11. Piecewise linear transformations of the unit interval and Cantor sets

11.1 Introduction

To study ergodic properties of dynamical sytems, the spectral study is one of the important tools. By the spectra of the unitary operator associated with transformations, we get the relations between eigenvalue 1 and the ergodicity, and also between point spectra and weak mixing.

Lasota and Yorke [218] studied for expanding maps of the interval the relations between the spectra of the Perron–Frobenius operator and its ergodicity. We summarize the definitions of the unitary operator and the Perron–Frobenius operator, and also the relations of the spectra of these operators and ergodicity of dynamical systems in page 302.

Hofbauer and Keller [171] showed the relations between the singularities of the Ruelle's zeta function and the spectra of the Perron–Frobenius operator. Baladi and Ruelle [45] and Mori [253] also proved the similar results. However, their methods are quite different.

In the first part of this chapter, along ([253]), though the Perron–Frobenius operator is not compact, we will give the 'Fredholm determinant' using generating functions and renewal equations. By this Fredholm determinant, we can give the spectra of the Perron–Frobenius operator and also the Ruelle's zeta function concretely. Thus we can determine the ergodic properties of the dynamical system such as ergodicity, mixing and decay rate of correlation, by this Fredholm determinant.

In a second part of this chapter, as an application of the first part, we compute the Hausdorff dimension of Cantor sets associated in a natural way with piecewise linear maps, by introducing the α–Fredholm matrix.

11.1.1 One–dimensional dynamical system

Consider a map F from the unit interval $[0,1]$ into itself. Our main interest is to study the asymptotic behavior of orbits x, $F(x)$, $F^2(x)$, $\cdots$. Here,

$$F^n(x) = \begin{cases} x & \text{if } n = 0, \\ F^{n-1}(F(x)) & \text{if } n \geq 1, \end{cases}$$

denotes the n–th iteration and not the n–th power $(F(x))^n$ of F. The cases which we want to study are the cases such that there exists a probability measure μ on $[0,1]$ which satisfies the following conditions:

1. μ is absolutely continuous with respect to the Lebesgue measure, i.e., there exists a *density function* $\varrho(x)$ (that we denote $\varrho(x) = \frac{d\mu}{dx}(x)$) such that $\int_0^1 \varrho\, dx = 1$ and for every measurable set A, $\mu(A) = \int_A \varrho(x)\, dx$;
2. for any measurable set A, $\mu(A) = \mu(F^{-1}(A))$ holds (μ is an *invariant measure*).

From the physical view point, F is a *time evolution*, and μ is an *equilibrium state*, i.e., we are studying a dynamical system in an equilibrium state. To be more precise, we should mention the σ–algebra which is generated by the intervals. However in our discussion, it does not play an essential role, thus we shall omit it.

The maps which we consider in this chapter are piecewise monotonic transformations. We call F a *piecewise monotonic transformation* if there exists a finite set $\mathcal{A}$ and a finite partition into intervals $\{\langle a\rangle\}_{a\in\mathcal{A}}$ of $[0,1]$ such that on each subinterval $\langle a\rangle$ the map F is monotone (increasing or decreasing).

Definition 11.1.1. (1) *we call a transformation F* piecewise linear *if the derivative F' is constant on each interval of monotonicity $\langle a\rangle$.*

(2) *We call a transformation F* Markov *if the following hold:*

$$\text{if } F(\langle a\rangle) \cap \langle b\rangle^o \neq \emptyset, \text{ then } \overline{F(\langle a\rangle)} \supset \langle b\rangle,$$

where we denote by J^o and $\overline{J}$ the interior and the closure of a set J, respectively. Namely, for any $a \in \mathcal{A}$, $F(\langle a\rangle)$ is essentially a union of several $\langle b\rangle$ $(b \in \mathcal{A})$.

(3) *We call a map F* expanding *if $\xi > 0$, where*

$$\xi = \liminf_{n\to\infty} \frac{1}{n} \operatorname*{ess\,inf}_{x\in[0,1]} \log |F^{n\prime}(x)|.$$

11.1.2 Symbolic Dynamics

Let $\mathcal{A}$ be a finite set. Let T be the shift on $\mathcal{A}^{\mathbb{N}}$. We call a triple (X, T, ν) a *symbolic dynamical system* if X is a closed invariant subset of $\mathcal{A}^{\mathbb{N}}$ (i.e., $T^{-1}X \supset X$) and ν is an invariant probability measure. Here, again, we should mention about the σ–algebra which is generated by cylinder sets, but we shall also omit it.

