 4.6 we proved that the scaling function for an expanding repeller in the line determines the $C^{1+\varepsilon}$-structure. In this chapter we will basically concentrate on nonlinear conformal expanding repellers, called CER's, proving that the class of equivalence of the geometric measure determines the conformal structure.

Section 1. Equivalent notions of linearity.

Definition 7.1.1. We call a CER (X, f) *linear* if one of the following conditions holds:

a) The Jacobian of f with respect to the Gibbs measure μ_X equivalent to a geometric measure m_X on X , Jf, is locally constant.

b) The function $\mathrm{HD}(X) \log |f'|$ is cohomologous to a locally constant function on X.

c) The conformal structure on X admits a conformal affine refinement so that f is affine (i.e., see Ch.4.3, there exists an atlas $\{\varphi_t\}$ that is a family of conformal injections $\varphi_t : U_t \to \mathbb{C}$ where $\bigcup_t U_t \supset X$ such that all the maps $\varphi_t \varphi_s^{-1}$ and $\varphi_t f \varphi_s^{-1}$ are affine)

Recall that as the conformal map f may change the orientation of $\mathbb{C}$ on some components of its domain we can write $|f'|$ but not f' unless f is holomorphic.

Proposition 7.1.2. The conditions a), b) and c) are equivalent.

Before we shall prove this proposition we distinguish among CER's *real- analytic* repellers:

Definition 7.1.3. We call (X, f) *real- analytic* if X is contained in the union of a finite family of real analytic open arcs and closed curves.

Lemma 7.1.4. If there exists a connected open domain U in $\mathbb{C}$ intersecting X for a CER (X, f) and if there exists a real analytic function k on it equal identically 0 on $U \cap X$ but not on U then (X, f) is real-analytic.

Proof. Pick an arbitrary $x \in U \cap X$. Then in a neighbourhood V of x the set $E = \{k = 0\}$ is a finite union of pairwise disjoint real- analytic curves and of the point x. This follows from the existence of a finite decomposition of the germ of E at x into irreducible germs and the form of each such germ, see for example Proposition 5.8 in the Malgrange book [Malgrange]. As the sets $f^n(X \cap V)$, $n \geq 0$ cover X, X is compact and f is open on X we conclude that X is contained in a finite union of real-analytic curves γ_j and a finite set of points A such that the closures of γ_j can intersect only in A.

Suppose that there exists a point $x \in X$ such that X is not contained in any real-analytic curve in every neighbourhood of x. Then the same is true for every point $z \in X \cap f^{-n}\{x\}$, $n \geq 0$, hence for an infinite number of points (because pre-images of x are dense in X by the topological exactness of f, see Ch.?). But we proved above that the number of such points is finite so we arived at a contradiction. We conclude that X is contained in a 1-dimensional real-analytic submanifold of $\mathbb{C}$.

Proof of Proposition 7.1.2.

a)$\Rightarrow$ b). Let u be the eigenfunction $\mathcal{L}u = u$ for the transfer operator $\mathcal{L} = \mathcal{L}_\varphi$ for the function $\varphi = -\kappa \log |f'|$, where $\kappa = \mathrm{HD}(X)$, as in Ch. 3.3.. Here the eigenvalue $\lambda = \exp P(f, \varphi)$ is equal to 1, see Ch.6.2..

For an arbitrary $z \in X$ we have in its neighbourhood in X

$$\text{Const} = \log Jf = \kappa \log |f'(x)| + \log u(f(x)) - \log u(x) \tag{7.1.1}$$

b)$\Rightarrow$ c). The function u extends to a real-analytic function $u_{\mathbb{C}}$ in a neighbourhood of X, see Ch.4.4, so the function $\log Jf$ extends to a real-analytic function $\log Jf_{\mathbb{C}}$ by the right hand side equality in the formula (7.1.1), for $u_{\mathbb{C}}$ instead of u. We have two cases: either $\log Jf_{\mathbb{C}}$ is not locally constant on every neighbourhood of X and then by Lemma 7.1.4 (X, f) is real-analytic or $\log Jf_{\mathbb{C}}$ is locally constant. Let us consider first the latter case.

Fix $z \in X$. Choose an arbitrary sequence of points $z_n \in X$, $n \geq 0$ such that $f(z_n) = z_{n-1}$ and choose branches f_ν^{-n} mapping z to z_n. Due to the expanding property of f they are all well defined on a common domain around z. For every x close to z denote $x_n = f_\nu^{-n}(x)$. We have $\mathrm{dist}(x_n, z_n) \to 0$ so by (7.1.1) for $\log Jf_{\mathbb{C}}$

$$\begin{aligned} &\sum_{n=1}^{\infty} \kappa(\log |f'(x_n)| - \log |f'(z_n)|) \\ &= \log u_{\mathbb{C}}(x) - \log u_{\mathbb{C}}(z) + \lim_{n\to\infty} (\log u_{\mathbb{C}}(z_n) - \log u_{\mathbb{C}}(x_n)) \\ &= \log u_{\mathbb{C}}(x) - \log u_{\mathbb{C}}(z) \end{aligned} \tag{7.1.2}$$

We conclude that $u_{\mathbb{C}}(x)$ is a harmonic function in a neighbourhood of z in $\mathbb{C}$ as the limit of a convergent series of harmonic functions; we use the fact that the compositions of harmonic functions with the conformal maps f_ν^{-n} are harmonic. Close to z we take a so-called harmonic conjugate function h so that $\log u(x) + ih(x)$ is holomorphic.

Write $F_z = \exp(\log u + ih)$ and denote by $\tilde{F}_z$ a primitive function for F_z in a neighbourhood of z. This is a chart because $F_z(z) \neq 0$. The *atlas* given by the charts $\tilde{F}_z$ is affine (conformal) by the construction. We have due to (7.1.1) for the extended u

$$|(\tilde{F}_{f(z)} \circ f \circ \tilde{F}_z^{-1})'(F_z(x))| = u_{\mathbb{C}}(f(x))|f'(x)|/u_{\mathbb{C}}(x) = \text{Const}$$

so the differential of f is locally constant in our atlas.

In the case (X, f) is real-analytic we consider just the charts φ_t being primitive functions of u on real-analytic curves containing X into $\mathbb{R}$ with unique complex extensions to neighbourhoods of these curves into a neighbourhood of $\mathbb{R}$ in $\mathbb{C}$. The equality $\log Jf_{\mathbb{C}} = \text{Const}$ holds on these curves so the derivatives of $\varphi_t f \varphi_s^{-1}$ are locally constant.

c)⇒ a). Denote the maps $\varphi_t f \varphi_s^{-1}$ by $\tilde{f}_{t,s}$. In a neighbourhood (in X) of an arbitrary $z \in X$ we have

$$
\begin{aligned}
u(x) &= \lim_{n\to\infty} \mathcal{L}^n(1)(x) = \lim_{n\to\infty} \sum_{y\in f^{-n}(x)} |(f^n)'(y)|^{-\kappa} \\
&= \lim_{n\to\infty} |\varphi'(x)|^{\kappa} \sum_y |\varphi'(y)|^{-\kappa} |\tilde{f}'(y)|^{-n\kappa} \\
&= \text{Const} \lim_{n\to\infty} |\varphi'(x)|^{\kappa} \sum_y \tilde{f}'(y)|^{-n\kappa} = |\varphi'(x)|^{\kappa}\text{Const}
\end{aligned}
\tag{7.1.3}
$$

To simplify the notation we omitted the indices at φ and $\tilde{f}$ here, of course they depend on z and y's more precisely on the branches of f^{-n} on our neighbourhood of z mapping z to y's . Const also depends on z. We could omit the functions $\varphi'(y)$ in the last line of (7.1.3) because the diameters of the domains of $\varphi'(y)$ which were involved converged to 0 when $n \to \infty$ due to the expanding property of f, so these functions were almost constant.

Hence due to (7.1.3) in a neighbourhood of every $x \in X$ we get

$$
Jf(x) = \text{Const } u(f(x))|f'(x)|^{\kappa}/u(x) = \text{Const}|\tilde{f}'(x)|^{\kappa} = \text{Const}
$$

♣

Remark 7.1.5. In the b)⇒c) part of the proof of Proposition (7.1.2) as $-\kappa \log |f'|$ is harmonic we do not need to refer to Ch.4.4 for the real-analyticity of u . The formula (7.1.2) gives a harmonic extension of u to a neighbourhood of an arbitrary $z \in X$, depending on the choice of the sequence (z_n). If two extensions u_1, u_2 do not coincide on a neighbourhood of z then in a neighbourhood of z, $X \subset \{u_1 - u_2 = 0\}$. If the equation (7.1.1) does not extend to a neighbourhood of z then again $X \subset \{v = \text{Const}\}$ for a harmonic function v extending the right hand side of (7.1.1).

In each of the both cases (X, f) happens to be real-analytic and to prove it we do not need to refer to Malgrange's book as in the proof of Lemma 7.1.3. Indeed, for any non-constant harmonic function k on a neighbourhood of $x \in X$ such that $X \subset \{k = 0\}$ we consider a holomorphic function F such that $k = \text{Re}F$ and $F(x) = 0$. Then $E = \{k = 0\} = \{\text{Re}F = 0\}$. If F has a d-multiple zero at x then it is a standard fact that E is a union of d analytic curves intersecting at x within the angle $\frac{\pi}{d}$.

We end this Section with giving one more condition implying the linearity.

Lemma 7.1.6. Suppose for a CER (X, f) that there exists a Hölder continuous line field in the tangent bundle on a neighbourhood of X invariant under the differential of f. In other words there exists a complex valued nowhere zero Hölder continuous function α such that for every x in a neighbourhood of X

$$\text{(7.1.4)} \qquad \mathrm{Arg}\alpha(x) + \mathrm{Arg}f'(x) = \mathrm{Arg}\alpha(f(x)) + \varepsilon(x)\pi$$

where $\varepsilon(x)$ is a locally constant function equal 0 or 1. This is in the case f preserves the orientation at x, if it reverses the orientation we replace in (7.2.1) $\mathrm{Arg}f'$ by $-\mathrm{Arg}\bar{f}'$.

Then (X, f) is linear.

Proof. As in Proof of Proposition 7.1.2, the calculation (7.1.2), if f is holomorphic we have for x in a neighbourhood of $z \in X$ in $\mathbb{C}$

$$\mathrm{Arg}\alpha(z) - \mathrm{Arg}\alpha(x) = \sum_{n=1}^{\infty}(\mathrm{Arg}(f'(z_n)) - \mathrm{Arg}(f'(x_n))),$$

if we allow f to reverse the orientation then we replace $\mathrm{Arg}f'$ by $-\mathrm{Arg}\bar{f}'$ in the above formula for such n that f changes the orientation in a neighbourhood of x_n. So $\mathrm{Arg}\alpha(x)$ is a harmonic function. Close to z we find a conjugate harmonic function h so we get a family of holomorphic functions $F_z = \exp(-h + i\mathrm{Arg}\alpha$ which primitive functions give an atlas we have looked for.

Remark 7.1.7. The condition for (X, f) in Lemma 7.1.6 is stronger than the linearity property. Indeed we can define f on the union of the discs $D_1 = \{|z| < 1\}$ and $D_2 = \{|z-3| < 1\}$ by $f(z) = 5\exp 2\pi\vartheta i$ on D_1 where ϑ is irrational, and $f(z) = 5(z-3)$ on D_2. This is an example of an iterated function system from Ch.4.5. We get a CER (X, f) where $X = \bigcap_{n=0}^{\infty} f^{-n}(\{|z| < 5\})$. It is linear because it satisfies the condition c). Meanwhile $0 \in X$, $f(0) = 0$ and $f'(0) = 5\exp 2\pi\vartheta i$, so the equation (7.1.4) has no solution at $x = 0$ even for any iterate of f.

Remark 7.1.8. If we assume in place of (7.1.4) that $\mathrm{Arg}f'(x) - \mathrm{Arg}\alpha(f(x)) - \mathrm{Arg}\alpha(x)$ is locally constant, then we get the condition equivalent to the linearity.

Section 2. Rigidity of nonlinear CER's

In this section we shall prove the main theorem of Chapter 7:

Theorem 7.2.1. Let $(X, f), ((Y, g)$ be two non-linear conformal expanding repellers in $\mathbb{C}$. Let h be an invertible mapping from X onto Y preserving Borel σ-algebras and conjugating f to g, $h \circ f = g \circ h$. Suppose that one of the following assumption is satisfied:

1. h and h^{-1} are Lipschitz continuous.

2. h and h^{-1} are continuous and preserve so-called Lyapunov spectra, namely for every periodic $x \in X$ and integer n such that $f^n(x) = x$ we have $|(f^n)'(x)| = |(g^n)'(h(x))|$.

3. h_* maps a geometric measure m_X on X to a measure equivalent to a geometric measure m_Y on Y.

Then h extends from X (or from a set of full measure m_X in the case 3.) to a conformal homeomorphism on a neighbourhood of X.

We start the proof with a discussion of the assumptions. The equivalence of the conditions 1. and 2. has been proved in Ch.4.3. The condition 1. implies 3. by the definition of geometric measures 5.6.5. One of the steps of the proof of Theorem will assert that 3. implies 1. under the non-linearity assumption. Without this assumption the assertion may happen false. A positive result is that if h is continuous then for a constant $C > 0$ and every $x_1, x_2 \in X$

$$C < \frac{|h(x_1) - h(x_2)|^{\mathrm{HD}(Y)}}{|x_1 - x_2|^{\mathrm{HD}(X)}} < C^{-1}.$$

(We leave the proof to the reader.)

It may happen that $\mathrm{HD}(X) \neq \mathrm{HD}(Y)$ for example if X is a 1/3 – Cantor set and for g we remove each time half of the interval from the middle.

A basic observation to prove Theorem 7.2.1 is that

$$Jg \circ h = Jf \text{ and moreover } Jg^j \circ h = Jf^j \tag{7.2.1}$$

for every integer $j > 0$. This follows from $g^j \circ h = h \circ f^j$ and $Jh \equiv 1$. We recall that we consider Jacobians with respect to the Gibbs measures equivalent to geometric measures.

Observe finally that (X,f) linear implies (Y,g) linear. Indeed, if (X, f) is linear then Jf hence Jg admit only a finite number of values in view of $Jg \circ h = Jf$. As Jg is continuous this implies that Jg is locally constant i.e. (Y, g) is linear.

Lemma 7.2.2. If a CER (X, f) is non-linear then there exists $x \in X$ such that $\mathrm{grad} Jf_{\mathbb{C}}(x) \neq 0$.

Proof. If $\mathrm{grad} Jf_{\mathbb{C}} \equiv 0$ on X then as $Jf_{\mathbb{C}}$ is real- analytic we have either $\mathrm{grad} Jf_{\mathbb{C}} \equiv 0$ ona neighbourhood of X in $\mathbb{C}$ or by Lemma 7.1.4 (X, f) is real-analytic and $\mathrm{grad} Jf_{\mathbb{C}} \equiv 0$ on real- analytic curves containing X. In both cases by integration we obtain Jf locally constant on X what contradicts the non-linearity assumption.

Now we can prove Theorem in the simplest case to show the reader the main idea working later also in the general case.

Proposition 7.2.3. The assertion of Theorem 7.2.1 holds if we suppose additionally that (X, f) and $Y, g)$ are real-analytic and the conjugacy h is continuous.

Proof. Let M, N be real analytic manifolds containing X, Y respectively. By the non-linearity of X and Lemma 7.2.2 there exists $x \in X$ and its neighbourhood U in M such that $F := Jf_{\mathbb{C}}|_U : U \to \mathbb{R}$ has a real-analytic inverse $F^{-1} : F(U) \to U$. Then in view of (7.2.1) $h^{-1} = F^{-1} \circ Jg_{\mathbb{C}}$ on $h(U \cap X)$ so h^{-1} on $h(U \cap X)$ extends to a real analytic map on a neighbourhood of $h(U \cap X)$ in N.

Now we use the assumption that h^{-1} is continuous so $h(U \cap X)$ contains an open set v in Y. There exists a positive integer n such that $g^n(V) = Y$ hence for every $y \in Y$

there exists a neighbourhood W of y in N such that a branch g_ν^{-n} of g^{-n} mapping y and even $W\cap Y$ into V is well defined. So we have $h^{-1}=f^n\circ h^{-1}\circ g_\nu^{-n}$ extended on W to a real-analytic map. This gives a real-analytic extension of h^{-1} on a neighbourhood of Y because two such extensions must coincide on the intersections of their domains by the real-analyticity and the fact that Y has no isolated points.

Similarly using the non-linearity of (Y,g) and the continuity of h we prove that h extends analytically. By the analyticity and again lack of isolated points in X and Y the extentions are inverse to each other, so h extends even to a biholomorphic map.

Now we pass to the general case.

Lemma 7.2.4. Suppose that there exists $x\in X$ such that $\mathrm{grad}Jf_{\mathbb{C}}(x)\neq 0$ in the case X is real-analytic, or there exists an integer $k\geq 1$ such that $\det(\mathrm{grad}Jf_{\mathbb{C}},\mathrm{grad}(Jf_{\mathbb{C}}\circ f^k))\neq 0$ in the other case .

(In other words we suppose that $Jf_{\mathbb{C}}$, respect. $(Jf_{\mathbb{C}},Jf_{\mathbb{C}}\circ f^k)$, give a coordinate system on a real, respect. complex neighbourhood of x.)

Suppose the analogous property for (Y,g).

Let $h:X\to Y$ satisfy the property 3. assumed in Theorem 7.2.1. Then h extends from a set of full geometric measure in X to a bi-Lipschitz homeomorphism of X onto Y conjugating f with g.

