MINICOURSE ON HAUSDORFF DIMENSION

Connection with Classical, Relativized and Conditional Variational Principles in Ergodic Theory

nuno luzia

30 may 2007

Lecture 1: 1-dimensional dynamics

- Definition of Hausdorff dimension.
- How to compute Hausdorff dimension; Volume Lemma; Example: self-affine Cantor sets; Moran formula.
- Nonlinear Cantor sets; Variational Principle for the Topological Pressure; Bowen's equation; Measure of maximal dimension.

References

[Bo1] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms (Lecture Notes in Mathematics, 470)*. Springer, 1975.

[Bo2] R. Bowen. Hausdorff dimension of quasi-circles. *Publ. Math. I.H.E.S.* **50** (1979), 11-26.

[F] K. Falconer, *Fractal geometry, Mathematical foundations and applications*. Wiley, 1990.

[M] R. Mañé. *Ergodic theory and differentiable dynamics*. Springer, 1987.

[PT] J. Palis and F. Takens. *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations*. Cambridge University Press, 1993.

[P] Ya. Pesin. *Dimension Theory in Dynamical Systems: Contemporary Views and Applications*. Chicago University Press, 1997.

[R1] D. Ruelle. Repellers for real analytic maps. *Ergod. Th. and Dynam. Sys.* **2** (1982), 99-107.

Lecture 2: 2-dimensional dynamics

- Hausdorff dimension of general Sierpinski carpets and extension to skew-product transformations;
- Connection with the relativized variational principle; Gauge functions and relativized Gibbs measures.

References

[Be] T. Bedford. Crinkly curves, Markov partitions and box dimension of self similar sets. *PhD Thesis*, University of Warwick, 1984.

[DG] M. Denker and M. Gordin. Gibbs measures for fibred systems. *Adv. Math.* **148** (1999), 161-192.

[DGH] M. Denker, M. Gordin and S. Heinemann. On the relative variational principle for fibred expanding maps. *Ergod. Th. & Dynam. Sys.* **22** (2002), 757-782.

[GL] D. Gatzouras and P. Lalley. Hausdorff and box dimensions of certain self-affine fractals. *Indiana Univ. Math. J.* **41** (1992), 533-568.

[GP1] D. Gatzouras and Y. Peres. Invariant measures of full dimension for some expanding maps. *Ergod. Th. and Dynam. Sys.* **17** (1997), 147-167.

[HL] I. Heuter and S. Lalley. Falconer's formula for the Hausdorff dimension of a self-affine set in $\mathbb{R}^2$. *Ergod. Th. & Dynam. Sys.* **15** (1997), 77-97.

[L1] N. Luzia. A variational principle for the dimension for a class of non-conformal repellers. *Ergod. Th. & Dynam. Sys.* **26** (2006), 821-845.

[L2] N. Luzia. Hausdorff dimension for an open class of repellers in $\mathbb{R}^2$. *Nonlinearity* **19** (2006), 2895-2908.

[LW] F. Ledrappier and P. Walters. A relativised variational principle for continuous transformations. *J. London Math. Soc.* **16** (1977), 568-576.

[Mu] C. McMullen. The Hausdorff dimension of general Sierpiński carpets. *Nagoya Math. J.* **96** (1984), 1-9.

Lecture 3: Measure of full dimension for nonlinear 2-dimensional dynamics

- Topological pressure and variational principle for some noncompact sets.
- Existence of an ergodic measure of full dimension for skew-product transformations: reductions to the previous types of variational principles.

References

[Bo3] R. Bowen. Topological entropy for non-compact sets. *Trans. Amer. Math. Soc.* **184** (1973), 125-136.

[GP2] D. Gatzouras and Y. Peres. The variational principle for Hausdorff dimension: a survey. *Ergodic theory of Z^d actions (Warwick, 1993–1994)*, 113–125, London Math. Soc. Lecture Note Ser., 228, *Cambridge Univ. Press, Cambridge*, 1996.

[L3] N. Luzia. Measure of full dimension for some nonconformal repellers. Preprint, 2006. (arXiv:0705.3604v1)

[PP] Ya. Pesin and B. Pitskel. Topological pressure and the variational principle for noncompact sets. *Funktsional. Anal. i Prilozhen.* **18** (1984), 50-63.

[R2] D. Ruelle. *Thermodynamic Formalism (Encyclopedia of Mathematics and its Applications, 5).* Addison-Wesley, 1978.

Lecture 1

We begin by defining the Hausdorff dimension of a set $F \subset \mathbb{R}^n$ (see the book [F] for more details). This is done by defining *t-Hausdorff measures*, where $t \geq 0$, and then choose the appropriate t for measuring F. When $t = n$, the n-Hausdorff measure is just the n-dimensional exterior Lebesgue measure. The diameter of a set $U \subset \mathbb{R}^n$ is denoted by $|U|$. We say that the countable collection of sets $\{U_i\}$ is a δ-cover of F if $F \subset \bigcup_{i=1}^{\infty} U_i$ and $|U_i| \leq \delta$ for each i. Given $t \geq 0$ and $\delta > 0$ let

$$\mathcal{H}^t_\delta(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^t : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\}.$$

Then the *t-Hausdorff measure* of F is given by

$$\mathcal{H}^t(F) = \lim_{\delta \to \infty} \mathcal{H}^t_\delta(F)$$

(this limit exists because $\mathcal{H}^t_\delta(F)$ is increasing in δ). So, given $F \subset \mathbb{R}^n$, what is the appropriate t-Hausdorff measure for measuring F? Let us look at the function $t \mapsto \mathcal{H}^t(F)$. It is not difficult to see that there is a critical value t_0 such that

$$\mathcal{H}^t(F) = \begin{cases} \infty & \text{if } t < t_0 \\ 0 & \text{if } t > t_0. \end{cases}$$

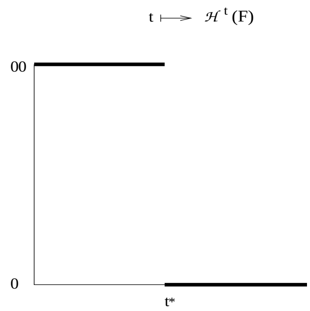

Figure 1

Definition 1. The *Hausdorff dimension* of F is

$$\dim_{\mathrm{H}} F = t_0.$$

Remark 1. In the definition of Hausdorff dimension we can use covers by *balls*. More precisely, let $\mathcal{B}^t_\delta(F)$ be the number obtained using covers by balls of diameters $\leq \delta$ instead of using any δ-cover when defining $\mathcal{H}^t_\delta(F)$. Of course, $\mathcal{H}^t_\delta(F) \leq \mathcal{B}^t_\delta(F)$ because we are restricting to a particular class of δ-covers and using inf over this class. Now, given a set U with $|U| \leq \delta$ there exists a ball $B \supset U$ with $|B| \leq 2|U|$. This implies that $\mathcal{B}^t_{2\delta}(F) \leq 2^t \mathcal{H}^t_\delta(F)$. So, letting $\delta \to 0$ we obtain

$$\mathcal{H}^t(F) \leq \mathcal{B}^t(F) \leq 2^t \mathcal{H}^t(F),$$

and so the critical value t_0 at which $\mathcal{H}^t(F)$ and $\mathcal{B}^t(F)$ jump is the same.

Using the same arguments, we can also use covers by squares. What we cannot do is to use only covers formed by sets which are "too distorted". This is why the computation of Hausdorff dimension of invariant sets for "nonconformal dynamics" reveals to be more complicated.

Problem: How to compute the Hausdorff dimension of a set?

- If, for every $\delta > 0$, we find *some* δ-cover $\{U_i\}$ of F such that

$$\sum_i |U_i| \leq C < \infty,$$

where C is a constant independent of δ, then $\mathcal{H}^t(F) < \infty$ which implies

$$\dim_{\mathrm{H}} F \leq t.$$

- How to obtain an estimation

$$\dim_{\mathrm{H}} F \geq t?$$

We must use *every* δ-cover... *A priori* this seems to be an untreatable problem. Is there some non-trivial mathematical object that takes acquaintance with this? As we shall see now, the answer is yes, and the object is *measure*.