The one–dimensional dynamical system which we considered in the former subsection can be expressed in terms of symbolic dynamics. Namely, we code in a natural way the orbits of the points of the unit inteval with respect to the partition into intervals of monotonicity. More precisely, let us take a piecewise monotonic transformation F. Let $\{\langle a\rangle\}_{a\in\mathcal{A}}$ be a partitions of $[0,1]$

into subintervals of monotonicity, i.e., F is monotone on each $\langle a\rangle$, $\cup_{a\in\mathcal{A}}\langle a\rangle = [0,1]$ and $\langle a\rangle \cap \langle b\rangle = \emptyset$ for $a \neq b$. Let $x \in [0,1]$. We then define the expansion $W^x = (w_n^x)_{n\in\mathbb{N}}$ of x by:

$$\forall n \in \mathbb{N}^+,\ F^n(x) \in \langle w_{n+1}^x\rangle.$$

We denote by X the closure of all the expansions of the points x: $X := \{W^x | x \in [0,1]\}$. Then the shift on X corresponds to F on $[0,1]$, that is, for a sequence $W^x \in X$ which is the expansion of x, the expansion of $F(x)$ equals $T(W^x)$.

11.1.3 Notation

Let us fix the notation. Let $W = w_1 \cdots w_n$ be a finite word in $\mathcal{A}^\star$. We denote the length of a word W by $|W|$, i.e., $|W| = n$. Let

$$\langle W\rangle = \bigcap_{k=1}^{n} F^{-k+1}(\langle w_k\rangle) \subset [0,1].$$

be defined as the interval coded by the word W, i.e.

$$\forall x \in [0,1],\ x \in \langle w_1\rangle, \dots, F^{n-1}(x) \in \langle w_n\rangle.$$

We say a word W is *admissible* if $\langle W\rangle \neq \emptyset$. For an admissible word W and for any point $x \in [0,1]$, we inductively define the point Wx as follows (if it exists):

1. $F^{|W|}(Wx) = x$,
2. $w_1 \cdots w_n x \in \langle w_1\rangle$,
3. $F(w_1 \cdots w_n x) = w_2 \cdots w_n x$ $(n \geq 2)$.

For a Markov map F, the following $\mathcal{A} \times \mathcal{A}$ matrix M plays an essential role, particularly when F is piecewise linear. Let for $a, b \in \mathcal{A}$

$$M_{a,b} = \begin{cases} 1 & \text{if } \overline{F(\langle a\rangle)} \supset \langle b\rangle, \\ 0 & \text{otherwise.} \end{cases}$$

We call this matrix the *structure matrix* of F. We call M *irreducible*, if for any $a, b \in \mathcal{A}$, there exists n such that $(M^n)_{a,b} > 0$, and we call M *periodic* if there exists an integer m such that if $(M^n)_{a,b} > 0$, then n can be divided by m. Otherwise, we call M *aperiodic*. When M is irreducible and aperiodic, there exists n such that all the components of M^n are positive: M is thus said to be *positive*.

We denote by L^1 L^∞ the set of functions f such that $\int_0^1 |f|\,dx < \infty$ and $\sup_{x\in[0,1]} |f(x)| < \infty$, respectively. To study the ergodic properties of the dynamical systems we consider here, the space of functions with *bounded variation* is one of the most important notions. We will denote it by BV.

Definition 11.1.2. *A function $f : [0,1] \to [0,1]$ is of bounded variation if its total variation*

$$Var(f) := \sup \sum_{k=1}^{n} |f(x_k) - f(x_{k+1})|,$$

is finite. The supremum is taken over all finite subdivisions of $[0,1]$ $(x_k)_{1\le k\le n}$.

A function $f \in L^1$ is of bounded variation if $v(f) = \inf Var(g)$ is finite. The infimum is taken over the class of f such that $f = g$ in L^1.

11.2 Ergodic Properties of some piecewise monotonic transformations

The aim of this section is to illustrate through the study of "simple" examples how to deduce from the spectral study of the Perron–Frobenius operator ergodic properties of the dynamical systems arising from piecewise monotonic transformations.

11.2.1 Piecewise linear Markov transformation

To avoid any technical difficulty, let us start with a simple example, namely piecewise linear Markov transformations.

Let

$$F(x) = \begin{cases} x/\eta_a & 0 \le x < \eta_a \\ (x-\eta_a)/\eta_b & \eta_a \le x \le 1, \end{cases}$$

where $1 - \eta_a = \eta_a\eta_b$, that is, $\lim_{x\uparrow 1} F(x) = \eta_a$. Let us take as alphabet $\mathcal{A} = \{a, b\}$, and corresponding intervals $\langle a\rangle = [0, \eta_a)$ and $\langle b\rangle = [\eta_a, 1]$. We are interested first in ergodic properties such as the existence of an invariant measure, and the second in statistical properties, for instance mixing. Let us recall that mixing deals with the convergence of $\int f(x)g(F^n(x)\,dx$ towards $\int f\,dx \int g\,d\mu$. For more details, see Chap. 3 or [367]. The rate of convergence is called the *decay of correlation.* For that purpose, let us introduce some tools and notations, as the Perron–Frobenius operator, the generating functions or the Fredholm matrix.

The Perron–Frobenius operator.