Proof. We can suppose that $\mathrm{HD}(X)\geq\mathrm{HD}(Y)$, recall that HD denotes Hausdorff dimension . Pick x with the property assumed in the Lemma. Let U be its neighbourhood in M (as in Proof of Proposition 7.2.3) or in $\mathbb{C}$ if (X,f) is not real-analytic, so that $F:=(Jf_{\mathbb{C}},Jf_{\mathbb{C}}\circ f^k)$ is an embedding on U. Let $y\in Y$ be a density point of the set $h(U\cap X)$ with respect to the Gibbs measure μ_Y equivalent to the geometric measure m_Y. (Recall that we have proved that almost every point is a density point for an arbitrary probability measure on a euclidean space in Ch.5.2 relying on Besicovitch's Theorem.) . So if we denote $(Jg_{\mathbb{C}},Jg_{\mathbb{C}}\circ g^k)$ in a neighbourhood (real or complex) of y by G, we have for every $\delta>0$ such $\varepsilon_0=\varepsilon_0(\delta)>0$ that for every $0<\varepsilon<\varepsilon_0$:

$$\frac{\mu_Y(B(y,\varepsilon)\cap h(U\cap X))}{\mu_Y(B(y,\varepsilon))}>1-\delta$$

and

$$h^{-1}=F^{-1}\circ G\qquad\text{on}\quad h(U\cap X).$$

(Observe that the last equality may happen false outside $h(U\cap X)$ even very close to y because h^{-1} may map such points to $(Jf_{\mathbb{C}},Jf_{\mathbb{C}}\circ f^k)^{-1}\circ G$ with a branch of $(Jf_{\mathbb{C}},Jf_{\mathbb{C}}\circ f^k)^{-1}$ different from F^{-1}.)

Now for every $\varepsilon>0$ small enough there exists an integer n such that $\mathrm{diam}g^nB(y,\varepsilon)$ is greater than a positive constant , $g^n|_{B(y,\varepsilon)}$ is injective and the distortion of g^n on $B(y,\varepsilon)$ is bounded by a constant C, both constants depending only on (Y,g). Then if $\varepsilon<\varepsilon_0(\delta)$ we obtain for $Y_\delta:=g^n(h(U\cap X)\cap B(y,\varepsilon))$,

$$\frac{\mu_Y(g^n(B(y,\varepsilon))\setminus Y_\delta}{\mu_Y(g^n(B(y,\varepsilon)))}<C\frac{\mu_Y(B(y,\varepsilon)\setminus h(U\cap X))}{\mu_Y(B(y,\varepsilon))}<C\delta.$$

So

$$(7.2.2) \qquad \frac{\mu_Y(Y_\delta)}{\mu_Y(g^n(B(y,\varepsilon)))} > 1 - C\delta$$

We have

$$|(f^n)'(h{-}1(y))|^{\mathrm{HD}(X)} \le \mathrm{Const}\, Jf(h{-}1(y)) = \mathrm{Const}\, Jg(y) \le \mathrm{Const}|(f^n)'(y)|^{\mathrm{HD}(Y)}.$$

As we assumed $\mathrm{HD}(X) \ge \mathrm{HD}(Y)$ we obtain

$$(7.2.3) \qquad |(f^n)'(h{-}1(y))| \le \mathrm{Const}|(f^n)'(y)|^{\mathrm{HD}(Y)/\mathrm{HD}(X)} \le \mathrm{Const}|(f^n)'(y)|$$

Then due to the bounded distortion property for iteration of f and g we obtain that $h^{-1} = f^n h^{-1} g^{-1}$ is Lipschitz on Y_δ with Lipschitz constant independent of δ, more precisely bounded by $\mathrm{Const}\sup \|D(F^{-1} \circ G\|$, where $F^{-1} \circ G$ is considered on a real (complex) neighbourhood of y and Const is that from (7.2.3).

There exists an integer $K > 0$ such that for every n, $g^K g^n B(y, \varepsilon(n))$ covers Y. Because Jg is bounded, separated from 0, this gives h^{-1} on $g^K(Y_\delta)$ Lipschitz with a Lipschitz constant independent from δ and $\mu(g^K(Y_\delta)) > 1 - \mathrm{Const}\,\delta$ for δ arbitrarily small. We conclude that h^{-1} is Lipschitz on a set of full measure μ_Y so it has a Lipschitz extension to Y.

We conclude also that $\mathrm{HD}(X) = \mathrm{HD}(Y)$. Otherwise $\mathrm{diam} h^{-1}(Y_\delta) \to 0$, so because $\mathrm{supp}\mu_X = X$ we would get $\mathrm{diam} X = 0$. So we can replace above the roles of (X, f) and (Y, g) and prove that h is Lipschitz.

The next step will assert that for non-linear repellers the assumptions of Lemma 7.2.4 about the existence of coordinate systems are satisfied.

Lemma 7.2.5. If (X,f) is a non-linear CER then there exists $x \in X$ such that either $\mathrm{grad} Jf_{\mathbb{C}}(x) \ne 0$ in the case X is real-analytic, or there exists an integer $k \ge 1$ such that $\det(\mathrm{grad} Jf_{\mathbb{C}}, \mathrm{grad}(Jf_{\mathbb{C}} \circ f^k)) \ne 0$ in the case (X,f) is not real-analytic.

Proof. We know already from Lemma 7.2.2 that there exists $\hat{x} \in X$ such that $\mathrm{grad} Jf_{\mathbb{C}}(z) \ne 0$ so we may restrict our considerations to the case (X, f) is not real-analytic.

Suppose Lemma is false. Then for all $k > 0$ the functions

$$\Phi_k := \det(\mathrm{grad} Jf_{\mathbb{C}}, \mathrm{grad}(Jf_{\mathbb{C}} \circ f^k))$$

are identically equal to 0 on X. Let W be a neighbourhood of $\hat{x}$ in $\mathbb{C}$ where $\mathrm{grad} Jf_{\mathbb{C}} \ne 0$.

Let us consider on W the line field $\mathcal{V}$ orthogonal to $\mathrm{grad} Jf_{\mathbb{C}}$. Due to the topological exactness of f on X for every $x \in X$ there exists $y \in W \cap X$ and $n \ge 0$ such that $f^n(y) = x$.

Thus define at x

(7.2.4) $$\mathcal{V}_x := Df^n(\mathcal{V}_y)$$

We shall prove now that if $x = f^k(y) = f^l(z)$ for some $y, z \in W \cap X,\ k, l \geq 0$, then

(7.2.5) $$Df^k(\mathcal{V}_y) = Df^l(\mathcal{V}_z).$$

If (7.2.5) is false, then close to x there exist $x' \in X$ and $m \geq 0$ such that $f^m(x') \in W$ (we again refer to the topological exactness of f) and $Df^k(\mathcal{V}_{y'}) \neq Df^l(\mathcal{V}_{z'})$, where $f^k(y') = f^l(z') = x'$, $y' \in X$ is close to y and $z' \in X$ is close to z. We obtain $Df^{k+m}(\mathcal{V}_{y'}) \neq Df^{l+m}(\mathcal{V}_{z'})$ so either $Df^{k+m}(\mathcal{V}_{y'}) \neq \mathcal{V}_{f^m(x')}$ or $Df^{l+m}(\mathcal{V}_{z'}) \neq \mathcal{V}_{f^m(x')}$. Consider the first case (the second is of course similar). We obtain that Jf and $Jf \circ f^{k+m}$ give a coordinate system in a neighbourhood of y' i.e. $\Phi_{k+m}(y') \neq 0$ contrary to the supposition.

Thus the formula (7.2.4) defines a line field at all points of X which is Df-invariant. Observe however that the same formula defines a real-analytic extension of the line field to a neighbourhood of x in $\mathbb{C}$ because $\mathcal{V}$ is real-analytic on a neighbourhood of $y \in W$ and f is analytic. Each two such germs of extensions related to two different pre-images of x must coincide because they coincide on X, otherwise (X, f) would be real-analytic. Now we can choose a finite cover $B_j = B(x_j, \delta_j)$ of a neighbourhood of X with discs, $x_j \in X$ so that for the respective F_j-branches of f^{-n_j} leading x_j into W, we have $F_j(3B_j) \subset W$ where $3B_j := B(x_j, 3\delta_j)$. Hence the formula (7.2.4) defines $\mathcal{V}$ on $3B_j$. So if $B_i \cap B_j \neq \emptyset$, then we have $3B_i \subset B_j$ or vice versa. So $3B_i \cap 3B_j \cap X \neq 0$ hence the extensions of $\mathcal{V}$ on $3B_i$ and on $3B_j$, in particular on B_i and on B_j, coincide on the intersection. This is so because they coincide on the intersection with X and $(X, f$ is not real-analytic.
(We made the trick with 3δ because it can happen that $B_i \cap B_j \neq \emptyset$ but $B_i \cap B_j \cap X = \emptyset$.)

Thus $\mathcal{V}$ extends real-analytically to a neighbourhood of X. This field is Df-invariant on a neighbourhood of X because we can define it in a neighbourhood of $x \in X$ and $f(x)$ by (7.2.4) taking the same $y \in W \cap X$ where $f^n(y) = x$, $f^{n+1}(y) = f(x)$. So by Lemma 7.1.7 (X, f) is linear what contradicts the assumption that (X, f) is non-linear.

Corollary 7.2.6. If for $(X, f), (Y, g)$ the assumptions of Theorem 7.2.1 are satisfied and if (y, g) is real-analytic then (X, f) is real-analytic too.

Proof. Due to Lemma 7.2.5 the assumptions of Lemma 7.2.4 are satisfied. So $h^{-1} = F^{-1} \circ G$ on a neighbourhood of $y \in Y$ by the continuity of h^{-1}, (see the notation in Proof of Lemma 7.2.4). Denote a real-analytic manifold Y is contained in by N. Then $Jg_{\mathbb{C}} \neq \mathrm{Const}$ on any neighbourhood of y in N. Otherwise h^{-1} would be constant, but y is not isolated in Y so h^{-1} would not be injective.

Remind that we can consider $F^{-1} \circ G$ as a real analytic extension of h^{-1} to a neighbourhood V of y in N. So the differential of $F^{-1}G$ is 0 at most at isolated points, so different from 0 at a point $y' \in V \cap Y$. We conclude due to the continuity of h that in a neighbourhood of $h^{-1}(y')$, X is contained in a real-analytic curve. So (X, f) is a real-analytic repeller.

Now we shall collect together what we have done and make a decidive step in proving Theorem 7.2.1, namely we shall prove that the conjugacy extends to a real-analytic diffeomorphism.

Proof of Theorem 7.2.1. If both (X, f) and (Y, g) are real- analytic then the conjugacy extends real-analytically to a real-analytic manifold so complex analytically to its neighbourhood by Proposition 7.2.3. Its assumptions hold by Lemmas 7.2.4 and 7.2.2. If both (X, f) and (Y, g) are not real-analytic (a mixed situation is excluded by Corollary 7.2.6), then by Lemma 7.2.4 which assumptions hold due to Lemma 7.2.5 we can assume the conjugacy h is a homeomorphism of X onto Y. But h^{-1} extends to a neighbourhood of $y \in Y$ in $\mathbb{C}$ to a real-analytic map. We use here again the notation of Proposition 7.2.4 and proceed precisely like in Proposition 7.2.3, Proposition 7.2.4 and Corollary 7.2.6 by writing $h^{-1} = F^{-1} \circ G$. This gives a real-analytic extension of h^{-1} to a neighbourhood of an arbitrary $y \in Y$ by the formula $f^n \circ h^{-1} \circ g_\nu^{-1}$ precisely as in Proof of Proposition 7.2.3.

For two different branches F_1, F_2 of g^{-n_1}, g^{-n_2} respectively, mapping y into the domain of $F^{-1} \circ G$ germs of the extensions must coincide because they coincide on the intersection with Y, see Lemma 7.1.4.

Now we build a real-analytic extension of h^{-1} to a neighbourhood of Y similarly as we extended $\mathcal{V}$ in Proof of Lemma 7.2.5, again using the assumption (Y, g) is not real-analytic.Similarly we extend h.

Denote the extensions by $\tilde{h}, \tilde{h^{-1}}$. We have $\tilde{h^{-1}} \circ \tilde{h}$ and $\tilde{h} \circ \tilde{h^{-1}}$ equal to the identity on X, Y respectively. The these compositions extend to the identities to neighbourhoods, otherwise (X, f) or (Y, g) would be real-analytic. We conclude that $\tilde{h}$ is a real-analytic diffeomorphism. Finally observe that $g\tilde{h} = \tilde{h} f$ on a neighbourhood of X because this equality holds on X itself and our functions are real-analytic, otherwise (X, f) would be real-analytic.

The only thing we should still prove is the following

Lemma 7.2.7. If (X, f) is a non-linear CER, not real-analytic , and there is a real-analytic diffeomorphism h on a neighbourhood of X to a neighbourhood of Y for another CER (Y, g) such that $h(X) = Y$ and h conjugates f with g in a neighbourhood of X then h is conformal.

Proof. Suppose for the simplification that f, g and h preserve the orientation of $\mathbb{C}$, we will comment the general case at the end.

For any orientation preserving diffeomorphism Φ of a domain in $\mathbb{C}$ into $\mathbb{C}$ denote the complex dilatation function by ω_Φ . We recall that $\omega_\Phi := \frac{d\Phi}{d\bar{z}} / \frac{d\Phi}{dz}$. (The reader not familiar with the complex dilatation and its properties is advised to read the first 10 pages of the classical Ahlfors book [Ahlfors].) The geometric meaning of the argument of $\omega_\Phi(z)$ may be explained by the equality $\frac{1}{2}\omega_\Phi = \alpha$ where α corresponds to the the direction in which the differential $D\Phi$ at z attains its maximum. In another words it is the direction of the smaller axis of the ellipse in the tangent space at z which is mapped by $D\Phi$ to the unit circle. Of course this makes sense if $\omega(z) \neq 0$. Observe finally that $\omega(z) = 0$ iff $\frac{d\Phi}{d\bar{z}} = 0$. Let go back now to our concrete maps.

If $\frac{dh}{d\bar{z}} \equiv 0$ on X then as $\frac{dh}{d\bar{z}}$ is a real-analytic function we have $\frac{dh}{d\bar{z}} \equiv 0$ on a neighbourhood of X, otherwise (X, f) would be real-analytic. But this means that h is holomorphic what proves our Lemma. It rests to prove that the case $\frac{dh}{d\bar{z}} \not\equiv 0$ on X is impossible.

Observe that if $\frac{dh}{d\bar{z}}(x) = 0$ then $\frac{dh}{d\bar{z}}(f(x)) = 0$ because $h = ghf_\nu^{-1}$ on a neighbourhood of $f(x)$ for the branch f_ν^{-1} of f^{-1} mapping $f(x)$ to x and because g and f_ν^{-1} are conformal. So if there exists $x \in X$ such that $\frac{dh}{d\bar{z}}(x) \neq 0$ then this holds also for all x's from a neighbourhood and as a consequence of the topological exactness of f for all x in a neighbourhood of X. Thus we have a complex-valued function ω_h nowhere zero on a neighbourhood of X.

Recall now that for any two orientation preserving diffeomorphisms Φ and Ψ, if Ψ is holomorphic then

$$\omega_{\Psi\circ\Phi} = \omega_\Phi$$

and if Φ is conformal then

$$\omega_\Psi \circ \Phi = (\frac{\Phi'}{|\Phi'|})^2 \omega_{\Psi\circ\Phi} = \omega_\Phi$$

Applying it to the equation $h \circ f = g \circ h$ we obtain

$$\omega_h \circ f = (\frac{f'}{|f'|})^2 \omega_{h\circ f} = (\frac{f'}{|f'|})^2 \omega_{g\circ h} = (\frac{f'}{|f'|})^2 \omega h.$$

Thus $\alpha(x) := \frac{1}{2}\omega_h(x)$ satisfies the equation (7.1.4) and by Lemma (7.1.6) (X, f) happens linear what contradics our assumption that it is non- linear.

In the case a diffeomorphism reverses the orientation we write everywhere above $\omega_{\bar{\Phi}}$ instead of ω_Φ and if Φ is conformal reversing orientation we write $\overline{\Phi'}$ instead of Φ'. Additionally some omegas should be conjugated in the formulas above. We also arive at (7.1.4). (In this situation the complex notation is not confortable. Everythig gets trivial if we act with differentials on line fields. We leave writing this down to the reader.)

version of Nov.25, 2002

CHAPTER 9
CONFORMAL MAPS WITH INVARIANT PROBABILITY MEASURES OF POSITIVE LYAPUNOV EXPONENT.

§9.1 RUELLE'S INEQUALITY.

Let X be a compact subset of the closed complex plane $\overline{\mathbb{C}}$ and let $\mathcal{A}(X)$ denote the set of all continuous maps $f : X \to X$ that can be analytically extended to an open neighbourhood $U(f)$ of X. In this section we only work with the standard spherical metric on $\overline{\mathbb{C}}$, normalized so that the area of $\overline{\mathbb{C}}$ is 1. In particular all the derivatives are computed with respect to this metric.

Let us recall and extend Definition 6.1.3. Let μ be an f-invariant Borel probability measure on X. Since $|f'|$ is bounded, the integral $\int \log|f'|\,d\mu$ is well-defined and moreover $\int \log|f'|\,d\mu < +\infty$. The number

$$\chi_\mu = \chi_\mu(f) = \int \log|f'|\,d\mu$$

is called the *Lyapunov characteristic exponent* of μ and f. Note that $\int \log|f'|d\mu = -\infty$ is not excluded. In fact it is possible, for example if $X = \{0\}$ and $f(z) = z^2$.

By Birkhoff Ergodic Theorem (Th 1.2.2) the *Lyapunov characteristic exponent* $\chi_\mu(x) = \lim_{n\to\infty} \frac{1}{n}\log|(f^n)'(x)|$ exists for a.e. x, compare Sec.6, and $\int \chi_\mu(x)\,d\mu(x) = \chi_\mu$. (In fact one allows $\log|f'|$ with integral $-\infty$ here, so one need extend slightly Th.1.2.2. This is not difficult.)

The section is devoted to prove the following.

Theorem 9.1.1. (Ruelle's inequality) If $f \in \mathcal{A}(X)$, then $\mathrm{h}_\mu(f) \le 2\int \max\{0, \chi_\mu(x)\}\,d\mu$. For ergodic μ this yields $\mathrm{h}_\mu(f) \le 2\max\{0, \chi_\mu\}$.