Mass Distribution Principle

Suppose there exists a Borel probability measure μ on $\mathbb{R}^n$ such that $\mu(F) = 1$, and there exist a constant $C > 0$ and $t > 0$ such that for every set $U \subset \mathbb{R}^n$,

$$\mu(U) \leq C|U|^t.$$

Then, if $\{U_i\}$ is any δ-cover of F,

$$\sum_i |U_i|^t \geq C^{-1} \sum_i \mu(U_i) \geq C^{-1}\mu(\sum_i U_i) \geq C^{-1}\mu(F) = C^{-1}.$$

This implies that $\mathcal{H}^t(F) \geq C^{-1} > 0$ and so

$$\dim_{\mathrm{H}} F \geq t.$$

There is a "non-uniform" version of this principle as we shall describe now. The main concept behind this is the *Hausdorff dimension of a probability measure* μ defined by L.-S. Young as

$$\dim_{\mathrm{H}} \mu = \inf\{\dim_{\mathrm{H}} F : \mu(F) = 1\}.$$

By definition, if $\mu(F) = 1$ then $\dim_{\mathrm{H}} F \geq \dim_{\mathrm{H}} \mu$. The "non-uniform" version of the Mass Distribution Principle deals with the following limit

$$\underline{d}_\mu(x) = \varliminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r},$$

where $B(x,r)$ stands for the open ball of radius r centered at the point x, which is called the *lower pointwise dimension of* μ. Then we have the following (see the book [P] for a proof)

Lemma 1. *(*Volume Lemma*)*

(1) *If* $\underline{d}_\mu(x) \geq t$ *for* μ*-a.e.* x *then* $\dim_{\mathrm{H}} \mu \geq t$.
(2) *If* $\underline{d}_\mu(x) \leq t$ *for* μ*-a.e.* x *then* $\dim_{\mathrm{H}} \mu \leq t$.
(3) *If* $\underline{d}_\mu(x) = t$ *for* μ*-a.e.* x *then* $\dim_{\mathrm{H}} \mu = t$.

Corollary 1. *If* $\mu(F) = 1$ *and* $\underline{d}_\mu(x) \geq t$ *for* μ*-a.e.* x *then* $\dim_{\mathrm{H}} F \geq t$.

Example 1. (Self-affine Cantor sets)
These are self-affine generalizations of the famous "middle-third" Cantor set. They are constructed as limit sets of *n-approximations*: the 1-approximation consists in m disjoint subintervals of $[0, 1]$; the 2-approximation consists in substituting each interval of the 1-approximation by a rescaled self-affine copy of the 1-approximation (from [0,1] to this interval), and so on. In the limit we get a Cantor set K.

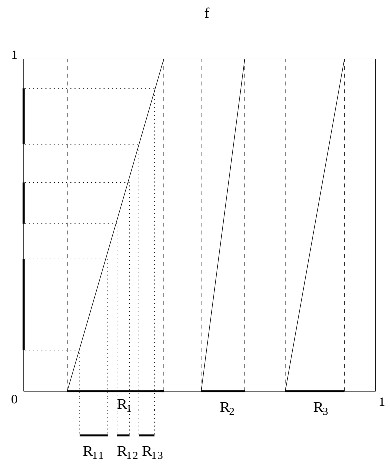

FIGURE 2

Until now we have not talked about "dynamical systems". These sets are dynamically defined (see Figure 2): Let $\{R_i\}_{i=1}^m$ be a collection of disjoint subintervals of $[0,1]$, which we call *Markov partition*. Consider the map $f\colon \bigcup_{i=1}^m R_i \to [0,1]$ that sends linearly each interval R_i onto $[0,1]$. Then K is the set of points that always remain in the Markov partition when iterated by f (the n-approximation is the set of points that remain in the Markov partition when iterated $n-1$ times). More precisely,

$$K = \bigcap_{n=0}^{\infty} \bigcup_{i_1,\dots,i_n} R_{i_1\dots i_n} \tag{1}$$

where $i_1, ..., i_n \in \{1, ...m\}$ and

$$R_{i_1\dots i_n} = R_{i_1} \cap f^{-1}(R_{i_2}) \cap \cdots \cap f^{-n+1}(R_{i_n}) \tag{2}$$

which we call *basic intervals of order n*.

Now we are going to compute the Hausdorff dimension of K in terms of the numbers $a_i = |R_i|, i = 1, .., m$, the lengths of the elements of the Markov partition.

Estimate $\leq$. Let $\delta > 0$ and n be such that $(\max a_i)^n \leq \delta$. So the basic sets of order n form a δ-cover of K. Note that

$$|R_{i_1\dots i_n}| = a_{i_1} a_{i_2} ... a_{i_n}.$$

So

$$\mathcal{H}^t_\delta(K) \leq \sum_{i_1=1}^m \sum_{i_2=1}^m \cdots \sum_{i_n=1}^m (a_{i_1} a_{i_2} ... a_{i_n})^t = \Big(\sum_{i=1}^m a_i^t\Big)^n. \tag{3}$$

What is the least t such that the expression on the right hand side of (3) is finite for every n? It is the unique solution $t = t_0$ of the following equation

$$\sum_{i=1}^m a_i^t = 1 \tag{4}$$

which is called *Moran formula* (note that $\sum_{i=1}^m a_i^t$ is strictly decreasing in t, has value $m > 1$ for $t = 0$ and value $\sum_{i=1}^m a_i < 1$ for $t = 1$). So $\mathcal{H}^{t_0}_\delta(K) \leq 1$ for every $\delta > 0$ which implies

$$\dim_{\mathrm{H}} K \leq t_0.$$

Estimate $\dim_{\mathrm{H}} K \geq t_0$. Let μ be the Bernoulli measure on K with weigths

$$\mu(R_i) = a_i^{t_0},\ i = 1, .., m$$

i.e.

$$\mu(R_{i_1 \dots i_n}) = \prod_{l=1}^{n} a_{i_l}^{t_0}$$

(note that $f|K$ is topologically conjugated to a full shift on m symbols; also remember that $\sum_{i=1}^{m} a_i^{t_0} = 1$). Then

$$\frac{\log \mu(R_{i_1 \dots i_n})}{\log |R_{i_1 \dots i_n}|} = t_0$$

and it follows from the Volume Lemma that $\dim_{\mathrm{H}} \mu = t_0$ and so

$$\dim_{\mathrm{H}} K \geq \dim_{\mathrm{H}} \mu = t_0.$$

Conclusion.

$$\dim_{\mathrm{H}} K = t_0$$

where t_0 is given by the Moran formula

$$\sum_{i=1}^{m} a_i^{t_0} = 1.$$

Example 2. If K is the middle-third Cantor set then its Hausdorff dimension is the solution of

$$2\left(\frac{1}{3}\right)^t = 1$$

i.e.

$$\dim_{\mathrm{H}} K = \frac{\log 2}{\log 3}.$$

This equation has a dynamical meaning of the form

$$\text{dimension} = \frac{\text{entropy}}{\text{Lyapunov exponent}}. \tag{5}$$

In general, the Lyapunov exponent of a 1-dimensional dynamics depends on invariant measures. For instance, is there some analogous formula for the Hausdorff dimension of self-affine Cantor sets? The answer is yes as we shall see now in a more general context. Also, a formula of the type (5) only makes sense for 1-dimensional dynamics, since for n-dimensional dynamics it is expected to exist several Lyapunov exponents.

Nonlinear Cantor sets

Now we consider sets Λ which are generated using a dynamics as for the self-affine Cantor sets, but now the generating dynamics need not be linear. Let $\{R_i\}_{i=1}^m$ be a collection of disjoint subintervals of $[0,1]$, and $f\colon \bigcup_{i=1}^m R_i \to [0,1]$ be a $\mathrm{C}^{1+\delta}$ transformation (for some $\delta > 0$) such that $f' \geq \sigma > 1$ and $f|R_i$ is a homeomorphism onto $[0,1]$ (see Figure 3).

As before we consider the f-invariant set Λ defined by (1) and (2). It follows from the intermediate value theorem that the basic intervals of order n satisfy

$$|R_{i_1 \dots i_n}| = ((f^n)'(x))^{-1} = \left(\prod_{j=0}^{n-1} f'(f^j x)\right)^{-1}$$

for some $x \in R_{i_1 \dots i_n}$. Using the fact that $f' \geq \sigma > 1$ and $f \in \mathrm{C}^{1+\delta}$ we obtain the following result (see the book [PT] for a proof)

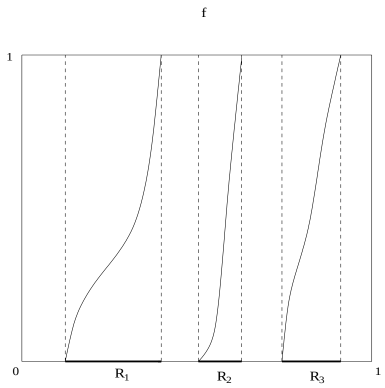

FIGURE 3

Lemma 2. *(*BOUNDED DISTORTION PROPERTY*)*
There exists $C > 0$ such that for every $n \in \mathbb{N}$ and every n-tuple $(i_1, ..., i_n)$,

$$\frac{(f^n)'(x)}{(f^n)'(y)} \leq C$$

for every $x, y \in R_{i_1 \dots i_n}$.