Definition 11.2.1. *The operator defined on L^1 by:*

$$Pf(x) = \sum_{y:\ F(y)=x} f(y)|F'(y)|^{-1},$$

is called the Perron–Frobenius operator *associated with F.*

For more details on this operator, see for instance [80, 196, 280]. This operator corresponds to the change of variables $F(x) \to x$ in the following integration:

$$\int f(x)g(F(x))\,dx = \int Pf(x)g(x)\,dx$$

for $f \in L^1$ and $g \in L^\infty$. However, restricting its domain to BV the set of bounded variation functions, the eigenvalues of this operator determine the ergodic properties of the dynamical system [218].

Let us review the properties of the Perron–Frobenius operator:

1. $P\colon L^1 \to L^1$ is positive,
2. the spectra of P on the unit circle coincide with the spectra of the unitary operator defined by
$$Uf(x) = f(F(x)).$$

From this fact, we can deduce the following:

1. 1 is an eigenvalue, and the eigenspace associated with eigenvalue 1 is contained in BV. Moreover, we can choose a basis with nonnegative functions $\varrho \in BV$ such that $\int \varrho\,dx = 1$, and these ϱ are the density functions of invariant probability measures ([218]);
2. if 1 is a simple eigenvalue, then the dynamical system is ergodic;
3. if there exists no other eigenvalue except 1 on the unit circle; then the dynamical system is mixing, i.e., for $f \in L^1$ and $g \in L^\infty$
$$\int f(x)g(F^n(x))\,dx \to \int f\,dx \int g\,d\mu,$$
where μ is the invariant probability measure absolutely continuous with respect to the Lebesgue measure.

As an operator from L^1 into itself, every $|z| < 1$ are eigenvalues of P with infinite multiplicity ([201]). Thus the rate of convergence of $\int f(x)g(F^n(x))\,dx$ to $\int f\,dx \int g\,d\mu$ has no meaning.

When we restrict P to BV, then every $|z| < e^{-\xi}$ are eigenvalues with infinite multiplicity. Thus the eigenvalues of P restricted to BV in the annulus $e^{-\xi} < |z| \leq 1$ play essential role to determine the ergodic properties (ergodicity, mixing, the decay rate of convergence and so on).

Remark. Let 1_J denote the indicator function of a set J. Note that

$$\begin{aligned} P1_{\langle a\rangle}(x) &= \sum_{y\colon F(y)=x} 1_{\langle a\rangle}(y)|F'(y)|^{-1} \\ &= \eta_a \\ P1_{\langle b\rangle}(x) &= \begin{cases} \eta_b & \text{if } x \in \langle a\rangle \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Generating functions and Fredholm matrix. Define for $g \in L^\infty$, the *generating functions* (these are formal power series) as

$$s_g^{\langle a\rangle}(z) = \sum_{n=0}^{\infty} z^n \int 1_{\langle a\rangle}(x) g(F^n(x))\,dx,$$
$$s_g^{\langle b\rangle}(z) = \sum_{n=0}^{\infty} z^n \int 1_{\langle b\rangle}(x) g(F^n(x))\,dx,$$
$$s_g(z) = \begin{pmatrix} s_g^{\langle a\rangle}(z) \\ s_g^{\langle b\rangle}(z) \end{pmatrix}.$$

Since

$$s_g^{\langle a\rangle}(z) = \sum_{n=0}^{\infty} z^n \int \left[P^n 1_{\langle a\rangle}\right](x)\, g(x)\,dx,$$

we get as a formal expression

$$s_g^{\langle a\rangle}(z) = \int \left[(I - zP)^{-1} 1_{\langle a\rangle}\right](x)\, g(x)\,dx.$$

This asserts, very roughly, that the reciprocals of eigenvalues of the Perron–Frobenius operator become singular points of $s_g^{\langle a\rangle}(z)$.

The aim of this section is thus to use the above generating functions, in order to study the ergodic properties of the map F. We are interested in particular in the eigenvalues of the Perron–Frobenius operator in order to prove the existence of an an invariant probability measure μ, absolutely continuous with respect to the Lebesgue measure such that the dynamical system $([0,1], \mu, F)$ is mixing.

Lemma 11.2.1. *Let $\Phi(z)$ be the matrix (called a* Fredholm matrix*)*

$$\Phi(z) = \begin{pmatrix} z\eta_a & z\eta_a \\ z\eta_b & 0 \end{pmatrix}.$$

Then taking $s_g(z) = \begin{pmatrix} s_g^{\langle a\rangle}(z) \\ s_g^{\langle b\rangle}(z) \end{pmatrix}$, we get

$$s_g(z) = \begin{pmatrix} \int_{\langle a\rangle} g\,dx \\ \int_{\langle b\rangle} g\,dx \end{pmatrix} + \Phi(z) s_g(z).$$