Proof. Consider a sequence of positive numbers $a_k \searrow 0$, and $\mathcal{P}_k$, $k = 1, 2, \ldots$ an increasing sequence of partitions of the sphere $\overline{\mathbb{C}}$ consisting of elements of diameters $\le a_k$ and of (spherical) areas $\ge \frac{1}{4}a_k^2$. Check that such partitions exist.

For every $g \in \mathcal{A}(X)$, $x \in X$ and $k \ge 1$ let

$$N(g, x, k) = \#\{P \in \mathcal{P}_k : g(P_k(x) \cap U(g)) \cap P \ne \emptyset\}$$

Our first aim is to show that for every $k > k(g)$ large enough

$$N(g, x, k) \le 4\pi(|g'(x)| + 2)^2 \tag{9.1.1}$$

Indeed, fix $x \in X$ and consider k so large that $\mathcal{P}_k(x) \subset U(g)$ and a Lipschitz constant of $g|_{\mathcal{P}_k(x)}$ does not exceed $|g'(x)|+1$. Thus the set $g(\mathcal{P}_k(x))$ is contained in the ball $B(g(x),(|g'(x)|+1)a_k)$. Therefore if $g(\mathcal{P}_k(x)) \cap P \neq \emptyset$, then

$$P \subset B(g(x),(|g'(x)|+1)a_k + a_k) = B(g(x),(|g'(x)|+2)a_k)$$

Hence $N(g,x,k) \leq \pi(|g'(x)|+2)^2 a_k^2 / \frac{1}{4}a_k^2 = 4\pi(|g'(x)|+2)^2$ and (9.1.1) is proved.

Let $N(g,x) = \sup_{k>k(g)} N(g,x,k)$. In view of (9.1.1) we get

$$N(g,x) \leq 4\pi(|g'(x)|+2)^2 \tag{9.1.2}$$

Now note that for every finite partition $\mathcal{A}$ one has

$$\mathrm{h}(g,\mathcal{A}) = \lim_{n\to\infty} \frac{1}{n+1}\mathrm{H}(\mathcal{A}^n) = \lim_{n\to\infty} \frac{1}{n+1}\Big(\mathrm{H}(g^{-n}(\mathcal{A})|\mathcal{A}^{n-1}) + \dots + \mathrm{H}(g^{-1}(\mathcal{A})|\mathcal{A}) + \mathrm{H}(\mathcal{A})\Big)$$

$$\leq \lim_{n\to\infty} \frac{1}{n}\Big(\mathrm{H}(g^{-n}(\mathcal{A})|g^{-(n-1)}(\mathcal{A})) + \dots + \mathrm{H}(g^{-1}(\mathcal{A})|\mathcal{A})\Big) = \mathrm{H}(g^{-1}(\mathcal{A})|\mathcal{A}). \tag{9.1.3}$$

(Compare this computation with the one done in Theorem 1.4.5 or in Proof of Theorem 1.5.4, which would result with $\mathrm{h}(g,\mathcal{A}) \leq \mathrm{H}(\mathcal{A}|g^{-1}(\mathcal{A}))$.) Going back to our situation, since

$$\mathrm{H}_{\mu_{\mathcal{P}_k(x)}}(g^{-1}(\mathcal{P}_k)|\mathcal{P}_k(x)) \leq \log \#\{P \in \mathcal{P}_k : g^{-1}(P) \cap \mathcal{P}_k(x) \neq \emptyset\} = \log N(g,x,k)$$

and by Theorem 1.8.7.a, we obtain

$$\mathrm{h}_\mu(g) \leq \limsup_{k\to\infty} \mathrm{H}_\mu(g^{-1}(\mathcal{P}_k)|\mathcal{P}_k) = \limsup_{k\to\infty} \int \mathrm{H}_{\mu_{\mathcal{P}_k(x)}}(g^{-1}(\mathcal{P}_k)|\mathcal{P}_k(x))\, d\mu(x)$$

$$\leq \limsup_{k\to\infty} \int \log N(g,x,k)\, d\mu(x) \leq \int \log N(g,x)\, d\mu(x).$$

Applying this inequality to $g = f^n$ ($n \geq 1$ an integer) and employing (9.1.2) we get

$$\mathrm{h}_\mu(f) = \frac{1}{n}\mathrm{h}_\mu(f^n) \leq \frac{1}{n}\int \log N(f^n,x)\, d\mu(x) = \int \frac{1}{n}\log N(f^n,x)\, d\mu(x)$$
$$\leq \int \frac{1}{n}\log 4\pi(|(f^n)'(x)|+2)^2\, d\mu(x)$$

Since $0 \leq \frac{1}{n}\log(|(f^n)'(x)|+2)^2 \leq 2(\log\big(\sup_X |f'|\big)+1)$ and $\lim_{n\to\infty}\frac{1}{n}\log(|(f^n)'(x)|+2) = \max\{0,\chi_\mu(x)\}$ for μ-a.e $x \in X$, it follows from the Dominated Convergence Theorem (Ch1.Sec.1) that

$$h_\mu(f) \leq \lim_{n\to\infty} \int \frac{1}{n}\log(|(f^n)'(x)|+2)^2\, d\mu(x) = \int \max\{0, 2\chi_\mu(x)\}\, d\mu.$$

The proof is completed. ♣

Exercise. Prove the following general version of Theorem 8.1.1: Let X be a compact f-invariant subset of a smooth Riemannian manifold for a C^1 mapping $f : U \to M$, defined on a neighbourhood U of X. Let μ be an f-invariant Borel probability measure X. Then

$$h_\mu(f) \le \int_X \max\{0, \chi_\mu^+(x)\}\, d\mu(x),$$

where $\chi_\mu^+(x) = \lim_{n\to\infty} \frac{1}{n} \log \|(Df^n)^\wedge\|$. Here Df^n is the differential and $(Df^n)^\wedge$ is the exterior power, the linear operator between the exterior algebras generated by the tangent spaces at x and $f^n(x)$. The norm is induced by the Riemann metric. Saying directly $\|(Df^n)^\wedge\|$ is supremum of the volumes of Df^n-images of unit cubes in k-dimensional subspaces of T_xM with $k = 0, 1, ..., \dim M$.

Note. Theorem 9.1.1 and Exercise rely on [Ruelle] D. Ruelle: An inequality of the entropy of differentiable maps. Bol. Soc. Bras. Mat. 9 (1978), 83-87.

§9.2. PESIN'S THEORY

In this section we work in the same setting and we follow the same notation as in Section 9.1.

Lemma 9.2.1. If μ is a Borel finite measure on $\mathbb{R}^n$, $n \ge 1$, a is an arbitrary point of $\mathbb{R}^n$ and the function $z \mapsto \log|z-a|$ is μ-integrable, then for every $C > 0$ and every $0 < t < 1$,

$$\sum_{n\ge1} \mu(B(a, Ct^n)) < \infty.$$

Proof. Since μ is finite and since given $t < s < 1$ there exists $q \ge 1$ such that $Ct^n \le s^n$ for all $n \ge q$, without loosing generality we may assume that $C = 1$. Recall that given $b \in \mathbb{R}^n$, and two numbers $0 \le r < R$, $R(b, r, R) = \{z \in \mathbb{C} : r \le |z-b| < R\}$. Since $-\log(t^n) \le -\log|z-a|$ for every $z \in B(a, t^n)$ we get the following.

$$\begin{aligned}\sum_{n\ge1} \mu(B(a,t^n)) &= \sum_{n\ge1} n\mu(R(a,t^{n+1},t^n)) = \frac{-1}{\log t}\sum_{n\ge1} -\log(t^n)\mu(R(a,t^{n+1},t^n)) \\ &\le \frac{-1}{\log t}\int_{B(a,t)} -\log|z-a|\, d\mu(z) < +\infty\end{aligned}$$

The proof is finished. ♣

Lemma 9.2.2. If μ is a Borel finite measure on $\overline{\mathbb{C}}$, $n \ge 1$, and $\log|f'|$ is μ integrable, then the function $z \mapsto \log|z-c| \in L^1(\mu)$ for every critical point c of f. If additionally μ is f-invariant, then also the function $z \mapsto \log|z - f(c)| \in L^1(\mu)$.

Proof. That $\log|z-c| \in L^1(\mu)$ follows from the fact that near c we have $C^{-1}|z-q|^{q-1} \leq |f'(z)| \leq C|z-c|^{q-1}$, where $q \geq 2$ is the order of the critical point c and $C \geq 1$ is a universal constant, and since out of any neighbourhood of the set of critical points of f, $|f'(z)|$ is uniformly bounded away from zero and infinity. In order to prove the second part of the lemma, consider a ray R emanating from $f(c)$ such that $\mu(R) = 0$ and a disk $B(f(c), r)$ such that $f_c^{-1} : B(f(c), r) \setminus R \to \overline{\mathbb{C}}$, an inverse branch of f sending $f(c)$ to c, is well-defined. Let $D = B(f(c), r) \setminus R$. We may additionally require $r > 0$ to be so small that $|z - f(c)| \asymp |f_c^{-1}(z) - c|^q$. It suffices to show that the integral $\int_D \log|z - f(c)|\, d\mu(z)$ is finite. And indeed, by f-invariance of μ we have

$$
\begin{aligned}
\int_D \log|z-f(c)|\, d\mu(z) &= \int_X 1_D(z) \log|z-f(c)|\, d\mu(z) \asymp \int_X 1_D(z) \log|f_c^{-1}(z)-c|^q\, d\mu(z) \\
&= \int_X (1_D \circ f)(z) \log|z-c|^q\, d\mu(z) = \int_X 1_{f^{-1}(D)} \log|z-c|^q\, d\mu(z)
\end{aligned}
$$

Notice here that the function $1_D(z)\log|f_c^{-1}(z)-c|^q$ is well-defined on X indeed and that unlike most of our comparability signs, the sign in the formula above means an additive comparability. The finiteness of the last integral follows from the first part of this lemma. ♣

Theorem 9.2.3. Let $(Z, \mathcal{F}, \nu)$ be a measure space with an ergodic measure preserving automorphism $T : Z \to Z$. Let $f : X \to X$ be a continuous map from a compact set $X \subset \overline{\mathbb{C}}$ onto itself having a holomorphic extention onto a neighbourhood of X ($f \in \mathcal{A}(X)$). Suppose that μ is an f-invariant ergodic measure on X with positive Lyapunov exponent. Suppose also that $h : Z \to X$ is a measurable mapping such that $\nu \circ h^{-1} = \mu$ and $h \circ T = f \circ h$ ν-a.e.. Then for ν-a.e. $z \in Z$ there exists $r(z) > 0$ such that for every $n \geq 1$ there exists $f_{x_n}^{-n} : B(x, r(z)) \to \overline{\mathbb{C}}$, an inverse branch of f^n sending $x = h(z)$ to $x_n = h(T^{-n}(z))$. In addition, for an arbitrary χ, $-\chi_\mu(f) < \chi < 0$, (not depending on z) and a constant $K(z)$

$$
|(f_{x_n}^{-n})'(y)| < K(z)e^{\chi n} \text{ and } \frac{|(f_{x_n}^{-n})'(w)|}{|(f_{x_n}^{-n})'(y)|} \leq K
$$

for all $y, w \in B(x, r(z))$. K is here the Koebe constant corresponding to the scale $1/2$.

Proof. Suppose first that $\mu\big(\bigcup_{n\geq 1} f^n(\mathrm{Crit}(f))\big) > 0$. Since μ is ergodic this implies that μ must be concentrated on a periodic orbit of an element $w \in \bigcup_{n\geq 1} f^n(\mathrm{Crit}(f))$. This means that $w = f^q(c) = f^{q+k}(c)$ for some $q, k \geq 1$ and $c \in \mathrm{Crit}(f)$, and

$$
\mu(\{f^q(c), f^{q+1}(c), \ldots, f^{q+k-1}(c)\}) = 1.
$$

Since $\int \log|f'|\, d\mu > 0$, $|(f^k)'(f^q(c))| > 1$. Thus the theorem is obviously true for the set $h^{-1}(\{f^q(c), f^{q+1}(c), \ldots, f^{q+k-1}(c)\})$ of ν measure 1.

So, suppose that $\mu\big(\bigcup_{n\geq 1} f^n(\mathrm{Crit}(f))\big) = 0$. Set $R = \min\{1, \mathrm{dist}(X, \overline{\mathbb{C}} \setminus U(f))\}$ and fix $\lambda \in (e^{\frac{1}{4}\chi}, 1)$. Consider $z \in Z$ such that $x = h(z) \notin \bigcup_{n\geq 1} f^n(\mathrm{Crit}(f))$,

$$
\lim_{n\to\infty} \frac{1}{n} \log|(f^n)'(h(T^{-n}(z))| = \chi_\mu(f),
$$

and $x_n = h(T^{-n}(z)) \in B(f(\mathrm{Crit}(f)), R\lambda^n)$ only for finitely many n's. We shall first demonstrate that the set of points satisfying these properties is of full measure ν. Indeed, the first requirement is satisfied by our hyphothesis, the second is due to Birkhoff's ergodic theorem. In order to prove that the set of points satisfying the third condition has ν measure 1 notice that

$$\begin{aligned}\sum_{n\geq 1} \nu\big(T^n(h^{-1}(B(f(\mathrm{Crit}(f)), R\lambda^n)))\big) &= \sum_{n\geq 1} \nu\big(h^{-1}(B(f(\mathrm{Crit}(f)), R\lambda^n))\big) \\ &= \sum_{n\geq 1} \mu(B(f(\mathrm{Crit}(f)), R\lambda^n)) < \infty,\end{aligned}$$

where the last inequality we wrote due to Lemma 9.2.2 and Lemma 9.2.1. The application of the Borel-Canteli lemma finishes now the demonstration. Fix now an integer $n_1 = n_1(z)$ so large that $x_n = h(T^{-n}(z)) \notin B(f(\mathrm{Crit}(f)), R\lambda^n)$ for all $n \geq n_1$. Notice that because of our choices there exists $n_2 \geq n_1$ such that $|(f^n)'(x_n)|^{-1/4} < \lambda^n$ for all $n \geq n_2$. Finally set $S = \sum_{n\geq 1} |(f^n)'(x_n)|^{-1/4}$, $b_n = \frac{1}{2}S^{-1}|(f^{n+1})'(x_{n+1})|^{\frac{-1}{4}}$, and

$$\Pi = \Pi_{n=1}^{\infty}(1-b_n)^{-1}$$

which converges since the series $\sum_{n\geq 1} b_n$ converges. Choose now $r = r(z)$ so small that $16r(z)\Pi K S^3 \leq R$, all the inverse branches $f_{x_n}^{-n} : B(x_0, \Pi r(z)) \to \overline{\mathbb{C}}$ are well-defined for all $n = 1, 2, \ldots, n_2$ and $\mathrm{diam}\big(f_{x_{n_2}}^{-n_2}(B\big(x_0, r\Pi_{k\geq n_2}(1-b_k)^{-1}\big)\big) \leq \lambda^{n_2}R$. We shall show by induction that for every $n \geq n_2$ there exists an analytic inverse branch $f_{x_n}^{-n} : B\big(x_0, r\Pi_{k\geq n}(1-b_k)^{-1}\big) \to \mathbb{C}$, sending x_0 to x_n and such that

$$\mathrm{diam}\big(f_{x_n}^{-n}(B\big(x_0, r\Pi_{k\geq n}(1-b_k)^{-1}\big)\big) \leq \lambda^n R.$$

Indeed, for $n = n_2$ this immediately follows from our requirements imposed on $r(z)$. So, suppose that the claim is true for some $n \geq n_2$. Since $x_n = f_{x_n}^{-n}(x_0) \notin B(\mathrm{Crit}(f), R\lambda^n)$ and since $\lambda^n R \leq R$, there exists an inverse branch $f_{x_{n+1}}^{-1} : B(x_n, \lambda^n R) \to \mathbb{C}$ sending x_n to x_{n+1}. Since $\mathrm{diam}\big(f_{x_n}^{-n}(B\big((x_0, r\Pi_{k\geq n}(1-b_k)^{-1}\big)\big) \leq \lambda^n R$, the composition $f_{x_{n+1}}^{-1} \circ f_{x_n}^{-n} B\big(x_0, r\Pi_{k\geq n}(1-b_k)^{-1}\big) \to \mathbb{C}$ is well-defined and forms the inverse branch of f^{n+1} that sends x_0 to x_{n+1}. By the Koebe distortion theorem we now estimate

$$\begin{aligned}&\mathrm{diam}\big(f_{x_{n+1}}^{-(n+1)}(B\big(x_0, r\Pi_{k\geq n+1}(1-b_k)^{-1}\big))\big) \\ &\qquad \leq 2r\Pi_{k\geq n+1}(1-b_k)^{-1}|(f^{n+1})'(x_{n+1})|^{-1}Kb_n^{-3} \\ &\qquad \leq 16r\Pi K S^3|(f^{n+1})'(x_{n+1})|^{-1}|(f^{n+1})'(x_{n+1})|^{\frac{3}{4}} \\ &\qquad = 16r\Pi K S^3|(f^{n+1})'(x_{n+1})|^{-\frac{1}{4}} \\ &\qquad \leq R\lambda^{n+1},\end{aligned}$$

where the last inequality sign we wrote due to our choice of r and the number n_2. Putting $r(z) = r/2$ the second part of this theorem follows now as a combined application of the equality $\lim_{n\to\infty} \frac{1}{n}\log|(f^n)'(x_n)| = \chi_\mu(f)$ and the Koebe distortion theorem. ♣

As an immediate consequence of Theorem 9.2.3 we get the following.