If $\varphi\colon \Lambda \to \mathbb{R}$, we use the following notation

$$(S_n \varphi)(x) = \sum_{j=0}^{n-1} \varphi(f^j(x)).$$

Then we can write

$$|R_{i_1 \dots i_n}|^t \sim \sup_{x \in R_{i_1 \dots i_n}} e^{-t(S_n \log f')(x)},$$

where $\sim$ means "up to a constant (independent of n)", due to the bounded distortion property. As for the self-affine case, for proving the estimate $\leq$ we use the cover of Λ by basic intervals of order n, so we want to find the "least" t such that

$$\sum_{i_1, \dots, i_n} |R_{i_1 \dots i_n}|^t \sim \sum_{i_1, \dots, i_n} \sup_{x \in R_{i_1 \dots i_n}} e^{-t(S_n \log f')(x)} \tag{6}$$

is finite for all n. Due to the uniform expansion of f, (6) grows at an exponential rate in n: negative, zero or positive, depending on the value of t. So we want to find the value $t = t_0$ for which

$$\overline{\lim_{n \to \infty}} \frac{1}{n} \log \sum_{i_1, \dots, i_n} \sup_{x \in R_{i_1 \dots i_n}} e^{-t_0(S_n \log f')(x)} = 0 \tag{7}$$

(for then, the exponential rate at $t = t_0 + \epsilon$, $\epsilon > 0$, is negative). This motivates the following definition.

Definition 2. Let $\varphi\colon \Lambda \to \mathbb{R}$ be a Hölder continuous function. The *Topological Pressure of* φ (with respect to the dynamics $f|\Lambda$) is

$$P_{f|\Lambda}(\varphi) = \overline{\lim_{n \to \infty}} \frac{1}{n} \log \sum_{i_1, \dots, i_n} \sup_{x \in R_{i_1 \dots i_n}} e^{S_n \varphi(x)}.$$

The condition of φ being Hölder continuous is to apply the bounded distortion property.

Remark 2. The *Topological Entropy* is the Topological Pressure of the constant zero function:

$$h_{top}(f|\Lambda) = P_{f|\Lambda}(0).$$

Since in our case we are dealing with a dynamics which is topologically congugated to a full shift in m symbols, we have that

$$h_{top}(f|\Lambda) = \log m.$$

So, equation (7) can be restated as: There is a unique solution $t = t_0$ of the following equation

$$P_{f|\Lambda}(-t \log f') = 0, \tag{8}$$

which is the celebrated *Bowen's equation*. So

$$\dim_{\mathrm{H}} \Lambda \leq t_0.$$

To prove the other inequality we shall need the well known *Variational Principle for the Topological Pressure*. Before stating this principle we must introduce the notion of *entropy of an invariant measure*. Denote by $\mathcal{M}(f|\Lambda)$ the set of all $f|\Lambda$-invariant probability measures, and by $\mathcal{M}_e(f|\Lambda)$ the subset of ergodic ones. Given $\mu \in \mathcal{M}(f|\Lambda)$, Shannon-McMillan-Breiman's theorem (see [M]) says that the limit

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu(R_{i_1 \ldots i_n})$$

exists for μ-a.e. $x \equiv \bigcap_{n=1}^{\infty} R_{i_1\ldots i_n}$. Moreover, if μ is ergodic then this limit is constant μ-a.e.

Definition 3. Let $\mu \in \mathcal{M}_e(f|\Lambda)$. The *Entropy of* μ (with respect to the dynamics $f|\Lambda$) is

$$h_\mu(f) = \lim_{n\to\infty} -\frac{1}{n} \log \mu(R_{i_1\ldots i_n}) \tag{9}$$

for μ-a.e. $x \equiv \bigcap_{n=1}^{\infty} R_{i_1\ldots i_n}$. If $\mu \in \mathcal{M}(f|\Lambda)$ then $h_\mu(f)$ is just the integral of the limit in (9) with respect to μ.

Note that, if $\mu \in \mathcal{M}_e(f|\Lambda)$ then by Shannon-McMillan-Breiman's theorem, Birkhoff's ergodic theorem and the bounded distortion property,

$$\begin{aligned}
\underline{d}_\mu(x) &= \lim_{n\to\infty} \frac{\log \mu(R_{i_1\ldots i_n})}{\log |R_{i_1\ldots i_n}|} \\
&= \lim_{n\to\infty} \frac{-\frac{1}{n}\log \mu(R_{i_1\ldots i_n})}{\frac{1}{n}(S_n \log f')(x)} \\
&= \frac{h_\mu(f)}{\int \log f' \, d\mu}
\end{aligned}$$

for μ-a.e. $x \equiv \bigcap_{n=1}^{\infty} R_{i_1\ldots i_n}$. So, by the Volume Lemma, if $\mu \in \mathcal{M}_e(f|\Lambda)$ then

$$\dim_{\mathrm{H}} \mu = \frac{h_\mu(f)}{\int \log f' \, d\mu}. \tag{10}$$

The next theorem is due to Sinai-Ruelle-Bowen and a proof can be seen in the excelent exposition [Bo1].

Theorem 1. (Variational Principle for the Topological Pressure)
Let $\varphi\colon \Lambda \to \mathbb{R}$ be Hölder continuous. Then

$$P_{f|\Lambda}(\varphi) = \sup_{\mu \in \mathcal{M}(f|\Lambda)} \left\{ h_\mu(f) + \int \varphi \, d\mu \right\}.$$

Moreover, this supremum is attained at a unique invariant measure, which is the "Gibbs state for the potential φ" (hence ergodic) and is denoted by μ_φ.

When $\varphi = 0$ we obtain the *Variational Principle for the Entropy.*

At this time the reader might guess that, for calculating Hausdorff dimension, we will use the potential $\varphi = -t_0 \log f'$ where t_0 is the solution of Bowen's equation (8). It follows from Theorem 1 and (10) that

$$0 = P_{f|\Lambda}(-t_0 \log f') = h_{\mu_{-t_0 \log f'}}(f) - t_0 \int \log f' \, d\mu_{-t_0 \log f'}$$

and

$$t_0 = \frac{h_{\mu_{-t_0 \log f'}}(f)}{\int \log f' \, d\mu_{-t_0 \log f'}} = \dim_{\mathrm{H}} \mu_{-t_0 \log f'}.$$

Thus

$$\dim_{\mathrm{H}} \Lambda \geq \dim_{\mathrm{H}} \mu_{-t_0 \log f'} = t_0$$

as we wish.

Conclusion: The Hausdorff dimension of Λ is given by the root t_0 of Bowen's equation

$$P_{f|\Lambda}(-t \log f') = 0$$

(for a complete discussion see [Bo2] and [R1]). Moreover, the Gibbs state for the potential $-t_0 \log f'$ denoted by

$$\mu_{-t_0 \log f'}$$

is the unique ergodic invariant measure on Λ of full dimension. Also, it is very important to keep in mind, in what comes, the validity of the "Variational Principle for the Dimension" i.e.

$$\dim_{\mathrm{H}} \Lambda = \sup_{\mu \in \mathcal{M}_e(f|\Lambda)} \dim_{\mathrm{H}} \mu, \tag{11}$$

and

$$\mu \in \mathcal{M}_e(f|\Lambda) \Rightarrow \dim_{\mathrm{H}} \mu = \frac{h_\mu(f)}{\int \log f' \, d\mu}. \tag{12}$$

Lecture 2

Here we deal with sets Λ which are generated by smooth plane transformations f which are nonconformal i.e. possess two different rates of expansion. As before we also consider sets which are generated using a Markov partition, and so the *basic sets of order* n (see (2)) form a natural cover of Λ by sets with diameter arbitrarily small. The problem is that, due to nonconformality, these sets are not "approximately" balls, and so covers formed by basic sets are not sufficiently for calculating Hausdorff dimension (see Remark 1).

