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 $[\frac{k_1}{2^{-n}},\frac{k_1+1}{2^{-n}}]\times\ldots\times[\frac{k_d}{2^{-n}},\frac{k_d+1}{2^{-n}}]$, $k_j\in\boldsymbol{Z}$ 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 $(U_i, i=1,2,\ldots)$ 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 $\{1/n, n=1,2,\ldots\}$ 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 $(i_0,i_1,\ldots)$ with $i_n\in\{1,2,\ldots,d\}$. Consider the metric

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

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

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

, equivalent to the previous one. Consider $\sigma:\Sigma^d\to\Sigma^d$ defined by $f((i_0,i_1,\ldots)=(i_1,\ldots)$. 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

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

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