Corollary 9.2.4. Assume the same notation and asumptions as in Theorem 9.2.3. Fix $\varepsilon > 0$. Then there exist a set $Z(\varepsilon) \subset Z$, the numbers $r(\varepsilon) \in (0,1)$ and $K(\varepsilon) \geq 1$ such that $\mu(Z(\varepsilon)) > 1 - \varepsilon$, $r(z) \geq r(\varepsilon)$ for all $z \in Z(\varepsilon)$ and with $x_n = h(T^{-n}(z))$

$$K(\varepsilon)^{-1} \exp(-(\chi_\mu + \varepsilon)n) \leq |(f_{x_n}^{-n})(y)| \leq K(\varepsilon) \exp(-(\chi_\mu - \varepsilon)n) \text{ and } \frac{|(f_{x_n}^{-n})'(w)|}{|(f_{x_n}^{-n})'(y)|} \leq K$$

for all $n \geq 1$, all $z \in Z(\varepsilon)$ and all $y, w \in B(x_0, r(\varepsilon))$. K is here the Koebe constant corresponding to the scale $1/2$.

Remark 9.2.5. In our future applications the system (Z, f, ν) will be usually given by the natural extension of the holomorphic system (f, μ).

§9.3 MAÑÉ'S PARTITION

In this section, basically following Mañé's book ???, we construct so called Mañé's partition which will play an important role in the proof of a part of the Volume Lemma given in the next section. We begin with the following elementary fact.

Lemma 9.3.1. If $x_n \in (0,1)$ for every $n \geq 1$ and $\sum_{n=1}^\infty n x_n < \infty$, then $\sum_{n=1}^\infty -x_n \log x_n < \infty$.

Proof. Let $S = \{n : -\log x_n \geq n\}$. Then

$$\sum_{n=1}^\infty -x_n \log x_n = \sum_{n \notin S} -x_n \log x_n + \sum_{n \in S} -x_n \log x_n \leq \sum_{n=1}^\infty n x_n + \sum_{n \in S} -x_n \log x_n$$

Since $n \in S$ means that $x_n \leq e^{-n}$ and since $\log t \leq 2\sqrt{t}$ for all $t \geq 1$, we have

$$\sum_{n \in S} x_n \log \frac{1}{x_n} \leq 2 \sum_{n=1}^\infty x_n \sqrt{\frac{1}{x_n}} \leq 2 \sum_{n=1}^\infty e^{-\frac{1}{2}n} < \infty$$

The proof is finished. ♣

The next lemma is the main and simultaneously the last result of this section.

Lemma 9.3.2. If μ is a Borel probability measure concentrated on a bounded subset M of a Euclidean space and $\rho : M \to (0,1]$ is a measurable function such that $\log \rho$ is integrable with respect to μ, then there exists a countable measurable partition, called Mañé's partition, $\mathcal{P}$ of M such that $\mathrm{H}_\mu(\mathcal{P}) < \infty$ and

$$\operatorname{diam}(\mathcal{P}(x)) \le \rho(x)$$

for μ-almost every $x \in M$.

Proof. Let q be the dimension of the Euclidean space containing M. Since M is bounded, there exists a constant $C > 0$ such that for every $0 < r < 1$ there exists a partition $\mathcal{P}_r$ of M of diameter $\le r$ and which consists of at most Cr^{-q} elements. For every $n \ge 0$ put $U_n = \{x \in M : e^{-(n+1)} < \rho(x) \le e^{-n}\}$. Since $\log \rho$ is a non-positive integrable function, we have

$$\sum_{n=1}^{\infty} -n\mu(U_n) \ge \sum_{n=1}^{\infty} \int_{U_n} \log \rho \, d\mu = \int_M \log \rho \, d\mu > -\infty$$

so that

$$\sum_{n=1}^{\infty} n\mu(U_n) < +\infty. \tag{9.3.1}$$

Define now $\mathcal{P}$ as the partition whose atoms are of the form $Q \cap U_n$, where $n \ge 0$ and $Q \in \mathcal{P}_{r_n}$, $r_n = e^{-(n+1)}$. Then

$$\mathrm{H}_\mu(\mathcal{P}) = \sum_{n=0}^{\infty} \Big(- \sum_{U_n \supset P \in \mathcal{P}} \mu(P) \log \mu(P) \Big).$$

But for every $n \ge 0$

$$\begin{aligned}
- \sum_{U_n \supset P \in \mathcal{P}} \mu(P) \log \mu(P) &= \mu(U_n) \sum_P - \frac{\mu(P)}{\mu(U_n)} \log\Big(\frac{\mu(P)}{\mu(U_n)}\Big) - \mu(U_n) \sum_P \frac{\mu(P)}{\mu(U_n)} \log(\mu(U_n)) \\
&\le \mu(U_n)(\log C - q \log r_n) - \mu(U_n) \log \mu(U_n) \\
&\le \mu(U_n) \log C + q(n+1)\mu(U_n) - \mu(U_n) \log \mu(U_n).
\end{aligned}$$

Thus, summing over all $n \ge 0$, we obtain

$$\mathrm{H}_\mu(\mathcal{P}) \le \log C + q + q \sum_{n=0}^{\infty} n\mu(U_n) + \sum_{n=0}^{\infty} -\mu(U_n) \log \mu(U_n).$$

Therefore looking at (9.3.1) and Lemma 9.3.1 we conclude that $\mathrm{H}_\mu(\mathcal{P})$ is finite. Also, if $x \in U_n$, then the atom $\mathcal{P}(x)$ is contained in some atom of P_{r_n} and therefore

$$\operatorname{diam}(\mathcal{P}(x)) \le r_n = e^{-(n+1)} < \rho(x).$$

Now the remark that the union of all the sets U_n is of measure 1 completes the proof. ♣

§9.4 VOLUME LEMMA AND THE FORMULA $\mathrm{HD}(\mu) = \mathrm{h}_\mu(f)/\chi_\mu(f)$

In this section we keep the notation of Sections 9.1 and 9.2 and our main purpose is to prove the following two results which generalize the respective results in Chapter 7.

Theorem 9.4.1. If $f \in \mathcal{A}(X)$ and μ is an ergodic f-invariant measure with positive Lyapunov exponent, then $\mathrm{HD}(\mu) = \mathrm{h}_\mu(f)/\chi_\mu(f)$.

Theorem 9.4.2. (Volume Lemma) With the assumptions of Theorem 9.4.1

$$\lim_{r\to 0} \frac{\log(\mu(B(x,r)))}{\log r} = \frac{\mathrm{h}_\mu(f)}{\chi_\mu(f)}$$

for μ-a.e. $x \in X$.

In view of Corollary 6.6.4, Theorem 9.4.1 follows from Theorem 9.4.2 and we only need to prove the latter one. Let us prove first

$$\liminf_{r\to 0} \frac{\log(\mu(B(x,r)))}{\log r} \geq \frac{\mathrm{h}_\mu(f)}{\chi_\mu(f)} \tag{9.4.1}$$

for μ-a.e. $x \in X$. By Corollary 7.1.9 there exists a finite partition $\mathcal{P}$ such that for an arbitrary $\varepsilon > 0$ and every x in a set X_o of full measure μ there exists $n(x) \geq 0$ such that for all $n \geq n(x)$.

$$B(f^n(x), e^{-\varepsilon n}) \subset \mathcal{P}(f^n(x)). \tag{9.4.2}$$

Let us work from now on in the natural extension $(\tilde{X}, \tilde{f}, \tilde{\mu})$. Let $\tilde{X}(\varepsilon)$ and $r(\varepsilon)$ be given by Corollary 9.2.4, i.e. $\tilde{X}(\varepsilon) = Z(\varepsilon)$. In view of Birkhoff's Ergodic Theorem there exists a measurable set $\tilde{F}(\varepsilon) \subset \tilde{X}(\varepsilon)$ such that $\tilde{\mu}(\tilde{F}(\varepsilon)) = \tilde{\mu}(\tilde{X}(\varepsilon))$ and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n-1} \chi_{\tilde{X}(\varepsilon)} \circ \tilde{f}^n(\tilde{x}) = \tilde{\mu}(\tilde{X}(\varepsilon))$$

for every $\tilde{x} \in \tilde{F}(\varepsilon)$. Let $F(\varepsilon) = \pi(\tilde{F}(\varepsilon))$. Then $\mu(F(\varepsilon)) = \tilde{\mu}(\pi^{-1}(F(\varepsilon)) \geq \tilde{\mu}(\tilde{F}(\varepsilon)) = \tilde{\mu}(\tilde{X}(\varepsilon))$ converges to 1 if $\varepsilon \searrow 0$. Consider now $x \in F(\varepsilon) \cap X_o$ and take $\tilde{x} \in \tilde{F}(\varepsilon)$ such that $x = \pi(\tilde{x})$. Then by the above there exists an increasing sequence $\{n_k = n_k(x) : k \geq 1\}$ such that $\tilde{f}^{n_k}(\tilde{x}) \in \tilde{X}(\varepsilon)$ and

$$\frac{n_{k+1} - n_k}{n_k} \leq \varepsilon \tag{9.4.3}$$

for every $k \geq 1$. Moreover, we can assume that $n_1 \geq n(x)$. Consider now an integer $n \geq n_1$ and the ball $B\big(x, Cr(\varepsilon)\exp(-(\chi_\mu + (2 + \log\|f'\|)\varepsilon)n)\big)$, where $0 < C < (Kr(\varepsilon))^{-1}$ is a constant (possibly depending on x) so small that

$$f^q\big(B(x, Cr(\varepsilon)\exp -(\chi_\mu + (2 + \log\|f'\|)\varepsilon)n)\big) \subset P(f^q(x)) \tag{9.4.4}$$

for every $q \le n_1$ and $K(\varepsilon) \ge 1$ is the constant appearing in Corollary 9.2.4. Take now any q, $n_1 \le q \le n$, and associate k such that $n_k \le q \le n_{k+1}$. Since $\tilde{f}^{n_k}(\tilde{x}) \in \tilde{X}(\varepsilon)$ and since $\pi(\tilde{f}^{n_k}(\tilde{x})) = f^{n_k}(x)$, Corollary 9.2.4 produces a holomorphic inverse branch $f_x^{-n_k} : B(f^{n_k}(x), r(\varepsilon)) \to \overline{\mathbb{C}}$ of f^{n_k} such that $f_x^{-n_k}(f^{n_k}(x)) = x$ and

$$f_x^{-n_k}\big(B(f^{n_k}(x), r(\varepsilon))\big) \supset B\big(x, K(\varepsilon)r(\varepsilon)^{-1}\exp(-(\chi_\mu + \varepsilon)n_k)\big).$$

Since $B(x, Cr(\varepsilon)\exp -(\chi_\mu + (2+\log\|f'\|)\varepsilon)n) \subset B\big(x, K(\varepsilon)^{-1}r(\varepsilon)\exp -(\chi_\mu+\varepsilon)n_k)\big)$, it follows from Corollary 9.2.4 that

$$\begin{aligned} f^{n_k}\big(B(x, Cr(\varepsilon)\exp -(\chi_\mu + (2+\log\|f'\|)\varepsilon)n)\big) &\subset \\ \subset B\big(f^{n_k}(x), CKr(\varepsilon)e^{-\chi_\mu(n-n_k)}&\exp(\varepsilon(n_k - (2+\log\|f'\|)n))\big). \end{aligned}$$

Since $n \ge n_k$ and since by (9.4.3) $q - n_k \le \varepsilon n_k$, we therefore obtain

$$\begin{aligned} &f^q\big(B(x, Cr(\varepsilon)\exp -(\chi_\mu + (2+\log\|f'\|)\varepsilon)n)\big) \subset \\ &\quad\subset B(f^q(x), CK(\varepsilon)r(\varepsilon)e^{-\chi_\mu(n-n_k)}\exp(\varepsilon(n_k - (2+\log\|f'\|)n))\exp((q-n_k)\log\|f'\|) \\ &\quad\subset B(f^q(x), CK(\varepsilon)r(\varepsilon)\exp\big(\varepsilon(n_k\log\|f'\| + n_k - 2n - n\log\|f'\|)\big) \\ &\quad\subset B(f^q(x), CK(\varepsilon)r(\varepsilon)e^{-\varepsilon n}) \subset B(f^q(x), e^{-\varepsilon q}). \end{aligned}$$

Combining this, (9.4.2), and (9.4.4), we get

$$B\big(x, Cr(\varepsilon)\exp -(\chi_\mu + (2+\log\|f'\|)\varepsilon)n)\big) \subset \bigvee_{j=0}^{n} f^{-j}(\mathcal{P})(x).$$

Therefore, applying Theorem 1.5.5 (the Shanon-McMillan-Breiman Theorem), we have

$$\liminf_{n\to\infty} -\frac{1}{n}\log\mu\big(B(x, Cr(\varepsilon)\exp -(\chi_\mu + (2+\log\|f'\|)\varepsilon)n)\big) \ge \mathrm{h}_\mu(f,\mathcal{P}) \ge \mathrm{h}_\mu(f) - \varepsilon$$

It means that denoting the number $Cr(\varepsilon)\exp -(\chi_\mu + (2+\log\|f'\|)\varepsilon)n)$ by r_n, we have

$$\liminf_{n\to\infty} \frac{\log\mu(B(x,r_n)}{\log r_n} \ge \frac{\mathrm{h}_\mu(f) - \varepsilon}{\chi_\mu(f) + (2+\log\|f'\|)\varepsilon}$$

Now, since $\{r_n\}$ is a geometric sequence and since $\varepsilon > 0$ can be taken arbitrarily small, we conclude that for μ-a.e. $x \in X$

$$\liminf_{n\to\infty} \frac{\log\mu(B(x,r)}{\log r} \ge \frac{\mathrm{h}_\mu(f)}{\chi_\mu(f)}$$

This completes the proof of (9.4.1). ♣

Remark. Since here $X \subset \mathbb{C}$, we could have considered a partition $\mathcal{P}$ of a neighbourhood of X in $\mathbb{C}$ where $\partial_{\mathcal{P},a}$ would have a more standard sense, see Remark after Corollary 7.1.8.

Now let us prove that

$$\limsup_{r\to 0} \frac{\log(\mu(B(x,r)))}{\log r} \le \mathrm{h}_\mu(f)/\chi_\mu(f) \tag{8.4.5}$$

for μ-a.e. $x \in X$.

In order to prove this formula we again work in the natural extension $(\tilde{X}, \tilde{f}, \tilde{\mu})$ and we apply Pesin theory. In particular the sets $\tilde{X}(\varepsilon)$, $\tilde{F}(\varepsilon) \subset \tilde{X}(\varepsilon)$ and the radius $r(\varepsilon)$, produced in Corollary 9.2.4 have the same neaning as in the proof of (9.4.1). To begin with notice that there exist two numbers $R > 0$ and $0 < Q < \min\{1, r(\varepsilon)/2\}$ such that the foloowing two conditions are satisfied.

(9.4.6) If $z \notin B(\mathrm{Crit}(f), R)$, then $f|_{B(z,Q)}$ is injective.

(9.4.7) If $z \in B(\mathrm{Crit}(f), R)$, then $f|_{B(z,Q\mathrm{dist}(z,\mathrm{Crit}(f)))}$ is injective.

Observe also that if z is sufficiently close to a critical point c, then $f'(z)$ is of order $(z-c)^{q-1}$, where $q \geq 2$ is the order of critical point c. In particular the quotient of $f'(z)$ and $(z-c)^{q-1}$ remains bounded away from 0 and ∞ and therefore there exists a constant number $B > 1$ such that $|f'(z)| \leq B\mathrm{dist}(z, \mathrm{Crit}(f))$. So, in view of Theorem 9.2.2, the logarithm of the function $\rho(z) = Q\min\{1, \mathrm{dist}(z, Crit(f))$ is integrable and consequently Lemma 9.3.2 applies. Let $\mathcal{P}$ be the Mañé's partition produced by this lemma. Then $B(x, \rho(x)) \supset \mathcal{P}(x)$ for μ-a.e. $x \in X$, say for a subset X_ρ of X of measure 1. Consequently

$$B_n(x,\rho) = \bigcap_{j=0}^{n-1} f^{-j}\big(B(f^j(x), \rho(f^j(x)))\big) \supset \mathcal{P}_0^n(x) \tag{9.4.8}$$

for every $n \geq 1$ and every $x \in X_\rho$. By our choice of Q and the definition of ρ, the function f is injective on all balls $B(f^j(x), \rho(f^j(x)))$, $j \geq 0$, and therefore f^k is injective on the set $B_n(x,\rho)$ for every $0 \leq k \leq n-1$. Now, let $x \in F(\varepsilon) \cap X_\rho$ and let k be the greatest subscript such that $q = n_k(x) \leq n-1$. Denote by f_x^{-q} the unique holomorphic inverse branch of f^q produced by Corollary 9.2.4 which sends $f^q(x)$ to x. Clearly $B_n(x,\rho) \subset f^{-q}(B(f^q(x), \rho(f^q(x))))$ and since f^q is injective on $B_n(x,\rho)$ we even have

$$B_n(x,\rho) \subset f_x^{-q}(B(f^q(x), \rho(f^q(x)))).$$

By Corollary 9.2.4 $\mathrm{diam}\big(f_x^{-q}(B(f^q(x), \rho(f^q(x))))\big) \leq K\exp(-q(\chi_\mu - \varepsilon))$. Since by (9.4.3), $n \leq q(1+\varepsilon)$ we finally deduce that

$$B_n(x,\rho) \subset B\big(x, K\exp(-n\frac{\chi_\mu - \varepsilon}{1+\varepsilon})\big).$$

Thus, in view of (9.4.8)

$$B\left(x, K\exp(-n\frac{\chi_\mu - \varepsilon}{1+\varepsilon})\right) \supset \mathcal{P}_0^n(x).$$

Therefore, denoting by r_n the radius of the ball above, it follows from Shanon-McMillan-Breiman theorem that for μ-a.e $x \in X$

$$\limsup_{n\to\infty} -\frac{1}{n}\log\mu(B(x, r_n) \leq \mathrm{h}_\mu(f, \mathcal{P}) \leq \mathrm{h}_\mu(f).$$

So

$$\limsup_{n\to\infty} \frac{\log \mu(B(x,r_n))}{\log r_n} \le \frac{\mathrm{h}_\mu(f)}{\chi_\mu(f)-\varepsilon}(1+\varepsilon).$$

Now, since $\{r_n\}$ is a geometric sequence and since ε can be taken arbitrarily small, we conclude that for μ-a.e. $x \in X$

$$\limsup_{n\to\infty} \frac{\log \mu(B(x,r))}{\log r} \le \frac{\mathrm{h}_\mu(f)}{\chi_\mu(f)}.$$

This completes the proof of (9.4.5) and because of (9.4.1) also the proof of Theorem 9.4.2. ♣

§9.5 PRESSURE-LIKE DEFINITION OF THE FUNCTIONAL $\mathrm{h}_\mu + \int \phi\, d\mu$.