Nevertheless, Bowen's equation

$$P_{f|\Lambda}(-t\log\|Df\|)=0$$

still makes sense but, in general, its root does not give us the Hausdorff dimension of Λ: it is greater than $\dim_{\mathrm{H}}\Lambda$ and depends on the dynamics f we use to generate Λ. The aim of this lecture is to convince the reader that the main tool for computing Hausdorff dimension in the nonconformal setting is not the Topological Pressure but the *Relativized Topological Pressure* and, especially, the corresponding *Relativized Variational Principle.*

Example 3. (General Sierpinski Carpets)
Let $\mathbb{T}^2=\mathbb{R}^2/\mathbb{Z}^2$ be the 2-torus and $f_0\colon\mathbb{T}^2\to\mathbb{T}^2$ be given by

$$f_0(x,y)=(lx,my)$$

where $l>m>1$ are integers. The grid of lines $[0,1]\times\{i/m\}$, $i=0,...,m-1$, and $\{j/l\}\times[0,1]$, $j=0,...,l-1$, form a set of rectangles each of which is mapped by f_0 onto the entire torus (these rectangles are the domains of invertibility of f_0). Now choose some of these rectangles, the Markov partition, and consider the fractal set Λ_0 consisting of those points that always remain in these chosen rectangles when iterating f_0. As before, Λ_0 is the limit (in the Hausdorff metric) of *n-approximations*: the 1-approximation consists of the chosen rectangles, the 2-approximation consists in dividing each rectangle of the 1-aproximation into $l\times m$ subrectangles and selecting those with the same pattern as in the begining, and so on (see Figure 4).

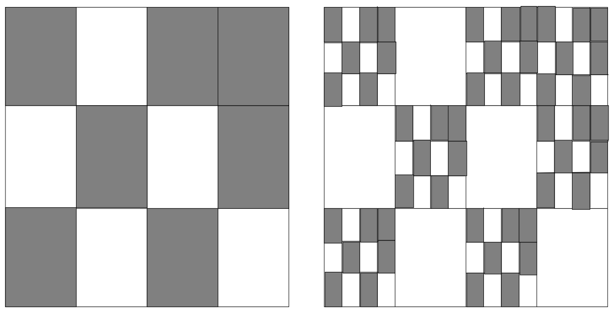

Figure 4. $l=4$, $m=3$; 1-approximation and 2-approximation

We say that (f_0,Λ_0) is a *general Sierpinski carpet.* The Hausdorff dimension of general Sierpinski carpets was computed by Bedford [Be] and McMullen [Mu], independently, obtaining the beautiful formula

$$\dim_{\mathrm{H}}\Lambda_0=\frac{\log(\sum_{i=1}^{m}n_i^{\alpha})}{\log m}$$

where n_i is the number of elements of the Markov partition in the horizontal strip i, and

$$\alpha = \frac{\log m}{\log l}.$$

See how it generalizes the 1-dimensional formula (5) it two ways: (i) put $m = 1$; (ii) put $n_i = 1$ whenever $n_i \neq 0$, $i = 1, ..., m$.

Self-affine fractals in the plane

Now we consider sets which are self-affine generalizations of general Sierpinski carpets (and generalize the corresponding 1-dimensional versions, the self-affine Cantor sets).

Let $S_1, S_2, ..., S_r$ be contractions of $\mathbb{R}^2$. Then there is a unique nonempty compact set Λ of $\mathbb{R}^2$ such that

$$\Lambda = \bigcup_{i=0}^{r} S_i(\Lambda).$$

This set is constructed like the general Sierpinski carpets (there, the contractions are the inverse branches of f_0 corresponding to the chosen rectangles, and the equation above simply means that the set is f_0-invariant). We will refer to Λ as the limit set of the semigroup generated by $S_1, S_2, ..., S_r$.

We shall consider the class of self-affine sets Λ that are the limit sets of the semigroup generated by the mappings A_{ij} given by

$$A_{ij} = \begin{pmatrix} a_{ij} & 0 \\ 0 & b_i \end{pmatrix} x + \begin{pmatrix} c_{ij} \\ d_i \end{pmatrix}, \quad (i,j) \in \mathcal{I}.$$

Here $\mathcal{I} = \{(i,j) : 1 \le i \le m \text{ and } 1 \le j \le n_i\}$ is a finite index set. We assume

$$0 < a_{ij} < b_i < 1, \tag{13}$$

for each pair (i,j), $\sum_{i=1}^{m} b_i \le 1$, and $\sum_{j=1}^{n_i} a_{ij} \le 1$ for each i. Also, $0 \le d_1 < d_2 < ... < d_m < 1$ with $d_{i+1} - d_i \ge b_i$ and $1 - d_m \ge b_m$ and, for each i, $0 \le c_{i1} < c_{i2} < ... < c_{in_i} < 1$ with $c_{i(j+1)} - c_{ij} \ge a_{ij}$ and $1 - c_{in_i} \ge a_{in_i}$. These hypotheses guarantee that the rectangles

$$R_{ij} = A_{ij}([0,1] \times [0,1])$$

have interiors that are pairwise disjoint, with edges parallel to the x- and y-axes, are arranged in "rows" of height b_i, and have height $b_i >$ width a_{ij} (see Figure 5).

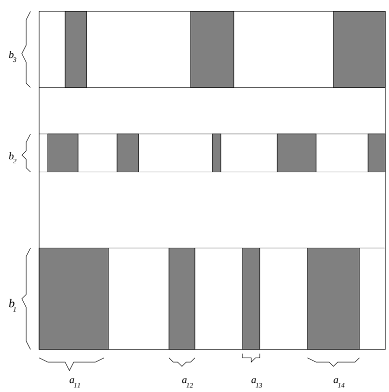

Figure 5

Theorem 2.
$$\dim_{\mathrm{H}} \Lambda = \sup_{\mathbf{p}} \left\{ \frac{\sum_{i=1}^{m} p_i \log p_i}{\sum_{i=1}^{m} p_i \log b_i} + t(\mathbf{p}) \right\} \tag{14}$$
where $\mathbf{p} = (p_1, p_2, ..., p_m)$ *is a probability vector, and* $t(\mathbf{p})$ *is the unique real in* $[0, 1]$ *satisfying*
$$\sum_{i=1}^{m} p_i \log \left(\sum_{j=1}^{n_i} a_{ij}^{t(\mathbf{p})} \right) = 0. \tag{15}$$

This theorem was proved in [GL] with a slightly different formula than (14), and later in [L1] using different methods aiming the *Relativized Variational Principle*. Here we give a sketch of proof of this theorem following [L1].

Before that, we make some remarks concerning formula (14). See how it generalizes the computation of Hausdorff dimension for self-affine Cantor sets in two ways: put $m = 1$; put $n_i = 1$ for $i = 1, ..., m$. This formula looks like the variational principle for dimension (see (11)). In fact, we shall see that, for each $\mathbf{p}$, there exists a Bernoulli measure $\mu_{\mathbf{p}}$ on Λ such that
$$\dim_{\mathrm{H}} \mu_{\mathbf{p}} = \frac{\sum_{i=1}^{m} p_i \log p_i}{\sum_{i=1}^{m} p_i \log b_i} + t(\mathbf{p}), \tag{16}$$
so it follows from formula (14) the validity of the variational principle for dimension. Also, since the functions
$$\mathbf{p} \mapsto \frac{\sum_{i=1}^{m} p_i \log p_i}{\sum_{i=1}^{m} p_i \log b_i} \quad \text{and} \quad \mathbf{p} \mapsto t(\mathbf{p})$$
are continuous (by convention $0 \log 0 = 0$; the continuity of $t(\mathbf{p})$ follows easily from the implicit function theorem), we obtain the following

Corollary 2. *There exists* $\mathbf{p}^*$ *such that*
$$\dim_{\mathrm{H}} \Lambda = \dim_{\mathrm{H}} \mu_{\mathbf{p}^*}.$$

A good way of looking at Λ is as a *(fibred) random 1-dimensional self-affine Cantor set*: If we had only one horizontal strip, then Λ would be exactly a 1-dimensional self-affine Cantor set characterized by the numbers $a_{1j}, j = 1, ..., n_1$ (the widths of the elements of the Markov partition), and so its Hausdorff dimension is given by the Moran formula (4). When there are several horizontal strips, say m, it is like we are working at the same time with m 1-dimensional self-affine transformations. When do we work with a particular 1-dimensional self-affine transformation and when do we work with another? We must look at the dynamics in the vertical axis. Given $y \in \pi(\Lambda)$ where $\pi(\Lambda)$ is the Cantor set obtained by projecting Λ onto the vertical axis (i.e the self-affine Cantor set characterized by the numbers $b_1, ..., b_m$), we iterate y using the vertical dynamics and see how the orbit of y *distributes* along the different horizontal strips, i.e. until time n we calculate the proportion of time the orbit of y stays in each horizontal strip. If the orbit of y distributes along the horizontal strips according to the distrubution $\mathbf{p} = (p_1, ..., p_m)$, then the 1-dimensional Cantor set
$$\Lambda_y = \Lambda \cap \{(x, y) : x \in [0, 1]\} \tag{17}$$
has Hausdorff dimension $t(\mathbf{p})$ given by the *random Moran formula* (15). Now the the subset of points $y \in \pi(\Lambda)$ that *distribute* according to the distribution $\mathbf{p}$ has Hausdorff dimension
$$\dim \nu_{\mathbf{p}} = \frac{\sum_{i=1}^{m} p_i \log p_i}{\sum_{i=1}^{m} p_i \log b_i},$$

where $\nu_{\mathbf{p}}$ is the Bernoulli measure on $\pi(\Lambda)$ which gives weigth p_i to the horizontal strip i. Finally, formula (14) says that "dimensions add".