Proof. Dividing the sum of the generating functions into the term corresponding to $n = 0$ and the remaining part, we get:

$$
\begin{aligned}
s_g^{\langle a\rangle}(z) &= \int 1_{\langle a\rangle}(x)g(x)\,dx + \sum_{n=1}^{\infty} z^n \int 1_{\langle a\rangle}(x)g(F^n(x))\,dx \\
&= \int 1_{\langle a\rangle}(x)g(x)\,dx + \sum_{n=1}^{\infty} z^n \int P1_{\langle a\rangle}(x)g(F^{n-1}(x))\,dx \\
&= \int_{\langle a\rangle} g(x)\,dx + \eta_a \sum_{n=1}^{\infty} z^n \int g(F^{n-1}(x))\,dx \\
&= \int_{\langle a\rangle} g(x)\,dx + z\eta_a \sum_{n=0}^{\infty} z^n \int g(F^{n}(x))\,dx \\
&= \int_{\langle a\rangle} g(x)\,dx + z\eta_a \left(s_g^{\langle a\rangle}(z) + s_g^{\langle b\rangle}(z)\right).
\end{aligned}
$$

Similarly, we get

$$
\begin{aligned}
s_g^{\langle b\rangle}(z) &= \int 1_{\langle b\rangle}(x)g(x)\,dx + \sum_{n=1}^{\infty} z^n \int 1_{\langle b\rangle}(x)g(F^n(x))\,dx \\
&= \int_{\langle b\rangle} g(x)\,dx + \sum_{n=1}^{\infty} z^n \int P1_{\langle b\rangle}(x)g(F^{n-1}(x))\,dx \\
&= \int_{\langle b\rangle} g(x)\,dx + \eta_b \sum_{n=1}^{\infty} z^n \eta_b \int_{\langle a\rangle} g(F^{n-1}(x))\,dx \\
&= \int_{\langle b\rangle} g(x)\,dx + z\eta_b s_g^{\langle a\rangle}(z).
\end{aligned}
$$

■

We thus get

$$
s_g(z) = (I - \Phi(z))^{-1} \begin{pmatrix} \int_{\langle a\rangle} g\,dx \\ \int_{\langle b\rangle} g\,dx \end{pmatrix}.
$$

We call this equation a *renewal equation.* In our case,

$$
s_g(z) = \frac{1}{(1-z)(1-(\eta_a-1)z)} \begin{pmatrix} 1 & z\eta_a \\ z\frac{1-\eta_a}{\eta_a} & 1-z\eta_a \end{pmatrix} \begin{pmatrix} \int_{\langle a\rangle} g\,dx \\ \int_{\langle b\rangle} g\,dx \end{pmatrix}.
$$

An invariant measure for the map F. Let us introduce now an invariant measure for the map F absolutely continuous with respect to the Lebesgue measure. Let us recall that the density of this measure is an eigenvector for the Perron–Frobenius operator associated with the eigenvalue 1, that is to say, it corresponds to the singularity 1 for the generating functions.

Taking

$$
\varrho(x) = \begin{cases} \frac{1}{(2-\eta_a)\eta_a} & x \in \langle a\rangle, \\ \frac{1}{2-\eta_a} & x \in \langle b\rangle, \end{cases}
$$

we get

$$
\begin{aligned}
s_g^{\langle a\rangle}(z) &= \frac{\int_{\langle a\rangle} g\,dx + z\eta_a \int_{\langle b\rangle} g\,dx}{(1-z)(1-(\eta_a-1)z)} \\
&= \frac{\eta_a \int g(x)\varrho(x)\,dx}{1-z} + \frac{(1-\eta_a)\int_{\langle a\rangle} g\,dx - \eta_a \int_{\langle b\rangle} g\,dx}{(1-(\eta_a-1)z)(2-\eta_a)}
\end{aligned}
$$

and

$$
s_g^{\langle b\rangle}(z) = \frac{(1-\eta_a)\int g(x)\varrho(x)\,dx}{1-z} + \frac{-\frac{1-\eta_a}{\eta_a}\int_{\langle a\rangle} g\,dx + (2-\eta_a)\int_{\langle b\rangle} g\,dx}{(1-(\eta_a-1)z)(2-\eta_a)}.
$$

We thus can imagine that $z = 1$ and $\eta_a - 1$ may be the eigenvalues of the Perron–Frobenius operator. As 1 is a singularity of order 1, 1 will be a simple eigenvalue of P, and there may be no other eigenvalue in the unit circle. This is what we hope to prove now in a more formal way.

Noticing $\int \varrho(x)\,dx = 1$, we denote by μ a probability measure of density function ϱ. The measure μ is an invariant probability measure for the transformation F.