In this section we prepare some general tools used in the next section to approximate topological pressure on hyperbolic sets. No smoothness is assumed here, we work in purely metric setting only. Our exposition is similar to that contained in Chapter 2.

Let $T : X \to X$ be a continuous map of a compact metric space (X,ρ) and let μ be a Borel probability measure on X. Given $\varepsilon > 0$ and $0 \le \delta \le 1$ a set $E \subset X$ is said to be $\mu-(n,\varepsilon,\delta)$-*spanning* if

$$\mu\Big(\bigcup_{x\in E} B_n(x,\varepsilon)\Big) \ge 1-\delta.$$

Let $\phi : X \to I\!R$ be a continuous function. We define

$$Q_\mu(T,\phi,n,\varepsilon,\delta) = \inf_E \Big\{\sum_{x\in E} \exp S_n\phi(x)\Big\}$$

where the infimum is taken over all $\mu-(n,\varepsilon,\delta)$-spanning sets E. The main result of this section is the following.

Theorm 9.5.1. For every $0 < \delta < 1$ and every ergodic measure μ

$$\mathrm{h}_\mu(T) + \int \phi\, d\mu = \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log Q_\mu(T,\phi,n,\varepsilon,\delta) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log Q_\mu(T,\phi,n,\varepsilon,\delta)$$

Proof. Denote the the number following the first equality sign by $\underline{\mathrm{P}}_\mu(T,\phi,\delta)$ and the number following the second equality sign by $\overline{\mathrm{P}}_\mu(T,\phi,\delta)$. First, following essentially the proof of the Part I of Theorem 2.3.1, we shall show that

$$\underline{\mathrm{P}}_\mu(T,\phi,\delta) \ge \mathrm{h}_\mu(T) + \int \phi\, d\mu \tag{9.5.1}$$

Indeed, similarly as in that proof consider a finite partition $\mathcal{U} = \{A_1, \ldots, A_s\}$ of X into Borel sets and compact sets $B_i \subset A_i$, $i = 1, 2, \ldots, A_s\}$, such that for the partition $\mathcal{V} = \{B_1, \ldots, B_s, X \setminus (B_1 \cup \ldots \cup B_s)\}$ we have $H_\mu(\mathcal{U}|\mathcal{V}) \leq 1$. For every $\theta > 0$ and $q \geq 1$, set

$$X_q = \{x \in X : -\frac{1}{n} \log \mu\big(\mathcal{V}^n(x)\big) \geq \mathrm{h}_\mu(T, \mathcal{V}) - \theta \quad \text{for all} \quad n \geq q \text{ and}$$
$$\frac{1}{n} S_n\phi(x) \geq \int \phi \, d\mu - \theta \quad \text{for all} \quad n \geq q\}$$

Fix now $0 \leq \delta < 1$. It follows from Shannon-McMillan-Breiman theorem and Birkhoff's ergodic theorem that for q large enough $\mu(X_q) > \delta$. Take $0 < \varepsilon < \frac{1}{2}\min\{\rho(B_i, B_j) : 1 \leq i < j \leq s\} > 0$ so small that

$$|\phi(x) - \phi(y)| < \theta$$

if $\rho(x, y) \leq \varepsilon$. Since for every $x \in X$ the set $B_n(x, \varepsilon) \cap X_q$ can be covered by at most 2^n elements of $\mathcal{V}^n$,

$$\mu(B_n(x, \varepsilon) \cap X_q) \leq \exp\big(n(\log 2 - \mathrm{h}_\mu(T, \mathcal{V}) + \theta)\big).$$

Now let E be a $\mu - (n, \varepsilon, \delta)$-spanning set for $n \geq q$, and consider the set $E' = \{x \in E : B_n(x, \varepsilon) \cap X_q \neq \emptyset\}$. Take any point $y(x) \in B_n(x, \varepsilon) \cap X_q$. Then by the choice of ε, $S_n\phi(x) - S_n\phi(y) > -n\theta$. Therefore we have

$$\begin{aligned}
&\sum_{x \in E} \exp S_n\phi(x) \exp\Big(-n\big(\mathrm{h}_\mu(T, \mathcal{V}) + \int \phi \, d\mu - 3\theta - \log 2\big)\Big) \geq \\
&\quad \geq \sum_{x \in E'} \exp S_n\phi(x) \exp\Big(-n\big(\mathrm{h}_\mu(T, \mathcal{V}) + \int \phi \, d\mu - 3\theta - \log 2\big)\Big) \\
&\quad = \sum_{x \in E'} \exp\Big(S_n\phi(x) - n \int \phi \, d\mu\Big) \exp\big(-n(\mathrm{h}_\mu(T, \mathcal{V}) - 3\theta - \log 2)\big) \\
&\quad = \sum_{x \in E'} \exp\Big(S_n\phi(x) - S_n\phi(y) + S_n\phi(y) - n \int \phi \, d\mu\Big) \exp\Big(-n\big(\mathrm{h}_\mu(T, \mathcal{V}) - 3\theta - \log 2\big)\Big) \\
&\quad \geq \sum_{x \in E'} \exp(-n\theta) \exp(-n\theta) \exp(2n\theta) \exp\big(-n(\mathrm{h}_\mu(T, \mathcal{V}) - \theta - \log 2)\big) \\
&\quad = \sum_{x \in E'} \exp\big(n(\log 2 - \mathrm{h}_\mu(T, \mathcal{V}) + \theta)\big) \\
&\quad \geq \sum_{x \in E'} \mu(B_n(x, \varepsilon) \cap X_q) \geq \mu(X_q) - \delta > 0
\end{aligned}$$

which implies that

$$Q_\mu(T, \phi, n, \varepsilon, \delta) \geq \mathrm{h}_\mu(T, \mathcal{V}) + \int \phi \, d\mu - 3\theta - \log 2.$$

Since $\theta > 0$ is an arbitrary number and since $\mathrm{h}_\mu(T, \mathcal{U}) \leq \mathrm{h}_\mu(T, \mathcal{V}) + H_\mu(\mathcal{U}|\mathcal{V}) \leq \mathrm{h}_\mu(T, \mathcal{V}) + 1$, letting $\varepsilon \to 0$, we get

$$\underline{\mathrm{P}}_\mu(T, \phi, \delta) \geq \mathrm{h}_\mu(T, \mathcal{U}) - 1 + \int \phi \, d\mu - \log 2$$

Therefore, by the definition of entropy of an automorphism, $\underline{\mathrm{P}}_\mu(T,\phi,\delta) \geq \mathrm{h}_\mu(T) + \int \phi\, d\mu - \log 2 - 1$. Using now the standard trick, actually always applied in the setting we are whose point is to replace T by its arbitrary iterates T^k and ϕ by $S_k\phi$, we obtain $k\underline{\mathrm{P}}_\mu(T,\phi,\delta) \geq k\mathrm{h}_\mu(T) + k\int \phi\, d\mu - \log 2 - 1$. So, dividing this inequality by k, and letting $k \to \infty$, we finally obtain

$$\underline{\mathrm{P}}_\mu(T,\phi,\delta) \geq \mathrm{h}_\mu(T) + \int \phi\, d\mu$$

Now let us prove that

$$\overline{\mathrm{P}}_\mu(T,\phi,\delta) \leq \mathrm{h}_\mu(T) + \int \phi\, d\mu \tag{9.5.2}$$

where $\overline{\mathrm{P}}_\mu(T,\phi,\delta)$ denotes limsup appearing in the statement of Theorm 9.5.1. Indeed, fix $0 < \delta < 1$, then $\varepsilon > 0$ and $\theta > 0$. Let P be a finite partition of X of diameter $\leq \varepsilon$. By Shannon-McMillan-Breiman theorem and Birkhoff's ergodic theorem there exists a Borel set $Z \subset X$ such that $\mu(Z) > 1 - \delta$ and

$$\frac{1}{n}S_n\phi(x) \leq \int \phi\, d\mu + \theta \quad \text{and} \quad -\frac{1}{n}\log \mu(\mathrm{P}^n(x)) \leq \mathrm{h}_\mu(T) + \theta \tag{9.5.3}$$

for every n large enough and all $x \in Z$. From each element of P^n having non-empty intersection with Z choose one point obtaining, say, a set $\{x_1, x_2, \dots, x_q\}$. Then $B_n(x_j,\varepsilon) \supset \mathrm{P}^n(x_j)$ for every $j = 1, 2, \dots, q$ and therefore the set $\{x_1, x_2, \dots, x_q\}$ is $\mu - (n,\varepsilon,\delta)$-spanning. By the second part of (9.5.3) we have $q \leq \exp(n(\mathrm{h}_\mu(T) + \theta))$. Using also the first part of (9.5.3), we get

$$\sum_{j=1}^{q} \exp S_n\phi(x_j) \leq \exp(n(\mathrm{h}_\mu(T) + \theta + \int \phi\, d\mu + \theta))$$

Therefore $Q_\mu(T,\phi,n,\varepsilon,\delta) \leq \exp(n(\mathrm{h}_\mu(T) + \theta + \int \phi\, d\mu + \theta))$ and letting consequtively $n \to \infty$ and $\varepsilon \to 0$, we obtain $\overline{\mathrm{P}}_\mu(T,\phi,\delta) \leq \mathrm{h}_\mu(T) + \int \phi\, d\mu + 2\theta$. Since θ is an arbitrary positive number, (9.5.2) is proved. This and (9.5.1) complete the proof of Theorem 9.5.1. ♣.

§9.6 KATOK'S THEORY – HYPERBOLIC SETS, PERIODIC POINTS, AND PRESSURE

In this section we again come back to the setting of Section 9.1. So, let X be a compact subset of the closed complex plane $\overline{\mathbb{C}}$ and let $f : X \to X$ be a continuous map that can be analytically extended to an open neighbourhood $U(f)$ of X.

Let μ be an f-invariant ergodic measure on X with positive Lyapunov exponent. $\mathrm{h}_\mu(f)$ and let $\phi : X \to \mathbb{R}$ be a real continuous function. Our first aim is to show that the number

$\mathrm{h}_\mu(f)+\int \phi\, d\mu$ can be approximated by the topological pressures of ϕ on hyperbolic subsets of X and then as a straightforward consequence we will obtain the same approximation for the topological pressure $\mathrm{P}(f,\phi)$.

Theorm 9.6.1. If μ is an f-invariant ergodic measure on X with positive Lyapunov expenent χ_μ and if $\phi : X \to I\!\!R$ is a real-valued continuous function, then there exists a sequence X_k, $k=1,2,\ldots$, of compact f- invariant subsets of U such that for every k the restriction $f|_{X_k}$ is a conformal expanding repeller,

$$\liminf_{k\to\infty} \mathrm{P}(f|_{X_k},\phi) \ge \mathrm{h}_\mu(f) + \int \phi\, d\mu$$

and if μ_k is any ergodic f-invariant measure on X_k, then the sequence μ_k, $k=1,2,\ldots$, converges to μ in the weak-*-topology on a closed neighbourhood of X.

Proof. Since $\mathrm{P}(f|_{X_k},\phi+c)=\mathrm{P}(f|_{X_k},\phi)+c$ and since $\mathrm{h}_\mu(f)+\int(\phi+c)\,d\mu = \mathrm{h}_\mu(f)+\int \phi\, d\mu + c$, adding a constant if necessary, we can assume that ϕ is positive, that is that $\inf\phi>0$. As in Section 9.2 we work in the natural extension $(\tilde X,\tilde f,\tilde\mu)$. Given $\delta>0$ let $\tilde X(\delta)$ and $r(\delta)$ be produced by Corollary 9.2.4. The set $\pi(\tilde X(\delta))$ is assumed to be compact. This corollary implies the existence of a constant $\chi'>0$ (possibly with a smaller radius $r(\delta)$) such that

$$\mathrm{diam}\big(f_{x_n}^{-n}(B(\pi(\tilde x),r(\delta))\big) \le e^{-n\chi'} \tag{9.6.1}$$

for all $\tilde x\in\tilde X(\delta)$ and $n\ge 0$. Fix a countable basis $\{\psi_j\}_{j=1}^\infty$ of the Banach space $C(X)$ of all continuous real-valued functions $C(X)$. Fix $\theta>0$ and an integer $s\ge 1$. In view of Theorem 9.5.1 and continuity of functions ϕ and ψ_i there exists $\varepsilon>0$ so small that

$$\liminf_{n\to\infty}\frac{1}{n}\log Q_\mu(T,\phi,n,\varepsilon,\delta) - (\mathrm{h}_\mu(f)+\int\phi\, d\mu) > -\theta, \tag{9.6.2}$$

if $|x-y|<\varepsilon$, then

$$|\phi(x)-\phi(y)|<\theta \tag{9.6.3}$$

and

$$|\psi_i(x)-\psi_i(y)| < \frac{1}{2}\theta \tag{9.6.4}$$

for all $i=1,2,\ldots,s$.

Set $\beta=r(\delta)/2$ and fix a finite $\beta/2$-spanning set of $\pi(\tilde X(\delta))$, say $\{x_1,\ldots,x_t\}$. That is $B(x_1,\beta/2)\cup\ldots\cup B(x_t,\beta/2)\supset\pi(\tilde X(\delta/2))$. Let $\mathcal U$ be a finite partion of X with diameter $<\beta/2$ and let n_1 be sufficiently large that

$$\exp(-n_1\chi') < \min\{\beta/3, K^{-1}\}. \tag{9.6.5}$$

Given $n \ge 1$ define

$$\begin{aligned}\tilde{X}_{n,s} = \{\tilde{x} \in \tilde{X}(\delta) : \tilde{f}^q(\tilde{x}) \in \tilde{X}(\delta) \ \text{ and } \ \pi(\tilde{f}^q(\tilde{x})) \in \mathcal{U}(\tilde{x}) \\ \text{for some } \ q \in [n+1, (1+\theta)n] \\ \text{and } \left|\frac{1}{k} S_k(\psi_i)(\pi(\tilde{x})) - \int \psi_i \, d\mu\right| < \frac{1}{2}\theta \\ \text{for every } k \ge n \ \text{ and all } \ i = 1, 2, \ldots, s\}.\end{aligned}$$

By Birkhoff's ergodic theorem $\lim_{n\to\infty} \mu(\tilde{X}_{n,s}) = \mu(\tilde{X}(\delta)) > 1 - \delta$. Therefore there exists $n \ge n_1$ so large that $\mu(\tilde{X}_{n,s}) > 1 - \delta$. Let $X_{n,s} = \pi((\tilde{X}_{n,s}))$. Then $\mu(X_{n,s}) > 1 - \delta$ and let $E_n \subset X_{n,s}$ be a maximal (n, ε)-separated subset of $X_{n,s}$. Then E_n is a spanning set of $X_{n,s}$ and therefore it follows from (8.6.2) that for all n large enough

$$\frac{1}{n} \log \sum_{x \in E_n} \exp S_n\phi(x) - (\mathrm{h}_\mu(f) + \int \phi \, d\mu) > -\theta.$$

Equivalently

$$\sum_{x \in E_n} \exp(S_n\phi(x)) > \exp(n(\mathrm{h}_\mu(f) + \int \phi \, d\mu - \theta)).$$

For every $q \in [n+1, (1+\theta)n]$ let

$$V_q = \{x \in E_n : f^q(x) \in \mathcal{U}(x)\}$$

and let $m = m(n)$ be a value of q that maximizes $\sum_{x \in V_q} \exp(S_n\phi(x))$. Since $\bigcup_{q=n+1}^{(1+\theta)n} V_q = E_n$, we thus obtain

$$\begin{aligned}\sum_{x \in V_m} \exp S_n\phi(x) &\ge (n\theta)^{-1} \sum_{q=n+1}^{(1+\theta)n} \sum_{x \in V_q} \exp S_n\phi(x) \\ &\ge (n\theta)^{-1} \sum_{x \in E_n} \exp(S_n\phi(x)) \ge \exp(n(\mathrm{h}_\mu(f) + \int \phi \, d\mu - 2\theta)).\end{aligned}$$

Consider now the sets $V_m \cap B(x_j, \beta/2)$, $1 \le j \le t$ and choose the value $i = i(m)$ of j that maximizes $\sum_{x \in V_m \cap B(x_j, \beta/2)} \exp(S_n\phi(x))$. Thus, writing D_m for $V_m \cap B(x_{i(m)}, \beta/2)$ we have $V_m = \bigcup_{j=1}^t V_m \cap B(x_i, \beta/2)$ and

$$\sum_{x \in D_m} \exp S_n\phi(x) \ge \frac{1}{t} \exp(n(\mathrm{h}_\mu(f) + \int \phi \, d\mu - 2\theta)).$$

Since ϕ is positive, this implies that

$$\text{(9.6.6)} \qquad \sum_{x \in D_m} \exp S_m\phi(x) \ge \frac{1}{t} \exp(n(\mathrm{h}_\mu(f) + \int \phi \, d\mu - 2\theta)).$$

Now, if $x \in D_m$, then $|f^m(x) - x_i| \le |f^m(x) - x| + |x - x_i| < \beta/2 + \beta/2 = \beta$ and therefore

$$f^m(x) \in B(x_i, \beta) \subset B(f^m(x), 2\beta).$$

Thus, by (9.6.1) and as $m \ge n \ge n_1$, we have $\operatorname{diam}\big(f_x{-}m(B(f^m(x), 2\beta)\big) \le \exp(-m\chi') < \beta/3$, where $\tilde{x} \in \pi^{-1}(x) \cap \tilde{X}_{n,s}$. Therefore

$$f_x^{-m}(B(x_i, \beta)) \subset B\left(x_i, \frac{\beta}{2} + \frac{\beta}{3}\right) = B\left(x_i, \frac{5}{6}\beta\right)$$