Sketch of proof of Theorem 2. Let

$$d(\mathbf{p}) = \frac{\sum_{i=1}^m p_i \log p_i}{\sum_{i=1}^m p_i \log b_i} + t(\mathbf{p}).$$

As usual, the proof divides in two parts.

Part 1: $\dim_{\mathrm{H}} \Lambda \geq \sup_{\mathbf{p}} d(\mathbf{p})$

For this part we use the Volume Lemma. As mentioned before, we construct Bernoulli measures $\mu_{\mathbf{p}}$ on Λ such that

$$\dim_{\mathrm{H}} \mu_{\mathbf{p}} = d(\mathbf{p}).$$

This gives what we want because $\dim_{\mathrm{H}} \Lambda \geq \dim_{\mathrm{H}} \mu_{\mathbf{p}}$. So, let $\mu_{\mathbf{p}}$ be the Bernoulli on Λ that gives weigth

$$p_{ij} = p_i \frac{a_{ij}^{t(\mathbf{p})}}{\sum_{k=1}^{n_i} a_{ik}^{t(\mathbf{p})}}$$

to the basic rectangle R_{ij}, for $(i,j) \in \mathcal{I}$. Note that our dynamics is topologically conjugated to a Bernoulli shift with $\#\mathcal{I}$ symbols. Denote by

$$R_{(i_1 j_1)\dots(i_n j_n)}, \quad (i_l, j_l) \in \mathcal{I},\, l = 1, \dots, n$$

the basic sets of order n so that

$$\Lambda = \bigcap_{n=1}^{\infty} \bigcup_{(i_1,j_1),\dots,(i_n,j_n)\in\mathcal{I}} R_{(i_1 j_1)\dots(i_n j_n)}.$$

Then

$$\mu_{\mathbf{p}}(R_{(i_1 j_1)\dots(i_n j_n)}) = \prod_{l=1}^{n} p_{i_l j_l}.$$

Problem: We cannot use directly the sets $R_{(i_1 j_1)\dots(i_n j_n)}$ for calculating the lower pointwise dimension of $\mu_{\mathbf{p}}$, because these sets are not "approximately" balls.

In fact, $R_{(i_1 j_1)\dots(i_n j_n)}$ is a rectangle with height $\prod_{l=1}^n b_{i_l}$ and width $\prod_{l=1}^n a_{i_l j_l}$ which decay exponentially to zero when n goes to ∞ with different rates (see (13)).

Solution: For each point in Λ, construct a set containning this point by making a union of several basic sets of order n, and which is "approximately" a ball in Λ (see Figure 6).

We use the following notation: given $z \in \Lambda$ we write $z \equiv (i_1 j_1, i_2 j_2, \dots)$ iff

$$z \in \bigcap_{n=1}^{\infty} R_{(i_1 j_1)\dots(i_n j_n)}.$$

Given $z \equiv (i_1 j_1, i_2 j_2, \dots)$, let

$$L_n(z) = \max\left\{ k \geq 1 : \prod_{l=1}^{n} b_{i_l} \leq \prod_{l=1}^{k} a_{i_l j_l} \right\} \tag{18}$$

and the *approximate square*

$$B_n(z) = \{z' \in \Lambda : i'_l = i_l,\, l = 1, \dots, n \text{ and } j'_l = j_l,\, l = 1, \dots, L_n(z)\}.$$

Note that $B_n(z)$ is a union of basic rectangles of order n and is the the intersection of Λ with a rectangle with height $\prod_{l=1}^{n} b_{i_l}$ and width $\prod_{l=1}^{L_n(\omega)} a_{i_l j_l}$ satisfying, by (18),

$$1 \le \frac{\prod_{l=1}^{L_n(z)} a_{i_l j_l}}{\prod_{l=1}^{n} b_{i_l}} \le \max a_{ij}^{-1}, \tag{19}$$

hence the term "approximate square".

Figure 6. Approximate square

Then we see that

$$\mu_{\mathbf{p}}(B_n(z)) = \prod_{l=1}^{n} p_{i_l} \cdot \prod_{l=1}^{L_n(z)} \frac{a_{i_l j_l}^{t(\mathbf{p})}}{\sum_{j=1}^{n_{i_l}} a_{i_l j}^{t(\mathbf{p})}}.$$

Exercise: Using Kolmogorov's Strong Law of Large Numbers prove that

$$\lim_{n\to\infty} \frac{\log \mu_{\mathbf{p}}(B_n(z))}{\sum_{l=1}^{n} \log b_{i_l}} = d(\mathbf{p}) \ \text{ for } \tilde{\mu}_{\mathbf{p}}\text{-a.e. } z.$$

By the Volume Lemma, this implies that $\dim_{\mathrm{H}} \mu_{\mathbf{p}} = d(\mathbf{p})$ (we can use approximate squares instead of balls) as we wish. As observed in the beginning of this part, this implies

$$\dim_{\mathrm{H}} \Lambda \ge \sup_{\mathbf{p}} d(\mathbf{p})$$

thus concluding Part 1.

Part 2: $\dim_{\mathrm{H}} \Lambda \le \sup_{\mathbf{p}} d(\mathbf{p})$

This is the harder part since, as observed in the beginning of this lesson, we cannot use covers by basic sets of order n to get a good estimate for the Hausdorff dimension. Also we do not want any estimate, we want the estimate

$$\sup_{\mathbf{p}} d(\mathbf{p}).$$

Remark 3. Surprisingly, if

$$b_i \le a_{ij} \le b_i^2 \tag{20}$$

then if follows from [HL] that we can actually use Bowen's equation to compute the Hausdorff dimension of Λ (condition (20) implies that the number of basic sets of order n we need to form the approximate square "is not significant"). See also [L2] for an extension of [HL] to nonlinear transformations.

The main idea is to use an extension of the following formula to our setting: Given $E \subset \mathbb{R}^k$ and $F \subset \mathbb{R}^n$ bounded,

$$\dim_{\mathrm{H}}(E \times F) \leq \dim_{\mathrm{H}} E + \overline{\dim}_{\mathrm{B}} F. \tag{21}$$

There is a new element in the scenario we need to introduce: the *box counting dimension* of a set F, written $\overline{\dim}_{\mathrm{B}} F$ (see the book [F] for more details). Given $\delta > 0$, let $N(F, \delta)$ be the smallest number of balls of radius δ needed to cover F. Then

Definition 4. The *box counting dimension* of F is

$$\overline{\dim}_{\mathrm{B}} F = \overline{\lim_{\delta \to 0}} \frac{\log N(F, \delta)}{-\log \delta}.$$

Since in the definition of Hausdorff dimension we use infimum over covers formed by balls of radius $\leq \delta$, it is easy to see that

$$\dim_{\mathrm{H}} F \leq \overline{\dim}_{\mathrm{B}} F.$$

There are good things and bad things about this new dimension.

Good:
(i) Formula (21);
(ii) It is relatively easy to compute: consider a partition of the ambient space into cubes of side-length 2^{-n}, count the number of these cubes that intersect the set F, and then see how this number grows when n goes to ∞ (the exponential rate in base 2);
(iii) For 1-dimensional dynamically defined sets F we have $\dim_{\mathrm{H}} F = \overline{\dim}_{\mathrm{B}} F$.

Bad:
(i) In higher dimensions, Hausdorff dimension and box counting dimension need not coincide: It is proved in [GL] that the coincidence of these dimensions for the sets we are considering here only happens for exceptional choices of the numbers b_i and a_{ij} !
(ii) $\overline{\dim}_{\mathrm{B}} F = \overline{\dim}_{\mathrm{B}} \overline{F}$. In particular, $\overline{\dim}_{\mathrm{B}}(\mathbb{Q} \cap [0,1]) = 1$. This does not happen with Hausdorff dimension because

$$\dim_{\mathrm{H}} \Big(\bigcup_{i=1}^{\infty} F_i \Big) = \sup_i \dim_{\mathrm{H}} F_i, \tag{22}$$

in particular, the Hausdorff dimension of a countable set is 0.

Suggestion: Try to prove formula (21). For a proof see [F].