Mixing. Returning to the definition of $s_g(z)$, we get by comparing the coefficient of z^n

$$
\begin{aligned}
\int 1_{\langle a\rangle}(x) g(F^n(x))\,dx &= \eta_a \int g\,d\mu + C_a(\eta_a-1)^n, \\
\int 1_{\langle b\rangle}(x) g(F^n(x))\,dx &= (1-\eta_a) \int g\,d\mu + C_b(\eta_a-1)^n,
\end{aligned}
$$

where

$$
\begin{aligned}
C_a &= \frac{(1-\eta_a)\int_{\langle a\rangle} g\,dx - \eta_a \int_{\langle b\rangle} g\,dx}{2-\eta_a}, \\
C_b &= \frac{-\frac{1-\eta_a}{\eta_a}\int_{\langle a\rangle} g\,dx + (2-\eta_a)\int_{\langle b\rangle} g\,dx}{2-\eta_a}.
\end{aligned}
$$

Therefore, for example, $\int 1_{\langle a\rangle}(x) g(F^n(x))\,dx$ converges to $\eta_a \int g\,d\mu$ exponentially as $n \to \infty$ with the rate $(\eta_a - 1)$. Moreover, since $\eta_a - 1 < 0$, it is oscillating. Note that the Lebesgue measure of $\langle a\rangle$ and $\langle b\rangle$ are η_a and $1-\eta_a$, respectively.

Consider now more generally for a word $W = w_1 \cdots w_m$

$$
\begin{aligned}
&\sum_{n=0}^{\infty} z^n \int 1_{\langle W\rangle}(x) g(F^n(x))\,dx \\
&= \sum_{n=0}^{m-1} z^n \int 1_{\langle W\rangle}(x) g(F^n(x))\,dx + \sum_{n=m}^{\infty} z^n \int 1_{\langle W\rangle}(x) g(F^n(x))\,dx \\
&= \sum_{n=0}^{m-1} z^n \int 1_{\langle W\rangle}(x) g(F^n(x))\,dx \\
&\quad + \begin{cases} z^m \eta_{w_1}\cdots\eta_{w_m}(s_g^{\langle a\rangle}(z) + s_g^{\langle b\rangle}(z)) & \text{if } w_m = a, \\ z^m \eta_{w_1}\cdots\eta_{w_m} s_g^{\langle a\rangle}(z) & \text{if } w_m = b. \end{cases}
\end{aligned}
$$

Take

$$
\begin{aligned}
\chi_g^W(z) &= \sum_{n=0}^{m-1} z^n \int 1_{\langle W\rangle}(x) g(F^n(x))\,dx, \\
\Phi^W(z) &= \begin{cases} \left(z^m \eta_{w_1}\cdots\eta_{w_m},\, z^m \eta_{a_1}\cdots\eta_{a_m}\right) & \text{if } a_m = a, \\ \left(z^m \eta_{w_1}\cdots\eta_{w_m},\, 0\right), & \text{if } w_m = b. \end{cases}
\end{aligned}
$$

Dividing f into the sum $f = \sum_w C_W 1_{\langle W\rangle}$, then

$$
\begin{aligned}
\sum_{n=0}^{\infty} z^n \int f(x) g(F^n(x))\,dx &= \sum_W C_W \sum_{n=0}^{\infty} z^n \int 1_{\langle W\rangle}(x) g(F^n(x))\,dx \\
&= \sum_W C_W \chi_g^W(z) + \sum_W C_W \Phi^W(z) s_g(z).
\end{aligned}
$$

Since $f \in BV$, there exists a decomposition $f = \sum_W C_W 1_{\langle W\rangle}$ such that

$$
\sum_{n=1}^{\infty} r^n \sum_{|W|=n} |C_W| < \infty
$$

for any $0 < r < 1$. This ensures the convergence of $\sum_W C_W \chi_g^W(z)$ and $\sum_W C_W \Phi^W(z)$ in $|z| < e^{\xi}$. We now prove this.

Suppose now $|z| < e^{\xi}$. Consider the eigenvalues of the Perron–Frobenius operator in BV which are greater than $e^{-\xi}$ in modulus. Let us prove that they coincide with the reciprocals of the solutions of $\det(I - \Phi(z)) = 0$.

Note that for any $\varepsilon > 0$, there exists a constant C such that

$$
\begin{aligned}
Lebes(\langle W\rangle) &\le Ce^{-(\xi-\varepsilon)n} \\
\eta_W = \eta_{w_1}\cdots\eta_{w_n} &\le Ce^{-(\xi-\varepsilon)n},
\end{aligned}
$$

where $Lebes$ is the Lebesgue measure. Hence,

$$|\sum_W C_W \chi_g^W(z)| = \left|\sum_{n=1}^{\infty}\sum_{|W|=n} |C_W|\chi_g^W(z)\right|$$
$$\leq \sum_{n=1}^{\infty}\sum_{|W|=n} |C_W| \sum_{m=0}^{n-1} |z|^m ||g||_\infty C e^{-(\xi-\varepsilon)n},$$
$$|\text{each component of } \sum_W \Phi^W(z)| \leq \sum_{n=1}^{\infty}\sum_{|W|=n} |C_W||z|^n C e^{-(\xi-\varepsilon)n}.$$

Therefore for $|z| < e^{\xi}$, $\sum_W C_W \chi_g^W(z)$ and every components of $\sum_W \Phi^W(z)$ are analytic.