In particular

$$\overline{f_x^{-m}(B(x_i, \beta))} \subset B(x_i, \beta) \tag{9.6.7}$$

Consider now two distinct points $y_1, y_2 \in D_m$. Then $f_{y_2}^{-m}(B(x_i, \beta)) \cap f_{y_1}^{-m}(B(x_i, \beta)) = \emptyset$ and decreasing β a little bit, if necessary, we may assume that

$$f_{y_2}^{-m}(\overline{B(x_i, \beta)}) \cap f_{y_1}^{-m}(\overline{B(x_i, \beta)}) = \emptyset.$$

Let

$$\xi = \min\left\{\beta, \min\left\{\operatorname{dist}\big(f_{y_2}^{-m}(\overline{B(x_i, \beta)}), f_{y_1}^{-m}(\overline{B(x_i, \beta)}\big) : y_1, y_2 \in D_m, y_1 \ne y_2\right\}\right\}.$$

Define now inductively the sequence of sets $\{X^{(j)}\}_{j=0}^{\infty}$ contained in $U(f)$ by setting

$$X^{(0)} = \overline{(B(x_i, \beta)} \quad \text{and} \quad X^{(j+1)} = \bigcup_{x \in D_m} f_{x_0}^{-m}(X^{(j)})$$

By (9.6.7),$\{X^{(j)}\}_{j=0}^{\infty}$, is a descending sequence of non-empty compact sets, and therefore the intersection

$$X^* = X^*(\theta, s) = \bigcap_{j=0}^{\infty} X^{(j)}$$

is also a non-empty compact set. Moreover, by the construction $f^m(X^*) = X^*$, $f^m|_{X^*}$ is topologically conjugate to the full one-sided shift generated by an alphabet consisting of $\#D_m$ elements and it immediately follows from Corollary 9.2.4 that $f^m|_{X^*}$ is an expanding map. Since $f^m|_{X^*}$ is an open map, it is straightforward to check that the triple (f^m, X^*, U_m) is a conformal expanding repeller with a sufficiently small neighborhood U_m of X^*. Thus $(f, X(\theta, s), W_s)$, is also a conformal expanding repeller, where

$$X(\theta, s) = \bigcup_{l=0}^{m-1} f^l(X^*) \quad \text{and} \quad W_s = \bigcup_{l=0}^{m-1} f^l(U_m)$$

Fix now an integer $j \geq 1$. For any j-tuple $(z_0, z_1, \ldots, z_{j-1})$, $z_l \in D_m$ choose exactly one point y from the set $f_{z_{j-1}}^{-m} \circ f_{z_{j-2}}^{-m} \circ \ldots \circ f_{z_0}^{-m}(X^*)$ and denote the made up set by A_j. Since by (9.6.3) and (9.6.5) $S_{jm}\phi(y) \geq \sum_{l=0}^{j-1} S_m\phi(z_l) - jm\theta$ we see that

$$\sum_{y \in A_j} \exp S_{jm}\phi(y) \geq \Big(\sum_{x \in D_m} \exp S_m\phi(x)\Big)^j \exp(-jm\theta)$$

and

$$\frac{1}{j} \log \sum_{y \in A_j} \exp S_{jm}\phi(y) \geq \log \sum_{x \in D_m} \exp S_m\phi(x) - m\theta$$

In view of the definition of ξ, the set A_j is (j, ξ)-separated for f^m and ξ is an expansive constant for f^m. Hence, letting $j \to \infty$ we obtain

$$\begin{aligned} \mathrm{P}(f^m|_{X^*}, S_m\phi) &\geq \log \sum_{x \in D_m} \exp S_m\phi(x) - m\theta \\ &\geq n\big(\mathrm{h}_\mu(f) + \int \phi \, d\mu - 2\theta\big) - \log t - m\theta \end{aligned}$$

where the last inequality was written in view of (9.6.6). Since $n+1 \leq m \leq n(1+\theta)$ and since $\inf \phi > 0$ (and consequently $\mathrm{h}_\mu(f) + \int \phi \, d\mu > 0$), we get

$$\begin{aligned} \mathrm{P}(f|_{X(\theta,s)}, \phi) &= \frac{1}{m}\mathrm{P}(f^m|_{X(\theta,s)}, \phi) \geq \frac{1}{m}\mathrm{P}(f^m|_{X^*}, S_m\phi) \\ &\geq \frac{1}{1+\theta}\left(\mathrm{h}_\mu(f) + \int \phi \, d\mu - 2\theta\right) - \frac{\log t}{m} - \theta \end{aligned}$$

Supposing now that n (and consequently also m) was choosen sufficiently large we get

$$\mathrm{P}(f|_X(\theta, s), \phi) \geq \frac{1}{1+\theta}(\mathrm{h}_\mu(f) + \int \phi \, d\mu) - 4\theta.$$

If now ν is any ergodic f-invariant measure on $X(\theta, s)$, then it follows from the definition of the set $\tilde{X}_{n,s}$, the construction of the set $X(\theta, s)$ and the Birkhoff ergodic theorem that $|\int \psi_i \, d\nu - \int \psi_i \, d\mu| < \theta$ for every $i = 1, 2, \ldots, s$. Therefore putting for example $X_k = X(1/k, k)$, completes the proof of Theorem 9.6.1. ♣

Remark 9.6.2. If the set X is repelling, that is if $\bigcap_{n \geq 0} f^{-n}(U) = X$, then the sets X_k constucted in the proof of Theorem 9.6.1 are all contained in X. In particular we get the following.

Corollary 9.6.3. If the set X is repelling and if $\mathrm{P}(f, \phi) > \sup \phi$, then there exists a sequence X_k, $k = 1, 2, \ldots$, of compact f-invariant subsets of X such that for every k, $f|_{X_k}$ is a conformal expanding repeller,

$$\lim_{k \to \infty} \mathrm{P}(f|_{X_k}, \phi) = \mathrm{P}(f, \phi)$$

and if μ_k is any ergodic f-invariant measure on X_k, then the sequence μ_k, $k = 1, 2, \ldots$, converges to μ in the weak-*-topology on X.

Remark 9.6.4. Of course in Corollary 9.6.3 was sufficient to assume that $\mathrm{P}(f, \phi) = \sup\{\mathrm{h}_\mu(f) + \int \phi \, d\mu\}$ where the supremum is taken over all ergodic invariant measures of positive entropy, which is assured for eaxample by inequality $\mathrm{P}(f, \phi) > \sup \phi$. Besides, if the function ϕ has an equilibrium state of positive entropy, then the sequence μ_k can be choosen to converge to this equilibrium state.

Our last immediate conclusion concerns periodic points.

Corollary 9.6.5. If $f : X \to X$ is repelling and $\mathrm{h}_{\mathrm{top}}(f) > 0$, then f has infinitely many periodic points. Moreover the number of periodic points of period n grows exponentialy fast with n.

CHAPTER 10

CONFORMAL MEASURES

§10.1. GENERAL NOTION OF CONFORMAL MEASURES.

Let $T : X \to X$ be a continuous map of a compact metric space (X, ρ) and let $g : X \to I\!\!R$ be a non–negative measurable function. A Borel probability measure m on X is said to be g–conformal for $T : X \to X$ if

$$m(T(A)) = \int_A g\, dm \tag{10.1.1}$$

for any Borel set $A \subset X$ such that $T|_A$ is injective. Sets with this property will be called special. There is a close relation between conformal measures and Perron-Frobenius type operators. In order to describe it notice first that

$$\int_{T(A)} \phi\, dm = \int_A (\phi \circ T) g\, dm \tag{10.1.2}$$

for any Borel function ϕ and any special set A. Assume now (only till Proposition 10.1.1) additionally that T is bounded–to–one (i.e. the numbers of preimages of points are uniformly bounded) and that g takes values in $I\!\!R_+$. Define then the Perron–Frobenius operator $\mathcal{L}_g$, associated to T and g, putting for a measurable function $\phi : X \to I\!\!R_+$

$$\mathcal{L}_g \phi(x) = \sum_{T(y)=x} \frac{\phi(y)}{g(y)}$$

$\mathcal{L}_g \phi$ is a well defined measurable function. We shall prove the following.

Proposition 10.1.1. Assume that there exists a finite partition of X into special sets X_i $(1 \le i \le s)$, such that all the maps $T : X_i \to T(X_i)$ are measurable isomorphisms. Then m is g–conformal if and only if $\mathcal{L}_g$ acts on $L^1(m)$ and $\mathcal{L}_g^* m = m$, where $\mathcal{L}_g^*$ is the operator conjugate with $\mathcal{L}_g$.

Proof. Let m be g–conformal and let $\phi \in L^1(m)$. By (10.1.2)

$$\int_{T(X_i)} \frac{\phi}{g} \circ \left(T|_{X_i}\right)^{-1} dm = \int_{X_i} \phi\, dm$$

for every $i = 1, \ldots, s$. Thus, summing over all i yields

$$\int_X \mathcal{L}_g \phi \, dm = \int_X \phi \, dm$$

Conversely, assume that $\mathcal{L}_g$ acts on $L^1(m)$ and that m is a fixed point of $\mathcal{L}_g^*$. Let A be a special set. Then, by the definition of the Perron–Frobenius operator

$$\int_A g \, dm = \int_A f \, d(\mathcal{L}_g^* m) = \int_X \mathcal{L}_g(1_A g) \, dm = \int_X \sum_{T(y)=x} 1_A(y) \, m(dx) = m(T(A))$$

Thus m is f-conformal. ♣

Now we shall provide a general method of constructing conformal measures. The construction will make use of the following simple analytical fact. For a sequence $\{a_n : n \geq 1\}$ of reals the number

$$c = \limsup_{n \to \infty} \frac{a_n}{n} \tag{10.1.3}$$

will be called the transition parameter of $\{a_n : n \geq 1\}$. It is uniquely determined by the property that

$$\sum_{n \geq 1} \exp(a_n - ns)$$

converges for $s > c$ and diverges for $s < c$. For $s = c$ the sum may converge or diverge. By a simple argument one obtains the following.

Lemma 10.1.2. There exits a sequence $\{b_n : n \geq 1\}$ of positive reals such that

$$\sum_{n=1}^{\infty} b_n \exp(a_n - ns) \begin{cases} < \infty & s > c \\ = \infty & s \leq c \end{cases}$$

and $\lim_{n \to \infty} \frac{b_n}{b_{n+1}} = 1$.

Proof. If $\sum \exp(a_n - nc) = \infty$, put $b_n = 1$ for every $n \geq 1$. If $\sum \exp(a_n - nc) < \infty$, choose a sequence $\{n_k : k \geq 1\}$ of positive integers such that $\lim_{k \to \infty} n_k n_{k+1}^{-1} = 0$ and $\varepsilon_k := a_{n_k} n_k^{-1} - c \to 0$. Setting

$$b_n = \exp\left(n \left(\frac{n_k - n}{n_k - n_{k-1}} \varepsilon_{k-1} + \frac{n - n_{k-1}}{n_k - n_{k-1}} \varepsilon_k \right) \right) \qquad \text{for } n_{k-1} \leq n < n_k,$$

it is easy to check that the lemma follows. ♣

Getting back to dynamics let $\{E_n\}_{n=1}^{\infty}$ be a sequence of finite subsets of X such that

$$T^{-1}(E_n) \subset E_{n+1} \qquad \text{for every} \quad n \geq 1 \tag{10.1.4}$$

and let

$$a_n = \log\big(\sum_{x\in E_n} \exp(S_n g(x))\big)$$

where $S_n g = \sum_{0\le k<n} g \circ T^k$. Denote by c the transition parameter of this sequence. Choose a sequence $\{b_n : n \ge 1\}$ of positive reals as in lemma 10.1.2 for the sequence $\{a_n : n \ge 1\}$. For $s > c$ define

$$M_s = \sum_{n=1}^{\infty} b_n \exp(a_n - ns) \tag{10.1.5}$$

and the normalized measure

$$m_s = \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in E_n} b_n \exp(S_n g(x) - ns)\delta_x, \tag{10.1.6}$$

where δ_x denotes the unit mass at the point $x \in X$. Let A be a special set. Using (10.1.4) and (10.1.6) it follows that

$$\begin{aligned} m_s(T(A)) &= \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in E_n\cap T(A)} b_n \exp(S_n g(x) - ns)\\ &= \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in A\cap T^{-1}E_n} b_n \exp(S_n g(T(x)) - ns)\\ &= \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in A\cap E_{n+1}} b_n \exp[S_{n+1} g(x) - (n+1)s]\exp(s - g(x))\\ &\quad - \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in A\cap(E_{n+1}\setminus T^{-1}E_n)} b_n \exp(S_n g(T(x)) - ns). \end{aligned} \tag{10.1.7}$$

Set

$$\Delta_A(s) =$$

$$= \left|\frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in A\cap E_{n+1}} b_n \exp[S_{n+1} g(x) - (n+1)s]\exp(s - g(x)) - \int_A \exp(c-g)\, dm_s\right|$$

and observe that

$$
\begin{aligned}
&\Delta_A(s) = \\
&\quad = \frac{1}{M_s}\Bigg|\sum_{n=1}^{\infty}\sum_{x\in A\cap E_{n+1}} \exp[S_{n+1}g(x)-(n+1)s]\exp(-g(x))(b_ne^s - b_{n+1}e^c) \\
&\qquad - b_1\sum_{x\in A\cap E_1} e^{c-s}\Bigg| \\
&\quad \le \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in A\cap E_{n+1}}\left|\frac{b_n}{b_{n+1}} - e^{c-s}\right| b_{n+1}\exp(s-g(x))\exp\big(S_{n+1}g(x)-(n+1)s\big) \\
&\qquad + \frac{1}{M_s}b_1\exp(c-s)\,\sharp(A\cap E_1) \\
&\quad \le \frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in E_{n+1}}\left|\frac{b_n}{b_{n+1}} - e^{c-s}\right| b_{n+1}\exp(s-g(x))\exp\big(S_{n+1}g(x)-(n+1)s\big) \\
&\qquad + \frac{1}{M_s}b_1\exp(c-s)\,\sharp E_1.
\end{aligned}
$$

By lemma 10.1.2 we have $\lim_{n\to\infty} b_{n+1}/b_n = 1$ and $\lim_{s\searrow c} M_s = \infty$. Therefore

$$\lim_{s\searrow c}\Delta_A(s) = 0 \tag{10.1.8}$$

uniformly for all special sets A.

Any weak accumulation point, when $s\downarrow c$, of the measures $\{m_s : s > c\}$ defined by (10.1.6) will be called a limit measure (associated to the function g and the sequence $\{E_n : n \ge 1\}$).

In order to find conformal measures among the limit measures, it is necessary to examine (10.1.7) in greater detail. To beginn with, for a Borel set $D \subset X$, consider the following condition

$$\lim_{s\searrow c}\frac{1}{M_s}\sum_{n=1}^{\infty}\sum_{x\in D\cap(E_{n+1}\setminus T^{-1}E_n)} b_n\exp\big(S_ng(T(x))-ns\big) = 0. \tag{10.1.9}$$

We will need the following definitions.

A point $x\in X$ is said to be singular for T if at least one of the following two conditions is satisfied:

(10.1.10) There is no open neighbourhood U of x such that $T|_U$ is injective.
(10.1.11) For all $\varepsilon > 0$ there exists an open set $U\subset B(x,\varepsilon)$ such that $T(U)$ is not an open subset of X.

The set of all singular points is denoted by $\mathrm{Sing}(T)$, the set of all points satisfying condition (10.1.10) is denoted by $\mathrm{Crit}(T)$ and the set of all points satisfying condition (10.1.11) is denoted by $X_0(T)$.

It is easy to give examples where $X_0 \cap \mathrm{Crit}(T) \neq \emptyset$. If $T : X \to X$ is an open map, no point satisfies condition (10.1.11), that is $X_0(T) = \emptyset$.

In spite of what was assumed in [ECM] and similarly as in [Sul], the set $\mathrm{Sing}(T)$ is not required to be finite. Let us prove the following.

Lemma 10.1.3. Let m be a Borel probability measure on X and let Γ be a compact set containing $\mathrm{Sing}(T)$. If (10.1.1) holds for every special set A whose closure is disjoint from Γ and such that $m(\partial A) = m(\partial T(A)) = 0$, then (10.1.1) continues to hold for every special set A disjoint from Γ.

Proof. Let A be a special set disjoint from Γ. Fix $\varepsilon > 0$. Since on the complement of Γ the map T is open, for each point $x \in A$ there exists an open neighbourhood $U(x)$ of x such that $T|_{U(x)}$ is a homeomorphism, $m(\partial U(x)) = m(\partial T(U(x))) = 0$, $\overline{U(x)} \cap \Gamma = \emptyset$ and such that

$$\int_{\cup U(x)\setminus A} g\, dm < \varepsilon$$

Choose a countable family $\{U_k\}$ from $\{U(x)\}$ which covers A and define recursively $A_1 = U_1$ and $A_n = U_n \setminus \bigcup_{k<n} U_k$. By the assumption of the lemma, each set A_k satisfies (10.1.1) and hence

$$\begin{aligned} m(T(A)) &= m(\bigcup_{k=1}^{\infty} T(A \cap A_k)) \leq \sum_{k=1}^{\infty} m(T(A_k)) \\ &= \sum_{k=1}^{\infty} \int_{A_k} g\, dm = \int_A g\, dm + \sum_{k=1}^{\infty} \int_{A_k \setminus A} g\, dm \\ &\leq \int_A g\, dm + \varepsilon. \end{aligned}$$

If $\varepsilon \to 0$, it follows that

$$m(T(B)) \leq \int_B g\, dm$$

for any special set B disjoint from Γ. Using this fact, the lower bound for $m(T(A))$ is obtained from the following estimate, if $\varepsilon \to 0$:

$$\begin{aligned} m(T(A)) &= m(\bigcup_{k=1}^{\infty} T(A \cap A_k)) = \sum_{k=1}^{\infty} m(T(A \cap A_k)) \\ &= \sum_{k=1}^{\infty} (m(T(A_k)) - m(T(A_k \setminus A))) \geq \sum_{k=1}^{\infty} \left(\int_{A_k} g\, dm - \int_{A_k \setminus A} g\, dm \right) \\ &= \int_{\cup_{k\geq 1} A_k} g\, dm - \int_{\cup_{k\geq 1} A_k \setminus A} g\, dm \geq \int_A g\, dm - \varepsilon. \end{aligned}$$

This proves the lemma. ♣

Lemma 10.1.4. Let m be a limit measure and let Γ be a compact set containing $\mathrm{Sing}(T)$. Assume that every special set $D \subset X$ with $m(\partial D) = m(\partial T(D)) = 0$ and $\bar{D} \cap \Gamma = \emptyset$ satisfies condition (10.1.9). Then $m(T(A)) = \int_A \exp(c - f)\, dm$ for every special set A disjoint from Γ.