Of course our set Λ need not be a product of two Cantor sets. We can write (see (17))

$$\Lambda = \bigcup_{y \in \pi(\Lambda)} \Lambda_y$$

in an attempt to fall into formula (21) with $\pi(\Lambda)$ playing the role of E and Λ_y playing the role of F. But this is no good because

$$\pi(\Lambda) \ni y \mapsto \overline{\dim}_{\mathrm{B}} \Lambda_y \tag{23}$$

is far from being constant. So the idea is to decompose $\pi(\Lambda)$ into subsets such that the function (23) restricted to these subsets is constant. In fact, as argued in remarks after Theorem 2, the value $\overline{\dim}_{\mathrm{B}} \Lambda_y$ should only depend on the vertical distribution of the orbit of y along the horizontal strips. In this direction, let (for more details see [L1])

$$G_{\mathbf{p}} = \{y \in \pi(\Lambda) : \text{“the orbit of } y \text{ has distribution } \mathbf{p}\text{”}\}.$$

Then

$$y \in G_{\mathbf{p}} \Rightarrow \overline{\dim}_{\mathrm{B}} \Lambda_y = t(\mathbf{p}).$$

So if we define

$$\Lambda_{\mathbf{p}} = \bigcup_{y \in G_{\mathbf{p}}} \Lambda_y,$$

then we have an extension of formula (21) to the set $\Lambda_{\mathbf{p}}$, with $G_{\mathbf{p}}$ playing the role of E and $\Lambda_y, y \in G_{\mathbf{p}}$ playing the role of F (these sets depend on y but their box counting dimension do not), namely

$$\dim_{\mathrm{H}} \Lambda_{\mathbf{p}} \leq \dim_{\mathrm{H}} G_{\mathbf{p}} + t(\mathbf{p}).$$

For proving this formula we use the domination condition $b_i < a_{ij}$. Moreover,

$$\dim_{\mathrm{H}} G_{\mathbf{p}} = \frac{\sum_{i=1}^m p_i \log p_i}{\sum_{i=1}^m p_i \log b_i},$$

and so

$$\dim_{\mathrm{H}} \Lambda_{\mathbf{p}} \leq \frac{\sum_{i=1}^m p_i \log p_i}{\sum_{i=1}^m p_i \log b_i} + t(\mathbf{p}).$$

Now since every $y \in \pi(\Lambda)$ has some distribution $\mathbf{p}$, we have

$$\Lambda = \bigcup_{\mathbf{p}} \Lambda_{\mathbf{p}},$$

and so, by formula (22),

$$\dim_{\mathrm{H}} \Lambda = \sup_{\mathbf{p}} \dim_{\mathrm{H}} \Lambda_{\mathbf{p}} \leq \sup_{\mathbf{p}} \left\{ \frac{\sum_{i=1}^m p_i \log p_i}{\sum_{i=1}^m p_i \log b_i} + t(\mathbf{p}) \right\}.$$

In fact, we are cheating a little bit because the set of all distributions $\{\mathbf{p}\}$ is not countable. Nevertheless, this set is compact and the function

$$\mathbf{p} \mapsto \dim_{\mathrm{H}} \Lambda_{\mathbf{p}}$$

is continuous, so we can use compacteness arguments to obtain what we want. This concludes Part 2 and the sketch of proof of Theorem 2.

□

Nonlinear skew-product transformations

In this section we consider $C^{1+\delta}$ expanding maps $f \colon \mathbb{T}^2 \to \mathbb{T}^2$ of the form

$$f(x, y) = (a(x, y), b(y)) \tag{24}$$

satisfying the domination condition

$$\min_x \partial_x a(x, y) > b'(y) > 1 \quad \text{for all } y, \tag{25}$$

and f-invariant sets Λ possessing a good Markov structure.

For example, if (f_0, Λ_0) is a general Sierpinski carpet, we take f $C^{1+\delta}$-close to f_0 satisfying (24) and Λ to be the *hyperbolic f-continuation* of Λ_0. Then f satisfies (25) and $f|\Lambda$ is topologically conjugated to a full shift.

In Lecture 1 we saw that the transition from self-afine Cantor sets to nonlinear Cantor sets was done using the bounded distortion property and the Moran formula was substituted by Bowen's equation in terms of the Topological Pressure. Now the self-affine sets that model the nonlinear transformations of skew-product type satisfying (24) and (25) are the ones considered in the previous section: the condition of basic rectangles being aligned along horizontal strips translates into (24), and the domination condition $b_i < a_{ij}$ translates into (25). So using bounded distortion arguments we should be able to find a formula for the Hausdorff dimension of Λ extending (14).

In the self-affine setting we considered probability vectors $\mathbf{p}$ that correspond to Bernoulli measures $\nu_{\mathbf{p}}$ on $\pi(\Lambda)$. As before π is the projection onto the y-axis, $\pi(x, y) = y$. Note that $\pi \circ f = b \circ \pi$, and we say that $b|\pi(\Lambda)$ is a factor of $f|\Lambda$ (in the self-affine case we called b the "vertical dynamics"). In this nonlinear setting, Bernoulli measures on $\pi(\Lambda)$ should be replaced by the bigger set $\mathcal{M}_e(b|\pi(\Lambda))$ (the set of "vertical" ergodic measures). By what expression should we replace

$$\frac{\sum_{i=1}^{m} p_i \log p_i}{\sum_{i=1}^{m} p_i \log b_i}\,? \tag{26}$$

By this time the reader should recognize this expression as the Hausdorff dimension of the Bernoulli measure $\nu_{\mathbf{p}}$ which is the quotient between the entropy of this measure by its Lyapunov exponent (remember (12)). So, expression (26) should be replaced by

$$\dim_{\mathrm{H}} \nu = \frac{h_\nu(b)}{\int \log b' d\nu}$$

for $\nu \in \mathcal{M}_e(b|\pi(\Lambda))$. Now the number $t(\mathbf{p})$ given by the random Moran formula (15) should be replaced by what? Since, in Lecture 1, Moran formula was replaced by Bowen's equation in terms of the Topological Pressure, now we should have a *random Bowen's equation* in terms of the *random Topological Pressure*.

As before, let

$$R_{(i_1 j_1)\dots(i_n j_n)}, \quad (i_l, j_l) \in \mathcal{I},\ l = 1, \dots, n$$

be the basic sets of order n and

$$R_{i_1 \dots i_n}, \quad i_l \in \{1, \dots, m\},\ l = 1, \dots, n$$

be the *basic intervals of order* n obtained by projecting the first ones onto the y-axis. We use the notation $y \equiv (i_1, i_2, \dots)$ iff

$$y \in \bigcap_{n=1}^{\infty} R_{i_1 \dots i_n}.$$

Definition 5. Let $\varphi \colon \Lambda \to \mathbb{R}$ be a Hölder continuous function. Given $y \in \pi(\Lambda)$ with $y \equiv (i_1, i_2, \dots)$, the *Relative Pressure* of φ with respect to the fibre $\pi^{-1}(y)$ is

$$P(f|\Lambda, \varphi, \pi^{-1}(y)) = \varlimsup_{n\to\infty} \frac{1}{n} \log \sum_{j_1,\dots,j_n} \sup_{x \in R_{(i_1 j_1)\dots(i_n j_n)} \cap \pi^{-1}(y)} e^{S_n \varphi(x,y)}.$$

Then, given $\nu \in \mathcal{M}_e(b|\pi(\Lambda))$, the number $t(\mathbf{p})$ should be replaced by the number $t(\nu)$ which is the root of

$$\int_{\pi(\Lambda)} P(f|\Lambda, -t(\nu) \log \partial_x a, \pi^{-1}(y))\, d\nu(y) = 0. \tag{27}$$

So we are led to the following question:

$$\dim_{\mathrm{H}} \Lambda = \sup_{\nu \in \mathcal{M}_e(b|\pi(\Lambda))} \left\{ \frac{h_\nu(b)}{\int \log b' d\nu} + t(\nu) \right\} ? \tag{28}$$

Before answering affirmatively to this question, we should interpret what is the expression between brackets in (28). In the self-affine case this expression was the Hausdorff dimension of a Bernoulli measure $\mu_{\mathbf{p}}$ which projects under π to $\nu_{\mathbf{p}}$: it is the vertical dimension plus the dimension of typical horizontal fibres. Now, this expression still is the Hausdorff dimension of an ergodic measure μ_ν projecting to

ν. This is the time for introducing the *Relativized Variational Principle* due to [LW] (in a more general context).