On the other hand, $s_g(z)$ has singularities at $z = 1$ and $z = 1/(\eta_a - 1)$. Therefore, $\sum_{n=0}^{\infty} z^n \int f(x)g(F^n(x))\,dx$ can be extended meromorphically outside of the unit disk and it has at most two singularities in $|z| < e^{\xi}$. This shows that the reciprocals of the solutions of $\det(I - \Phi(z)) = 0$ coincide with the eigenvalues of the Perron–Frobenius operator which is grater than $e^{-\xi}$ in modulus. Namely, $\det(I - \Phi(z))$ plays the role of the Fredholm determinant for nuclear operators. This is the reason why we call $\Phi(z)$ the Fredholm matrix.

Comparing the coefficients as before,

$$\int f(x)g(F^n(x))\,dx \to \begin{cases} \sum_W C_W \eta_{w_1}\cdots\eta_{w_m} \int g\,d\mu & w_m = a \\ \sum_W C_W \eta_{w_1}\cdots\eta_{w_m}\eta_a \int g\,d\mu & w_m = b \end{cases}$$
$$= \int f\,dx \int g\,d\mu.$$

The order of convergence is proved to be $\min\{(\eta_a - 1), e^{-\xi}\}$. Hence taking $f \in BV$ and $g \in L^\infty$, we get

$$\int f(x)g(F^n(x))\,d\mu = \int f \cdot \varrho(x)g(F^n(x))\,dx$$
$$\to \int f \cdot \varrho\,dx \int g\,d\mu = \int f\,d\mu \int g\,d\mu.$$

We call this property of the dynamical system *mixing*, and $\min\{(\eta_a - 1), e^{-\xi}\}$ the *decay rate of correlation*.

Remark.

We can note now that the Fredholm matrix $\Phi(z)$ is essentially a 'weighted structure matrix'. Here the structure matrix of this dynamical system is $M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Namely, the first row expresses that $F(\langle a \rangle)$ contains both $\langle a \rangle$ and $\langle b \rangle$, and the second row expresses that $F(\langle b \rangle)$ contains only $\langle b \rangle$. The trace $\operatorname{tr} M = 1$ means $F(\langle a \rangle) \supset \langle a \rangle$, and this corresponds to the fixed point $x = 0$. At the same time, since $M^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, its trace equals 3. This means there exists a fixed point $a \to a \to a$ and another period 2 periodic orbit $a \to b \to a$ ($b \to a \to b$). In this way, the trace of the product of the structure matrix expresses a number of periodic orbits.

Ruelle's zeta function.

Similarly, the trace of $\Phi^n(z)$ also corresponds to periodic orbits with period n. Let us now define *Ruelle's zeta function*:

$$\zeta(z) = \exp\left[\sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{F^n(p)=p} |F^{n\prime}(p)^{-1}|\right].$$

Then

$$\begin{aligned}\zeta(z) &= \exp\left[\sum_{n=1}^{\infty} \frac{1}{n} \operatorname{tr} \Phi^n(z)\right] \\ &= \exp[-\operatorname{tr}\log(I - \Phi(z))] \\ &= (\det(I - \Phi(z))^{-1}.\end{aligned}$$

This shows that the reciprocals of the singularities of the zeta function become eigenvalues of the Perron–Frobenius operator.

The above discussions can be easily generalized to general piecewise linear Markov transformations. We can construct a Fredholm matrix $\Phi(z)$ by considering generating functions, and we get

Theorem 11.2.1. *In $|z| < e^{\xi}$, the reciprocals of the solutions of* $\det(I - \Phi(z)) = 0$ *coincide with the eigenvalues of the Perron–Frobenius operator restricted to BV which are greater than $e^{-\xi}$ in modulus. Moreover,* $\det(I - \Phi(z)) = 1/\zeta(z)$.

Thus, if the solutions of the equation $\det(I - \Phi(z)) = 0$ satisfy

1. 1 is simple,
2. there exists no others eigenvalues on the unit circle,

then there exists an invariant probability measure μ absolutely continuous with respect to the Lebesgue measure and the dynamical system $([0,1], \mu, F)$ is mixing.

11.2.2 β–expansion

As an example of a non-Markov transformation, let us study the β–transformations. Let $1 < \eta_a + \eta_b \leq 2$, and

$$\begin{aligned}&\mathcal{A} = \{a, b\}, \\ &\langle a\rangle = [0, \eta_a), \quad \langle b\rangle = [\eta_a, 1], \\ &F(x) = \begin{cases} x/\eta_a & x \in \langle a\rangle \\ (x - \eta_a)/\eta_b & x \in \langle b\rangle. \end{cases}\end{aligned}$$

The map F is called a *β–transformation* if $1/\eta_a = 1/\eta_b = \beta$. The above case is a generalization of a β–transformation.