Proof. Let $D \subset X$ be a special set such that $\bar{D} \cap \Gamma = \emptyset$ and $m(\partial D) = m(\partial T(D)) = 0$. It follows immediately from (10.1.7)–(10.1.9) that $m(T(D)) = \int_D \exp(c - f)\, dm$. Applying now Lemma 10.1.3 completes the proof. ♣

Lemma 10.1.5 Let m be a limit measure. If condition (10.1.9) is satisfied for $D = X$, then $m(T(A)) \geq \int_A \exp(c - f)\, dm$ for every special set A disjoint from $\mathrm{Crit}(T)$.

Proof. Suppose first that A is compact and $m(\partial A) = 0$. From (10.1.7), (10.1.8) and the assumption one obtains

$$\lim_{s \in J} \mid m_s(T(A)) - \int_A \exp(c - f)\, dm_s \mid = 0$$

where J denotes the subsequence along which m_s converges to m. Since $T(A)$ is compact, this implies

$$m(T(A)) \geq \liminf_{s \in J} m_s(T(A)) = \lim_{s \in J} \int_A \exp(c - f)\, dm_s = \int_A \exp(c - f)\, dm$$

Now, drop the assumption $m(\partial A) = 0$ but keep A compact and assume additionally that for some $\varepsilon > 0$ the ball $B(A, \varepsilon)$ is also special. Choose a descending sequence A_n of compact subsets of $B(A, \varepsilon)$ whose intersection equals A and $m(\partial A_n) = 0$ for every $n \geq 0$. By what has been already proved

$$m(T(A)) = \lim_{n \to \infty} m(T(A_n)) \geq \int_{A_n} \exp(c - f)\, dm = \int_A \exp(c - f)\, dml$$

The next step is to prove the lemma for A, an arbitrary open special set disjoint from $\mathrm{Crit}(T)$ by partitioning it by countably many compact sets. Then one approximates from above special sets of sufficiently small diameters by special open sets and the last step is to partition an arbitrary special set disjoint from $\mathrm{Crit}(T)$ by sets of so small diameters that the lemma holds. ♣

Lemma 10.1.6 Let Γ be a compact subset of X containing $\mathrm{Sing}(T)$. Suppose that for every integer $n \geq 1$ there are a continuous function $g_n : X \to X$ and a measure m_n on X satisfying (10.1.1) for $g = g_n$ and for every special set $A \subset X$ with

(a) $$\overline{A} \cap \Gamma = \emptyset$$

and satisfying

$$m_n(B) \geq \int_B g_n\, dm_n$$

for any special set $B \subset X$ such that $B \cap \mathrm{Crit}(T) = \emptyset$. Suppose, moreover, that the sequence $\{g_n\}_{n=1}^\infty$ converges uniformly to a continuous function $g : X \to I\!R$. Then for any weak accumulation point m of the sequence $\{m_n\}_{n=1}^\infty$ we have

$$\text{(b)} \qquad m(T(A)) = \int_A g\, dm$$

for all special sets $A \subset X$ such that $A \cap \Gamma = \emptyset$ and

$$\text{(c)} \qquad m(T(B)) \geq \int_B g\, dm$$

for all special sets $B \subset X$ such that $B \cap \mathrm{Crit}(T) = \emptyset$.

Moreover, if (a) is replaced by

$$\text{(a')} \qquad \overline{A} \cap (\Gamma \setminus (\mathrm{Crit}(T) \setminus X_0(T))) = \emptyset,$$

then for any $x \in \mathrm{Crit}(T) \setminus X_0(T)$

$$\text{(d)} \qquad m(\{T(x)\}) \leq g(x) m(\{x\}) \leq q(x) m(\{T(x)\})$$

where $q(x)$ denotes the maximal number of preimages of single points under the transformation T restricted to a sufficiently small neighbourhood of x.

The proof of property (b) is a simplification of the proof of Lemma 10.1.4 and the proof of property (c) is a simplification of the proof of Lemma 10.1.5. The proof of (d) uses the same technics and is left for the reader.

§10.2. SULLIVAN'S CONFORMAL MEASURES AND DYNAMICAL DIMENSION, I.

Let, as in Chapter 9, X denote a compact subset of the extended complex plane $\overline{\mathbb{C}}$ and let $f \in \mathcal{A}(X)$ which means that $f : X \to X$ is a continuous map that can be analytically extended to an open neighbourhood $U(f)$ of X.

Let $t \geq 0$. Any $|f'|^t$–conformal measure for $f : X \to X$ is called a t–conformal Sullivan's measure or even shorter a t–conformal measure. Rewritting the defintion (10.1.1) it means that

$$\text{(10.2.1)} \qquad m(f(A)) = \int_A |f'|^t\, dm$$

for every special set $A \subset X$. An obvious but important property of conformal measures is formulated in the following

Lemma 10.2.1. If $f : X \to X$ is locally eventually onto, then every Sullivan's conformal measure is positive on nonempty open sets of X.

In particular it follows from this lemma that if f is locally eventually onto, then for every $r > 0$

$$M(r) = \inf\{m(B(x,r)) : x \in X\} > 0 \tag{10.2.2}$$

Denote by $\delta(f)$ the infinium over all exponents $t \geq 0$ for which a t–conformal measure for $f : X \to X$ exists.

Our aim in the two subsequent sections is to show the existence of conformal measures and even more to establish more explicite dynamical characterization of the number $\delta(f)$. As a matter of fact we are going to prove that under some additional assumptions $\delta(f)$ concides with the dynamical dimension $\mathrm{DD}(X)$ of X and the hyperbolic dimension $\mathrm{HyD}(X)$ of X which is defined as follows.

$$\mathrm{DD}(X) = \sup\{\mathrm{HD}(\mu) : \mu \in M_e^+(f)\}$$
$$\mathrm{HyD}(X) = \sup\{\mathrm{HD}(Y) : f|_Y \text{ is a conformal expanding repeller}\}$$

In this section we shall prove the following two results.

Lemma 10.2.2. If $f : X \to X$ is locally eventually onto, then $\mathrm{DD}(X) \leq \delta(f)$.

Proof. Our main idea "to get to a large scale" is the same as in [SulDU]. However to carry it out we use Pesin theory described in Ch.8.2 instead of Mane's partition introduced in [Mane] and applied in [SulDU]. So, let $\mu \in M_e^+(f)$ and let m be a t–conformal measure. We again work in the natural extension $(\tilde{X}, \tilde{f}, \tilde{\mu})$. Fix $\varepsilon > 0$ and let $\tilde{X}(\varepsilon)$ and $r(\varepsilon)$ be given by Corollary 9.2.4. In view of the Birkhoff ergodic theorem there exist a measurable set $\tilde{F}(\varepsilon) \subset \tilde{X}(\varepsilon)$ such that $\tilde{\mu}(\tilde{F}(\varepsilon)) = \tilde{\mu}(\tilde{X}(\varepsilon))$ and an increasing sequence $\{n_k = n_k(\tilde{x}) : k \geq 1\}$ such that $\tilde{f}^{n_k}(\tilde{x}) \in \tilde{X}(\varepsilon)$ for every $k \geq 1$. Let $F(\varepsilon) = \pi(\tilde{F}(\varepsilon))$. Then $\mu(F(\varepsilon)) = \tilde{\mu}(\pi^{-1}(F(\varepsilon)) \geq \tilde{\mu}(\tilde{F}(\varepsilon)) \geq 1 - 2\varepsilon$. Consider now $x \in F(\varepsilon)$ and take $\tilde{x} \in \tilde{F}(\varepsilon)$ such that $x = \pi(\tilde{x})$. Since $\tilde{f}^{n_k}(\tilde{x}) \in \tilde{X}(\varepsilon)$ and since $\pi(\tilde{f}^{n_k}(\tilde{x}) = f^{n_k}(x)$, Corollary 9.2.4 produces a holomorphic inverse branch $f_x^{-n_k} : B(f^{n_k}(x), r(\varepsilon)) \to \overline{\mathbb{C}}$ of f^{n_k} such that $f_x^{-n_k} f^{n_k}(x) = x$ and

$$f_x^{-n_k}\big(B(f^{n_k}(x), r(\varepsilon))\big) \subset B\big(x, K|(f^{n_k})'(x)|^{-1} r(\varepsilon)\big)$$

Set $r_k(x) = K|(f^{n_k})'(x)|^{-1} r(\varepsilon)$. Then by Corollary 9.2.4 and t– conformality of m

$$m(B(x, r_k(x))) \geq K^{-t}|(f^{n_k})'(x)|^{-t} m\big(B(f^{n_k}(x), r(\varepsilon))\big) \geq M(r(\varepsilon))^{-1} K^{-2t} r(\varepsilon)^{-t} r_k(x)^t$$

Therefore, it follows from Theorem 5.5.1 (Besicovitch covering theorem) that $\mathrm{H}_t(F(\varepsilon)) \leq M(r(\varepsilon))K^{2t} r(\varepsilon)^t b(2) < \infty$. Hence $\mathrm{HD}(F(\varepsilon)) \leq t$. Since $\mu\big(\bigcup_{n=1}^\infty F(1/n)\big) = 1$, it implies that $\mathrm{HD}(\mu) \leq t$. This finishes the proof. ♣

Theorem 10.2.3. If $f : X \to X$ is locally eventually onto and X is a repelling set for f, then $\mathrm{HyD}(X) = \mathrm{DD}(X)$.

Proof. In order to see that $\mathrm{HyD}(X) \le \mathrm{DD}(X)$ notice only that in view of Theorem 7.1.1 there exists $\mu \in M_e^+(f|_Y) \subset M_e^+(f)$ such that $\mathrm{HD}(\mu) = \mathrm{HD}(Y)$. In order to prove that $\mathrm{DD}(X) \le \mathrm{CD}(X)$ we will use Katok's theory from Section 9.6 applied to μ, an arbitrary ergodic invariant measure of positive entropy. First, for every integer $n \ge 0$ define on X a new continuous function

$$\phi_n = \max\{-n, \log|f'|\}.$$

Then $\phi_n \ge \log|f'|$ and $\phi_n \searrow \log|f'|$ pointwise on X. Since in addition $\phi_n \le \log||f'||$, it follows from the Lebesgue monotone convergence theorem that $\lim_{n\to\infty}\int \phi_n\,d\mu = \chi_\mu(f) = \int \log|f'|\,d\mu > 0$. Fix $\varepsilon > 0$. Then for all n sufficiently large, say $n \ge n_0$, $\int \phi_n d\mu \le \chi_\mu/(1-\varepsilon)$ which implies that

$$\mathrm{h}_\mu(f) = \mathrm{HD}(\mu)\chi_\mu \ge (1-\varepsilon)\mathrm{HD}(\mu)\int \phi_n\,d\mu. \tag{10.2.4}$$

Fix such $n \ge n_0$. Let $X_k \subset X$, $k \ge 0$, be the sequence of conformal expanding repellers produced in Theorem 9.6.1 for the measure μ and the function $-\mathrm{HD}(\mu)\phi_n$ and let μ_k be an equilibrium state of the map $f|_{X_k}$ and the potential $-\mathrm{HD}(\mu)\phi_n$ restricted to X_k. It follows from the second part of Theorem 9.6.1 that $\lim_{k\to\infty}\int\phi_n\,d\mu_k = \int\phi_n\,d\mu > 0$. Thus by Theorem 10.6.1 and (10.2.4)

$$\begin{aligned}\liminf_{k\to\infty}\big(\mathrm{h}_{\mu_k} - \mathrm{HD}(\mu)\int\phi_n\,d\mu_k\big) &= \liminf_{k\to\infty}\mathrm{P}\big(f|_{X_k}, -\mathrm{HD}(\mu)\phi_n\big)\\ &\ge \mathrm{h}_\mu(f)\mathrm{HD}(\mu)\int\phi_n\,d\mu\\ &\ge -\mathrm{HD}(\mu)\int\phi_n\,d\mu\end{aligned}$$

Hence, for all k large enough

$$\begin{aligned}\mathrm{h}_{\mu_k} &\ge \mathrm{HD}(\mu)\int\phi_n\,d\mu_k - 2\varepsilon\mathrm{HD}(\mu)\int\phi_n\,d\mu \ge \mathrm{HD}(\mu)\int\phi_n\,d\mu_k - 3\varepsilon\mathrm{HD}(\mu)\int\phi_n\,d\mu_k\\ &= (1-3\varepsilon)\mathrm{HD}(\mu)\int\phi_n\,d\mu_k \ge (1-3\varepsilon)\mathrm{HD}(\mu)\int\log|f'|\,d\mu_k.\end{aligned}$$

Thus

$$\mathrm{HD}(X_k) \ge \mathrm{HD}(\mu_k) = \frac{\mathrm{h}_{\mu_k}(f)}{\chi_{\mu_k}} \ge (1-3\varepsilon)\mathrm{HD}(\mu).$$

So, letting $\varepsilon \to 0$ finishes the proof. ♣

§10.3. SULLIVAN'S CONFORMAL MEASURES AND DYNAMICAL DIMENSION, II.

In this section $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ is assumed to be a rational map of degree ≥ 2 and X is its Julia set $J(f)$. Neverthless it is worth to mention that some results proved here continue to hold under weaker assumption that $f|_X$ is open or X is a perfect locally maximal set for f. By $Crit(f)$ we denote the set of all critical points contained in the Julia set $J(f)$.

Lemma 10.3.1. If $z \in J(f)$ and $\overline{\{f^n(z) : n \geq 0\}} \cap \mathrm{Crit}(f) = \emptyset$, then the series $\sum_{n=1}^{\infty} |(f^n)'(z)|^{\frac{1}{3}}$ diverges.

Proof. By the assumption there exists $\varepsilon > 0$ such that for every $n \geq 0$ the map f restricted to the ball $B(f^n(z), \varepsilon)$ is injective. Since f is uniformly continuous there exists $0 < \alpha < 1$ such that for every $x \in \overline{\mathbb{C}}$

$$f(B(x, \alpha\varepsilon)) \subset B(f(x), \varepsilon). \tag{10.3.1}$$

Suppose that the series $\sum_{n=1}^{\infty} |(f^n)'(z)|^{\frac{1}{3}}$ converges. Then there exists $n_0 \geq 1$ such that $\sup_{n \geq n_0} (2|(f^n)'(z)|)^{\frac{1}{3}} < 1$. Choose $0 < \varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_{n_o} < \alpha\varepsilon$ so small that for every $n = 1, 2, \ldots, n_0$

$$f^n \text{ restricted to the ball } B(z, \varepsilon_n) \text{ is injective.} \tag{10.3.2}$$

and

$$f^n(B(z, \varepsilon_n)) \subset B(f^n(z), \varepsilon) \tag{10.3.3}$$

For every $n \geq n_0$ define ε_{n+1} inductively by

$$\varepsilon_{n+1} = (1 - (2|(f^n)'(z)|)^{\frac{1}{3}})\varepsilon_n. \tag{10.3.4}$$

Then $0 < \varepsilon_n < \alpha\varepsilon$ for every $n \geq 1$. Assume that (10.3.2) and (10.3.3) are satisfied for some $n \geq n_0$. Then by the Köbe Distortion Theorem ??? and (10.3.4) the set $f^n(B(z, \varepsilon_{n+1}))$ is contained in the ball centered at $f^n(z)$ and of radius

$$\varepsilon_{n+1}|(f^n)'(z)| \frac{2}{(1 - \varepsilon_{n+1}/\varepsilon_n)^3} = \frac{2\varepsilon_{n+1}|(f^n)'(z)|}{2|(f^n)'(z)|} = \varepsilon_{n+1} < \alpha\varepsilon.$$

Therefore, since f is injective on $B(f^n(z), \varepsilon)$, formula (10.3.2) is satisfied for $n+1$ and using also (10.3.1) we get

$$f^{n+1}(B(z, \varepsilon_{n+1})) = f\big(f^n(B(z, \varepsilon_{n+1}))\big) \subset f(B(f^n(z), \alpha\varepsilon)) \subset B(f^{n+1}(z), \varepsilon).$$

Thus (10.3.3) is satisfied for $n+1$.

Let $\varepsilon_n \downarrow \varepsilon_0$. Since the series $\sum_{n=1}^{\infty} |(f^n)'(z)|^{\frac{1}{3}}$ converges, it follows from (10.3.4) that $\varepsilon_0 > 0$. Clearly (10.3.2) and (10.3.3) remain true with ε_n replaced by ε_0. It follows that the family $\{f^n|_{B(z, \frac{1}{2}\varepsilon_0)}\}_{n=1}^{\infty}$ is normal and consequently $z \notin J(f)$. This contradiction finishes the proof. ♣

As an immediate consequence of this lemma and of Birkhoff's Ergodic Theorem we get the following.

Corollary 10.3.2. If μ be an ergodic f–invariant measure for which there exists a compact set $Y \subset J(f)$ such that $\mu(Y) = 1$ and $Y \cap Crit(f) = \emptyset$, then $\chi_\mu \geq 0$.