Theorem 3. (Relativized Variational Principle)
Given $\varphi\colon \Lambda \to \mathbb{R}$ *continuous and* $\nu \in \mathcal{M}(b|\pi(\Lambda))$,

$$\sup_{\substack{\mu\in\mathcal{M}(f|\Lambda)\\ \mu\circ\pi^{-1}=\nu}} \left\{h_\mu(f) - h_\nu(b) + \int_\Lambda \varphi\, d\mu\right\} = \int_{\pi(\Lambda)} P(f|\Lambda, \varphi, \pi^{-1}(y))\, d\nu(y). \tag{29}$$

We say that μ is an *equilibrium state* for (29) if the supremum is attained at μ.

It follows from the work of [DG, DGH] the existence of a unique equilibrium state for (29) relative to any $\nu \in \mathcal{M}(b|\pi(\Lambda))$ and any Hölder-continuous φ. Moreover, the unique equilibrium state is ergodic if ν is ergodic, and has an important *Gibbs property*, as we shall describe now.

A bad thing about working with the relative pressure is that, in general, the function

$$y \mapsto P(f|\Lambda, \varphi, \pi^{-1}(y))$$

is only measurable. Now if φ is Hölder continuous then [DG, DGH] constructed a Hölder continuous function $A_\varphi\colon \pi(\Lambda) \to \mathbb{R}$, called *Gauge function*, such that

$$\int_{\pi(\Lambda)} \log A_\varphi\, d\nu = \int_{\pi(\Lambda)} P(f|\Lambda, \varphi, \pi^{-1}(y))\, d\nu(y) \tag{30}$$

for every $\nu \in \mathcal{M}(b|\pi(\Lambda))$. Moreover there is a *Gibbs family of measures* $\{\mu_{\varphi,y}\}_{y\in\pi(\Lambda)}$ such that: the unique equilibrium state μ for (29) relative to $\nu \in \mathcal{M}(b|\pi(\Lambda))$ and φ is $\mu = \nu \times \mu_{\varphi,y}$, and there exist positive constants c_1, c_2 such that, for all $y \in \pi(\Lambda)$,

$$c_1 \le \frac{\mu_{\varphi,y}(R_n \cap \pi^{-1}(y))}{\exp\{S_n\varphi(z) + S_n(\log A_\varphi)(y)\}} \le c_2 \tag{31}$$

for all $n \in \mathbb{N}$, basic rectangle of order n R_n, and $z \in R_n \cap \pi^{-1}(y)$.

Let μ_ν be the equilibrium state for (29) relative to $\nu \in \mathcal{M}(b|\pi(\Lambda))$ and the potential $\varphi = -t(\nu)\log\partial_x a$, where $t(\nu)$ is given by (27) so that

$$\int_{\pi(\Lambda)} \log A_{-t(\nu)\log\partial_x a}\, d\nu = 0. \tag{32}$$

Then, using the Gibbs property (31) and (32), we can prove (see [L1]):

Lemma 3. *If* $\nu \in \mathcal{M}_e(b|\pi(\Lambda))$ *then*

$$\dim_{\mathrm{H}} \mu_\nu = \frac{h_\nu(b)}{\int \log b' d\nu} + t(\nu).$$

And

Theorem 4.

$$\dim_{\mathrm{H}} \Lambda = \sup_{\nu\in\mathcal{M}_e(b|\pi(\Lambda))} \dim_{\mathrm{H}} \mu_\nu.$$

In particular, the variational principle for dimension holds.

Problem: Is there an ergodic invariant measure of full dimension?

This problem is more difficult than it looks at first sight. In fact, we can see that the map

$$\mathcal{M}_e(b|\pi(\Lambda)) \ni \nu \mapsto \dim_{\mathrm{H}} \mu_\nu = \frac{h_\nu(b)}{\int \log b'\, d\nu} + t(\nu)$$

is upper-semicontinuous. However, we cannot conclude there is an invariant measure of full dimension because the subset $\mathcal{M}_e(b|\pi(\Lambda)) \subset \mathcal{M}(b|\pi(\Lambda))$ is not closed (is dense).

Lecture 3

This lecture is devoted to answer positively to the problem proposed at the end of the previous lecture: *there is an ergodic invariant measure of full dimension.*

This will follow from the abstract result of [L3]:

Theorem 5. *Let (X,T) and (Y,S) be mixing subshifts of finite type, and $\pi\colon X\to Y$ be a continuous and surjective mapping such that $\pi\circ T=S\circ\pi$ (S is a factor of T). Let $\varphi\colon X\to\mathbb{R}$ and $\psi\colon Y\to\mathbb{R}$ be positive Hölder continuous functions. Then the maximum of*

$$\frac{h_{\mu\circ\pi^{-1}}(S)}{\int\psi\circ\pi\,d\mu}+\frac{h_\mu(T)-h_{\mu\circ\pi^{-1}}(S)}{\int\varphi\,d\mu} \tag{33}$$

over all $\mu\in\mathcal{M}(T)$ is attained on the set $\mathcal{M}_e(T)$.

Remark 4. This theorem answers positively to Problem 2 raised in [GP2] for mixing subshifts of finite type and Hölder continuous potentials. So, it also applies to obtain an invariant ergodic measure of full dimension for a class of transformations treated in [GP1].

For proving this theorem we will need a new variational principle for the topological pressure of certain noncompact sets.

Conditional Variational Principle for the Topological Pressure

The notion of topological entropy for noncompact sets was introduced by Bowen in the beautiful paper [Bo3]. Later this notion was generalized for the topological pressure of noncompact sets in [PP], for a definition we refer the reader to this paper (note the resemblance with the definition of Hausdorff dimension).

Given a continuous map $T\colon X\to X$ of a compact metric space, we denote by $P(\psi,K)$ the topological pressure associated to a continuous function $\psi\colon X\to\mathbb{R}$ and a T-invariant set K (not necessarily compact), as defined in [PP]. Let

$$I_\psi=\left(\inf_{\mu\in\mathcal{M}(T)}\int\psi\,d\mu,\ \sup_{\mu\in\mathcal{M}(T)}\int\psi\,d\mu\right).$$

Theorem 6. ([L3])
Let (X,T) be a mixing subshift of finite type, and $\varphi,\psi\colon X\to\mathbb{R}$ Hölder continuous functions. For $\alpha\in\mathbb{R}$ let

$$K_\alpha=\left\{x\in X:\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\psi(T^i(x))=\alpha\right\}.$$

If $0\notin\partial I_\psi$ and $\alpha\in I_\psi$ then

$$P(\varphi,K_\alpha)=\sup\left\{h_\mu(T)+\int\varphi\,d\mu:\mu\in\mathcal{M}(T)\ \textit{and}\ \int\psi\,d\mu=\alpha\right\}.$$

Morever, the supremum is attained at a unique measure μ_β which is the Gibbs state with respect to the potential $\varphi+\beta\psi$, for a unique $\beta\in\mathbb{R}$.

Proof. Here we just prove the 'moreover' part since we will not use equality with the pressure of the noncompact set. We have, for all $\beta \in \mathbb{R}$,

$$\sup\left\{h_\mu(T) + \int \varphi\, d\mu : \mu \in \mathcal{M}(T) \text{ and } \int \psi\, d\mu = \alpha\right\}$$
$$= \sup\left\{h_\mu(T) + \int \varphi\, d\mu + \beta \int \psi\, d\mu : \mu \in \mathcal{M}(T) \text{ and } \int \psi\, d\mu = \alpha\right\} - \beta\alpha$$
$$\leq \sup\left\{h_\mu(T) + \int \varphi\, d\mu + \beta \int \psi\, d\mu : \mu \in \mathcal{M}(T)\right\} - \beta\alpha.$$

Now it is well known (see [Bo1]) that the last supremum is uniquely attained at the *Gibbs state* μ_β associated to the potential $\varphi + \beta\psi$ (for the classical variational principle). So we must find a unique β such that $\int \psi\, d\mu_\beta = \alpha$.