The generating function $s_g^{\langle a\rangle}(z)$ satisfies the same renewal equation as before. If $F(\langle b\rangle) \subset \langle a\rangle$, we put $J_1 = F(\langle b\rangle)$. Next if $F(J_1) \supset \langle a\rangle$, we put $J_2 = F(J_1) \cap \langle b\rangle$. Then

$$\begin{aligned} s_g^{\langle b\rangle}(z) &= \int 1_{\langle b\rangle}(x)g(x)\,dx + \eta_b z s_g^{J_1}(z) \\ &= \int 1_{\langle b\rangle}(x)g(x)\,dx + \eta_b z\Big(\int 1_{J_1}(x)g(x)\,dx + z\eta_a\big(s_g^{\langle a\rangle}(z) + s_g^{J_2}(z)\big)\Big). \end{aligned}$$

Similarly, taking

$$J_n = \begin{cases} F(J_{n-1}) & \text{if } F(J_{n-1}) \subset \langle a\rangle, \\ F(J_{n-1}) \cap \langle b\rangle & \text{if } F(J_{n-1}) \supset \langle a\rangle, \end{cases}$$

and

$$\phi(n) = \begin{cases} 0 & \text{if } F(J_n) \subset \langle a\rangle \ (F^n 1 < \eta_a), \\ \prod_{k=1}^n \eta_{w_k^1} & \text{if } F(J_n) \supset \langle a\rangle \ (F^n 1 \geq \eta_a), \end{cases}$$

we get

$$s_g^{\langle b\rangle}(z) = \chi_g^b(z) + \sum_{n=1}^{\infty} z^n \phi(n) s_g^{\langle a\rangle}(z).$$

Here $W^1 = (w_n^1)_{n\in\mathbb{N}}$ is the expansion of 1 and

$$\chi_g^b(z) = \sum_{n=0}^{\infty} z^n \prod_{k=1}^{n} \eta_{w_k^1} \int_{J_n} g\,dx,$$

where we take $J_0 = \langle b\rangle$. Note that every components of $\chi_g(z)$ and $\Phi(z)$ are analytic in $|z| < e^{\xi}$. Therefore, putting

$$\Phi(z) = \begin{pmatrix} z\eta_a & z\eta_a \\ \sum_{n=1}^{\infty} z^n \phi(n) & 0 \end{pmatrix},$$

we get a renewal equation

$$s_g(z) = (I - \Phi(z))^{-1} \chi_g(z)$$

as before. Here $\chi_g(z) = \begin{pmatrix} \int_{\langle a\rangle} g\,dx \\ \chi_g^b(z) \end{pmatrix}$.

Through a similar discussion as in the Markov cases, we obtain that the reciprocals of the solutions of $\det(I - \Phi(z)) = 0$ coincide with the eigenvalues of the Perron–Frobenius operator which are greater than $e^{-\xi}$ in modulus. Using this expression, we can prove that the dynamical system is mixing, and we can also calculate the density function ϱ of the invariant probability measure μ absolutely continuous with respect to the Lebesgue measure, using

$$\lim_{z\uparrow 1}(1-z)(I-\Phi(z))^{-1}\begin{pmatrix}\int_{\langle a\rangle} g\,dx\\ \chi_g^b(1)\end{pmatrix}=\begin{pmatrix}\eta_a\int g\,d\mu\\ (1-\eta_a)\int g\,d\mu\end{pmatrix}.$$

Therefore,

$$\begin{pmatrix}\eta_a\int g\,d\mu\\ (1-\eta_a)\int g\,d\mu\end{pmatrix}$$
$$=\lim_{z\uparrow 1}\frac{1-z}{\det(I-\Phi(z))}\begin{pmatrix}1 & \eta_a\\ \sum_{n=1}^{\infty} z^n\phi(n) & 1-\eta_a\end{pmatrix}\begin{pmatrix}\int_{\langle a\rangle} g\,dx\\ \chi_g^b(1)\end{pmatrix}.$$

Hence,

$$\varrho(x)=C^{-1}\left[1_{\langle a\rangle}(x)+\eta_a\sum_{n=0}^{\infty}\prod_{k=1}^{n}\eta_{w_k^1}1_{J_n}(x)\right],$$

where C is the normalizing constant. In general, we can also prove that

$$\zeta(z)=\frac{1}{\det(I-\Phi(z))}.$$

We need unessential detailed discussion to prove this result (cf. [253]). Anyway, we get

$$\frac{1}{\zeta(z)}=1-z\eta_a\sum_{n=0}^{\infty}z^n\phi(n),$$

where $\phi(0)=1$. Using the above expression, we can calculate the spectrum of the Perron–Frobenius operator P. Thus by Theorem 11.2.1, we can determine the ergodicity of the dynamical system.