Let now Ω be a finite subset of $\overline{\bigcup_{n=1}^\infty f^n(\mathrm{Crit}(f))}$ such that

$$\Omega \cap \overline{\{f^n(c) : n = 1, 2 \ldots\}} \neq \emptyset \qquad \text{for every } c \in \mathrm{Crit}(f) \tag{10.3.5}$$

and

$$\Omega \cap \mathrm{Crit}(f) = \emptyset. \tag{10.3.6}$$

Sets satisfying these conditions exist since no critical point of f lying in $J(f)$ can be periodic. Now let $V \subset J(f)$ be an open neighbourhood of Ω and define $K(V)$ to be the set of those points of $J(f)$ whose forward trajectory avoids V. Equivalently this means that

$$K(V) = \{z \in J(f) : f^n(z) \notin V \text{ for every } n \geq 0\} = \bigcap_{n=0}^{\infty} f^{-n}(J(f) \setminus V)$$

Hence $K(V)$ is a compact subset of $J(f)$ and $f(K(V)) \subset K(V)$. Consequently we can consider dynamical system $f|_{K(V)} : K(V) \to K(V)$. Note that $f(K(V)) = K(V)$ does not hold for all sets V and that usually $f^{-1}(K(V)) \not\subset K(V)$. Simple considerations based on (10.3.5) and the definition of sets $K(V)$ give the following.

Lemma 10.3.3. $\mathrm{Crit}(f|_{K(V)}) \subset \mathrm{Crit}(f) \cap K(V) = \emptyset$, $K(V)_0(f) = \mathrm{Sing}(f) \subset \partial V$, and $-t \log |f'|$ is a well–defined continuous function on $K(V)$.

Fix now $z \in K(V)$ and set $E_n = f|_{K(V)}^{-n}(z)$, $n \geq 0$. Then $E_{n+1} = f|_{K(V)}^{-1}(E_n)$ and therefore the sequence $\{E_n\}$ satisfies (10.1.9) with $D = K(V)$. Take $t \geq 0$ and let $c(t, V)$ be the transition parameter associated to this sequence and the function $-t \log |f'|$. Put $\mathrm{P}(t, V) = \mathrm{P}(f|_{K(V)}, -t \log |f'|)$. We shall prove the following.

Lemma 10.3.4. $c(t, V) \leq \mathrm{P}(t, V)$.

Proof. Since $K(V)$ is a compact set disjoint from $\mathrm{Crit}(f)$, the map $f|_{K(V)}$ is locally !-to-1 which means that there exists $\delta > 0$ such that $f|_{K(V)}$ restricted to any set with diameter $\leq \delta$ is !to-1. Consequently, all the sets E_n are (n, ε)-separated for $\varepsilon < \delta$. Hence, the required inequality $c(t, V) \leq \mathrm{P}(t, V)$ follows immediately from Theorem .2.2.10. ♣

The standard straightforward arguments showing continuity of topological pressure prove also the following.

Lemma 10.3.5. The function $t \mapsto c(t, V)$ is continuous.

Set

$$s(V) = \inf\{t \geq 0 : c(t, V) \leq 0\} < +\infty$$

We shall prove the following.

Lemma 10.3.6. $s(V) \le \mathrm{DD}(J(f))$.

Proof. Suppose that $\mathrm{DD}(J(f)) < s(V)$ and take $0 \le \mathrm{DD}(J(f)) < t < s(V)$. From this choice and by Lemma 10.3.4 we have $0 < c(t,V) \le \mathrm{P}(t,V)$ and by Variational Principle ??? there exists $\mu \in M_e(f_{K(V)}) \subset M_e(f)$ such that $\mathrm{P}(t,V) \le \mathrm{h}_\mu(f) - t\chi_\mu(f) + c(t,V)/2$. Therefore, by Corollary 10.3.2 and Lemma 10.3.3 we get $\mathrm{h}_\mu(f) \ge c(t,V)/2 > 0$ and applying additionally Theorem 9.1.1 (Ruelle's inequality), $\chi_\mu(f) > 0$. Hence, it follows from Theorem 9.4.1 that

$$t \le \mathrm{HD}(\mu) - \frac{1}{2}\frac{c(t,V)}{\chi_\mu} < \mathrm{HD}(\mu) \le \mathrm{DD}(J(f))$$

This contradiction finishes the proof. ♣

Let m be a limit measure on $K(V)$ associated to the sequence E_n and the function $-s(V)\log|f'|$. Since $c(0,V) \ge 0$ and $s(V) < \infty$, it follows from Lemma 10.3.5 that $c(s(V),V) = 0$. Therefore, applying Lemma 10.1.4 and Lemma 10.1.5 with $\Gamma = \partial V$ we see that $m(f(A)) \ge \int_A |f'|^{s(V)}\,dm$ for any special set $A \subset K(V)$ and $m(f(A)) = \int_A |f'|^{s(V)}\,dm$ for any special set $A \subset K(V)$ such that $A \cap \partial V = \emptyset$. Treating now m as a measure on $J(T)$ and using straightforward measure–theoretic arguments we deduce from this that

$$m(f(A)) \ge \int_A |f'|^{s(V)}\,dm \tag{10.3.7}$$

for any special set $A \subset J(f)$ and

$$m(f(A)) = \int_A |f'|^{s(V)}\,dm \tag{10.3.8}$$

for any special set $A \subset J(f)$ such that $A \cap \bar{V} = \emptyset$. Now we are in position to prove the following.

Lemma 10.3.7. For every Ω there exist $0 \le s(\Omega) \le \mathrm{DD}(J(f))$ and a Borel probability measure m on $J(f)$ such that

$$m(f(A)) \ge \int_A |f'|^{s(\Omega)}\,dm$$

for any special set $A \subset J(f)$ and

$$m(f(A)) = \int_A |f'|^{s(\Omega)}\,dm$$

for any special set $A \subset J(f)$ disjoint from Ω.

Proof. For every $n \geq 1$ let $V_n = B(\Omega, \frac{1}{n})$ and let m_n be the measure on $J(f)$ satisfying (10.3.7) and (10.3.8) for the neighbourhood V_n. Using Lemma 10.1.6 we shall show that any weak–* limit m of the sequence of measures $\{m_n\}_{n=1}^{\infty}$ satisfies the requirements of Lemma 10.3.7. Indeed, first observe that the sequence $\{s(V_n)\}_{n=1}^{\infty}$ is nondecreasing and denote its limit by $s(\Omega)$. Therefore the sequence of continuous functions $g_n = |f'|^{s(V_n)}$, $n = 1, 2, \ldots$, defined on $J(f)$ converges uniformly to the continuous function $g = |f'|^{s(\Omega)}$. Let A be a special subset of $J(f)$ such that

$$\overline{A} \cap (\mathrm{Sing}(f) \cup \Omega) = \emptyset. \tag{10.3.9}$$

Then one can find a compact set $\Gamma \subset J(f)$ disjoint from $\overline{A}$ and such that $\mathrm{Int}(\Gamma) \supset \mathrm{Sing}(f) \cup \Omega$. So, using also Lemma 10.3.3, we see that for any n sufficiently large, say $n \geq q$,

$$\overline{V_n} \subset \Gamma \qquad \text{and} \qquad \overline{V_n} \cap \mathrm{Crit}(f) = \emptyset. \tag{10.3.10}$$

Therefore, by (10.3.7) and (10.3.8), we conclude that Lemma 10.1.6 applies to the sequence of measures $\{m_n\}_{n=q}^{\infty}$ and the sequence of functions $\{g_n\}_{n=q}^{\infty}$. Hence, the first property required in our lemma is satisfied for any special subset of $J(f)$ disjoint from $\mathrm{Crit}(f)$ and since $A \cap \Gamma = \emptyset$, the second property is satisfied for the set A. So, since any special subset of $J(f)$ disjoint from $\mathrm{Sing}(f) \cup \Omega$ can be expressed as a disjoint union of special sets satisfying (10.3.9), an easy computation shows that the second property is satisfied for all special sets disjoint from $\mathrm{Sing}(f) \cup \Omega$. Therefore, in order to finish the proof, it is enough to show that the second requirement of the lemma is satisfied for every point of the set $\mathrm{Sing}(f)$. First note that by (10.3.10) and (10.3.8), formula (a') in Lemma 10.1.6 is satisfied for every $n \geq q$ and every $x \in \mathrm{Crit}(f) \setminus J(f)_0(f)$. As $f : J(f) \to J(f)$ is an open map, the set $J(f)_0(f)$ is empty and $\mathrm{Sing}(f) = \mathrm{Crit}(f)$. Consequently formula (d) of Lemma 10.1.6 is satisfied for any critical point $c \in J(f)$ of f. Since $g(c) = |f'(c)|^{s(\Omega)} = 0$, this formula implies that $m(f()) \leq 0$. Thus $m(\{f(c)\}) = 0 = |f'(c)|^{s(\Omega)} m(\{c\})$. The proof is finished. ♣

Lemma 10.3.8. Let m be a the me the measure constructed in Lemma 10.3.7. If for some $z \in J(f)$ the series $S(t, z) = \sum_{n=1}^{\infty} |(f^n)'(z)|^t$ diverges then $m(\{z\}) = 0$ or a positive iteration of z is a parabolic point of f. Moreover, if z itself is periodic then $m(\{f(z)\}) = |f'(z)|^t m(\{z\})$.

Proof. Suppose that $m(\{z\}) > 0$. Assume first that the point z is not eventually periodic. fhen by the definition of a conformal measure on the complement of some finite set we get $1 \geq m(\{f^n(z) : n \geq 1\}) \geq m(\{z\}) \sum_{n=1}^{\infty} |(f^n)'(z)|^t = \infty$, which is a contradiction. Hence z is eventually periodic and therefore there exist positive integers k and q such that $f^k(f^q(z)) = f^q(z)$. Since $f^q(z) \in J(f)$ and since the family of of all iterates of f on a sufficiently small neighbourhood of an attractive periodic point is normal, this implies that $|(f^k)'(f^q(z))| \geq 1$. If $|(f^k)'(f^q(z))| = \lambda > 1$ then, again by the definition of a conformal measure on the complement of some finite set, $m(\{f^q(z)\}) > 0$ and $m(\{f^{kn}(f^q(z))\}) \geq \lambda^{nt} m(\{f^q(z)\})$. Thus $m(\{f^{kn}(f^q(z))\})$ converges to ∞, which is a contradiction. Therefore $|(f^k)'(f^q(z))| = 1$ which finishes the proof of the first assertion of the lemma. In order to prove the second assertion assume that $q = 1$. Then, using the definition of conformal measures on the complement of some finite set again, we get $m(\{f(z)\}) \geq m(\{z\})|f'(z)|^t$ and on the other hand

$$m(\{z\}) = m(\{f^{k-1}(f(z))\}) \geq m(\{f(z)\})|(f^{k-1})'(f(z))|^t = m(\{f(z)\})|f'(z)|^{-t}.$$

Therefore $m(\{f(z)\}) = m(\{z\})|f'(z)|^t$. The proof is finished. ♣

Corollary 10.3.9. If for every $x \in \mathrm{Crit}(f)$ one can find $y(x) \in \overline{\{f^n(x) : n \geq 0\}}$ such that the series $S(t, y(x))$ diverges for every $0 \leq t \leq \mathrm{DD}(J(f))$, then there exists an s-conformal measure for $f : J(f) \to J(f)$ with $0 \leq s \leq \mathrm{DD}(J(f))$.

Proof. Let m be a the me the measure constructed in Lemma 10.3.7. Since $S(t, y(x))$ diverges for every $0 \leq t \leq \mathrm{DD}(J(f))$, we see that $y(x) \notin \mathrm{Crit}(f)$. If for some $x \in \mathrm{Crit}(f)$, $y(x)$ is a non-periodic point eventually falling into a parabolic point, then let $z(x)$ be this parabolic point; otherwise put $z(x) = y(x)$. The set $\Omega = \{z(x) : x \in \mathrm{Crit}(f)\}$ meets the conditions (10.3.5), (10.3.6) and is contained in $\overline{\bigcup_{n=1}^{\infty} f^n(\mathrm{Crit}(f))}$. Since for every $t \geq 0$ and $z \in J(f)$ the divergence of the series $S(t, z)$ implies the divergence of the series $S(t, f(z))$, it follows immediately from Lemma 10.3.7 and Lemma 10.3.8 that the measure m is s-conformal. ♣

Now we are in position to prove the following main result of this section.

Theorem 10.3.10. $\mathrm{HyD}(J(f)) = \mathrm{DD}(J(f)) = \delta(f)$ and there exists a $\delta(f)$–conformal measure for $f : J(f) \to J(f)$.

Proof. For every $x \in \mathrm{Crit}(f)$ the set $\overline{\{f^n(x) : n \geq 0\}}$ is closed and forward invariant under f. Therefore, in view of Theorem 2.1.8 (Bogolubov-Krylov theorem) there exists $\mu \in M_e(f)$ supported on $\overline{\{f^n(x) : n \geq 0\}}$. By Corollary A of [Przyt, Lyap] there exists at least one point $y(x) \in \overline{\{f^n(x) : n \geq 0\}}$ such that $\limsup_{n\to\infty} |(f^n)'(y(x))| \geq 1$ and consequently the series $S(t, y(x))$ diverges for every $t \geq 0$. So, in view of Corollary 10.3.9 there exists an s-conformal measure for $f : J(f) \to J(f)$ with $0 \leq s \leq \mathrm{DD}(J(f))$. Combining this with Lemma 10.2.2 and Theorem 10.2.3 complete the proof. ♣

§10.4. PESIN'S FORMULA.

In this section our aim is to prove two main theorems. The first one is as follows.

Theorem 10.4.1. (Pesin's formula) Assume that X is a compact subset of the closed complex plane $\overline{\mathbb{C}}$ and that $f \in \mathcal{A}(X)$. If m is a t- conformal measure for f and $\mu \in M_e^+(f)$ is absolutely continuous with respect to m, then $\mathrm{HD}(\mu) = t$.

Proof. In view of Lemma 10.2.2 we only need to prove that $t \leq \mathrm{HD}(\mu)$ and in order to do this we essentially combine the arguments from the proof of Lemma 10.2.2 and from the proof of formula (9.4.1). So, we work in the natural extension $(\tilde{X}, \tilde{f}, \tilde{\mu})$. Fix $0 < \varepsilon < \chi_\mu/3$

and let $\tilde{X}(\varepsilon)$ and $r(\varepsilon)$ be given by Corollary 9.2.4. In view of the Birkhoff ergodic theorem there exists a measurable set $\tilde{F}(\varepsilon) \subset \tilde{X}(\varepsilon)$ such that $\tilde{\mu}(\tilde{F}(\varepsilon)) \geq 1 - 2\varepsilon$ and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n-1} \chi_{\tilde{X}(\varepsilon)} \circ \tilde{f}^n(\tilde{x}) = \tilde{\mu}(\tilde{X}(\varepsilon))$$

for every $\tilde{x} \in \tilde{F}(\varepsilon)$. Let $F(\varepsilon) = \pi(\tilde{F}(\varepsilon))$. Then $\mu(F(\varepsilon)) = \tilde{\mu}(\pi^{-1}(F(\varepsilon)) \geq \tilde{\mu}(\tilde{F}(\varepsilon)) \geq 1-2\varepsilon$. Consider now $x \in F(\varepsilon) \cap X_o$ and take $\tilde{x} \in \tilde{F}(\varepsilon)$ such that $x = \pi(\tilde{x})$. Then by the above there exists an increasing sequence $\{n_k = n_k(x) : k \geq 1\}$ such that $\tilde{f}^{n_k}(\tilde{x}) \in \tilde{X}(\varepsilon)$ and

$$\frac{n_{k+1} - n_k}{n_k} \leq \varepsilon \tag{10.4.1}$$

for every $k \geq 1$. Moreover Corollary 9.2.4 produces holomorphic inverse branches $f_x^{-n_k} : B(f^{n_k}(x), r(\varepsilon)) \to \overline{\mathbb{C}}$ of f^{n_k} such that $f_x^{-n_k} f^{n_k}(x) = x$ and

$$f_x^{-n_k}\big(B(f^{n_k}(x), r(\varepsilon))\big) \subset B\big(x, K|(f^{n_k})'(x)|^{-1} r(\varepsilon)\big)$$

Set $r_k = r_k(x) = K^{-1}|(f^{n_k})'(x)|^{-1} r(\varepsilon)$. By Corollary 9.2.4 $r_k \leq K^{-2}\exp\big(-(\chi_\mu - \varepsilon)n_k\big)r(\varepsilon)$. So, using Corollary 9.2.4 again and (10.4.1) we can estimate

$$\begin{aligned}
r_k &= r_{k+1}|(f^{n_{k+1}-n_k})'(f^{n_k}(x))| \leq r_{k+1} K \exp\big(\chi_\mu + \varepsilon)(n_{k+1} - n_k)\big) \\
&\leq r_{k+1} K \exp\big(\chi_\mu + \varepsilon)n_{k+1}\varepsilon\big) \leq K r_{k+1} \exp\big(\chi_\mu - \varepsilon)2n_{k+1}\varepsilon\big) \leq r_{k+1} K (K^{-2} r(\varepsilon) r_{k+1}^{-1})^{2\varepsilon} \\
&= K^{1-4\varepsilon} r(\varepsilon)^{2\varepsilon} r_{k+1}^{1-2\varepsilon}
\end{aligned}$$

Take now any $0 < r \leq r_1$ and find $k \geq 1$ such that $r_{k+1} < r \leq r_k$. Then using this estimate, t-conformality of m, and invoking Corollary 9.2.4 once more we get

$$\begin{aligned}
m(B(x,r)) \leq m(B(x,r_k)) &\leq K^t |(f^{n_k})'(x)|^{-t} m(B(x, r(\varepsilon))) \\
&\leq K^{2t} r(\varepsilon)^{-t} r_k^t \\
&\leq K^{(3-4\varepsilon)t} r(\varepsilon)^{2\varepsilon t} r^{(1-2\varepsilon)t}
\end{aligned}$$

So, by Theorem 5.5.1 (Besicovitch covering theorem) $\mathrm{H}_{(1-2\varepsilon)t}(X) \geq \mathrm{H}_{(1-2\varepsilon)t}(F(\varepsilon)) > 0$, whence $\mathrm{HD}(X) \geq (1-2\varepsilon)t$. Letting $\varepsilon \to 0$ completes the proof. ♣