We use the abbreviation $P(\cdot) = P(\cdot, X)$ (in Lecture 1 we used the notation $P_T(\cdot)$). It is proved in [R2] that $P(\cdot)$ is a real analytic function on the space of Hölder continuous functions and that

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} P(h_1 + \varepsilon h_2) = \int h_2\, d\mu_{h_1},$$
$$\frac{\partial^2 P(h + \varepsilon_1 h_1 + \varepsilon_2 h_2)}{\partial\varepsilon_1 \partial\varepsilon_2}\Big|_{\varepsilon_1=\varepsilon_2=0} = Q_h(h_1, h_2),$$

where Q_h is the bilinear form defined by

$$Q_h(h_1, h_2) = \sum_{n=0}^{\infty} \left(\int h_1 (h_2 \circ T^n)\, d\mu_h - \int h_1\, d\mu_h \int h_2\, d\mu_h \right), \tag{34}$$

and μ_h is the Gibbs measure for the potential h. Moreover, $Q_h(h_1, h_1) \geq 0$ and $Q_h(h_1, h_1) = 0$ if and only if h_1 is *cohomologous* to a constant function. From this we get that

$$\frac{d}{d\beta} \int \psi\, d\mu_\beta = \frac{d^2}{d\beta^2} P(\varphi + \beta\psi) = Q_{\varphi+\beta\psi}(\psi, \psi) > 0 \tag{35}$$

(the hypothesis $\alpha \in I_\psi$ prevents ψ being cohomologous to a constant). So we must see that

$$\lim_{\beta\to\infty} \int \psi\, d\mu_\beta = \sup_{\mu\in\mathcal{M}(T)} \int \psi\, d\mu, \tag{36}$$

$$\lim_{\beta\to-\infty} \int \psi\, d\mu_\beta = \inf_{\mu\in\mathcal{M}(T)} \int \psi\, d\mu. \tag{37}$$

Proof of (36): We use the notation

$$p(\beta) \sim q(\beta) \quad (\beta \to \infty) \quad \text{means} \quad \lim_{\beta\to\infty} \frac{p(\beta)}{q(\beta)} = 1.$$

We have that

$$\int \psi\, d\mu_\beta = \frac{d}{d\beta} P(\varphi + \beta\psi) = \frac{d}{d\beta} \sup_{\mu\in\mathcal{M}(T)} \left\{h_\mu(T) + \int \varphi\, d\mu + \beta \int \psi\, d\mu\right\}. \tag{38}$$

Since $\mu \mapsto h_\mu(T) + \int \varphi\, d\mu$ is bounded, it is easy to see that

$$\sup_{\mu\in\mathcal{M}(T)} \left\{h_\mu(T) + \int \varphi\, d\mu + \beta \int \psi\, d\mu\right\} \sim \beta \sup_{\mu\in\mathcal{M}(T)} \int \psi\, d\mu, \tag{39}$$

so, using L'Hospital's rule applied to (38) and (39), we obtain (36). The proof of (37) is similar. This concludes the proof of the theorem. □

Now we have everything we need to prove Theorem 5.

Proof of Theorem 5. Since φ is positive then, given $\nu \in \mathcal{M}(S)$, there is a unique real $t(\nu) \in [0,1]$ such that

$$\int_Y P(T, -t(\nu)\varphi, \pi^{-1}(y))\, d\nu(y) = 0 \tag{40}$$

(note that $t \mapsto P(T, -t\varphi, \pi^{-1}(y))$ is strictly decreasing). Denote by μ_ν the unique equilibrium state for (29) relative to ν and $-t(\nu)\varphi$. Then it follows from the relativized variational principle that, for $\mu \in \mathcal{M}(T)$ such that $\mu \circ \pi^{-1} = \nu$,

$$\frac{h_\mu(T) - h_\nu(S)}{\int \varphi\, d\mu} \leq t(\nu) \tag{41}$$

with equality if and only if $\mu = \mu_\nu$. Put

$$D(\mu) = \frac{h_{\mu\circ\pi^{-1}}(S)}{\int \psi \circ \pi\, d\mu} + \frac{h_\mu(T) - h_{\mu\circ\pi^{-1}}(S)}{\int \varphi\, d\mu}$$

and

$$D = \sup_{\mu \in \mathcal{M}(T)} D(\mu). \tag{42}$$

Then it follows by (41) that

$$D = \sup_{\nu \in \mathcal{M}(S)} \left\{ \frac{h_\nu(S)}{\int \psi\, d\nu} + t(\nu) \right\}, \tag{43}$$

and if this supremum is attained at $\nu_0 \in \mathcal{M}_e(S)$ then the supremum in (42) is attained at $\mu_{\nu_0} \in \mathcal{M}_e(T)$ as we wish.

It follows from (43) that

$$\sup_{\nu \in \mathcal{M}(S)} \left\{ h_\nu(S) + (t(\nu) - D) \int \psi\, d\nu \right\} = 0, \tag{44}$$

and if this supremum is attained at $\nu_0 \in \mathcal{M}_e(S)$ then so is the supremum in (43) and thus the supremum in (42) is attained at $\mu_{\nu_0} \in \mathcal{M}_e(T)$ as we wish. Let

$$\underline{t} = \inf_{\nu \in \mathcal{M}(S)} t(\nu) \quad \text{and} \quad \overline{t} = \sup_{\nu \in \mathcal{M}(S)} t(\nu).$$

The supremum in (44) can be rewritten as

$$\sup_{\underline{t} \leq t \leq \overline{t}} \ \sup_{\substack{\nu \in \mathcal{M}(S) \\ t(\nu) = t}} \left\{ h_\nu(S) + \int (t - D)\psi\, d\nu \right\}. \tag{45}$$

According to [DG, DGH], there is a Hölder continuous function $A_{-t\varphi} \colon Y \to \mathbb{R}$ such that

$$\int \log A_{-t\varphi}\, d\nu = \int P(T, -t\varphi, \pi^{-1}(y))\, d\nu(y), \tag{46}$$

so by (40),

$$t(\nu) = t \Leftrightarrow \int \log A_{-t\varphi}\, d\nu = 0.$$

So, the supremum in (45) can be rewritten as

$$\sup_{\underline{t} \leq t \leq \overline{t}} \sup \left\{ h_\nu(S) + \int (t - D)\psi\, d\nu : \nu \in \mathcal{M}(S) \text{ and } \int \log A_{-t\varphi}\, d\nu = 0 \right\}. \tag{47}$$

We assume the supremum above is not attained at $\underline{t}$ or $\overline{t}$ (see [L3] for more details). Applying Theorem B (we should verify its hypotheses) we get that the intermediate supremum in (47) is attained at the Gibbs measure (hence ergodic) $\nu_{\beta(t)}$ for the potential $(t-D)\psi + \beta(t)\log A_{-t\varphi}$, and the value of this supremum is, with $P(\cdot) = P(\cdot, Y)$,

$$h(t) = P((t-D)\psi + \beta(t)\log A_{-t\varphi}), \tag{48}$$

where $\beta(t)$ is the unique real satisfying

$$\int \log A_{-t\varphi}\, d\nu_{\beta(t)} = 0 \quad \text{i.e.} \quad t(\nu_{\beta(t)}) = t.$$

So we should see that the function $(\underline{t}, \overline{t}) \ni t \mapsto h(t)$ is continuous. This follows from the continuity of $t \mapsto A_{-t\varphi}$ and $\beta(t)$ (see [L3] for more details), together with $|P(\varphi_1) - P(\varphi_2)| \le \|\varphi_1 - \varphi_2\|$ (see [R2]). So if the supremum of $(\underline{t}, \overline{t}) \ni t \mapsto h(t)$ is attained at $t^* \in (\underline{t}, \overline{t})$ then

$$D = D(\mu_{\nu_{\beta(t^*)}}),$$

and this concludes the proof of the theorem.

□

Measure of full dimension

Let $f\colon \mathbb{T}^2 \to \mathbb{T}^2$, $f(x,y) = (a(x,y), b(y))$ and Λ such that $f(\Lambda) = \Lambda$ be as in Lecture 2. Let $\pi\colon \mathbb{T}^2 \to \mathbb{T}^1$ be the projection given by $\pi(x,y) = y$. Then $\pi \circ f = b \circ \pi$, and we are in the conditions of Theorem 5 with

$$T = f|\Lambda, \quad S = b|\pi(\Lambda), \quad \varphi = \log \partial_x a, \quad \psi = \log b'.$$

Now we improve Theorem 4 in Lecture 2.

Theorem 7. *There exists an ergodic invariant measure μ on Λ such that*

$$\dim_{\mathrm{H}} \Lambda = \dim_{\mathrm{H}} \mu.$$

Morever, μ is a Gibbs state for a relativized variational principle.

Proof. By the proof of Theorem 5,

$$\sup_{\nu \in \mathcal{M}_e(b|\pi(\Lambda))} \left\{ \frac{h_\nu(b)}{\int \log b'\, d\nu} + t(\nu) \right\}$$

is attained at some $\nu_0 \in \mathcal{M}_e(b|\pi(\Lambda))$. Then, by Lemma 3 and Theorem 4,

$$\dim_{\mathrm{H}} \Lambda = \dim_{\mathrm{H}} \mu_{\nu_0}.$$

□

Question: Is there a *unique* ergodic measure of full dimension ?

It follows from the proof of Theorem 5 that ergodic measures of full dimension are given by $\mu_{\nu_{\beta(t)}}$ where t is a zero (i.e. a maximizing point) of the function h defined by (48). For instance, this is the case if h is C^2 and $h'' < 0$.