11.3 Cantor sets

The usual Cantor set is generated by the map $F(x)=3x \pmod 1$ by discarding the inverse images of $[\frac{1}{3},\frac{2}{3}]$. Namely, first we discard the interval $[\frac{1}{3},\frac{2}{3}]$. Next we discard its inverse images $[\frac{1}{9},\frac{2}{9}]$ and $[\frac{7}{9},\frac{8}{9}]$. Then we discard their inverse images $[\frac{1}{27},\frac{2}{27}]$ $[\frac{7}{27},\frac{8}{27}]$, $[\frac{19}{27},\frac{20}{27}]$ and $[\frac{25}{27},\frac{26}{27}]$, and continue this procedure.

We consider here sets obtained by an analogous device of construction and show how to compute their Hausdorff dimension by introducing the α–Fredholm matrix.

11.3.1 Hausdorff Dimension

The aim of this section is to measure the size of a Cantor set by computing its Hausdorff dimension.

Definition 11.3.1. *Let us consider a set $\mathcal{C} \subset [0,1]$. Let $\delta > 0$. A covering of $\mathcal{C}$ by a family of intervals which is at most countable $\{I_i\}$ ($\cup_i I_i \supset \mathcal{C}$), where the length of each I_i is less than δ is called a δ–*covering.

Now for $\alpha \geq 0$, put

$$\nu_\alpha(\mathcal{C}, \delta) = \inf \sum_i |I_i|^\alpha,$$

where the infimum is taken over all δ–coverings. Then, since $\nu_\alpha(\mathcal{C}, \delta)$ is monotone increasing as $\delta \downarrow 0$, the limit

$$\nu_\alpha(\mathcal{C}) := \lim_{\delta \downarrow 0} \nu_\alpha(\mathcal{C}, \delta)$$

exists.

Lemma 11.3.1. *Let us assume that $\nu_\alpha(\mathcal{C}) < \infty$ for $\alpha \geq 0$. Then for any $\alpha' > \alpha$, $\nu_{\alpha'}(\mathcal{C}) = 0$.*

Proof. Let $\{I_i\}$ be a δ–covering such that $\sum_i |I_i|^\alpha < \nu_\alpha(\mathcal{C}) + 1$. Then

$$\nu_{\alpha'}(\mathcal{C}) \leq \sum_i |I_i|^{\alpha'} \leq \delta^{\alpha'-\alpha} \sum_i |I_i|^\alpha \leq \delta^{\alpha'-\alpha}(\nu_\alpha(\mathcal{C}) + 1).$$

Take $\delta \downarrow 0$, then we get $\nu_{\alpha'}(\mathcal{C}) = 0$. ∎

The Hausdorff measure. We can construct a measure space over $(\mathcal{C}, \nu_\alpha)$ with a suitable σ–algebra. We will call ν_α a *Hausdorff measure* with coefficient α. However, as one can see from Lemma 11.3.1, this Hausdorff measure usually has no meaning. Namely, there exists $\alpha_0 \geq 0$ such that for any $\alpha < \alpha_0$ every set has measure infinity with respect to ν_α, and for any $\alpha > \alpha_0$ they have measure 0. We call this critical value α_0 the *Hausdorff dimension* of a set $\mathcal{C}$.

11.3.2 A Cantor set associated with a piecewise linear Markov transformation

Let

$$F(x) = \begin{cases} x/\eta_a & x \in [0, \eta_a), \\ (x - \eta_a)/\eta_b & x \in [\eta_a, \eta_a + \eta_b), \\ (x - \eta_a - \eta_b)/\eta_c & x \in [\eta_a + \eta_b, 1], \end{cases}$$

and $0 < \eta_a, \eta_b, \eta_c < 1$ and

$$\eta_a < (1 - \eta_a - \eta_b)/\eta_c < \eta_a + \eta_b.$$

Put $\lambda_a = \eta_a$, $\lambda_c = 1 - \eta_a - \eta_b$. These are the lengths of $\langle a \rangle$ and $\langle c \rangle$, respectively. Now, let us define the α–Fredholm matrix by

$$\Phi_\alpha(z) = \begin{array}{cc} & \begin{array}{cc} a & c \end{array} \\ \begin{array}{c} a \\ c \end{array} & \begin{pmatrix} z\eta_a^\alpha & z\eta_a^\alpha \\ z\eta_c^\alpha & 0 \end{pmatrix} \end{array}.$$

Let $\mathcal{C}$ be the Cantor set of points such that their positive orbit under the action of F never enters the interval $\langle b \rangle$.

As a rough discussion, we will calculate the Hausdorff dimension by covering the set $\mathcal{C}$ only by words with the same length.

As a first approximation, let us consider the covering by intervals $\langle a \rangle$ and $\langle c \rangle$. This can be expressed by

$$\eta_a^\alpha + \eta_c^\alpha = (1,1)\begin{pmatrix} \lambda_a^\alpha \\ \lambda_c^\alpha \end{pmatrix}.$$

The second approximation is the covering by intervals $\langle aa \rangle$, $\langle ac \rangle$ and $\langle ca \rangle$. The subintervals corresponding to the word cc does not exist. Therefore, the second approximation equals

$$(\eta_a\lambda_a)^\alpha + (\eta_a\lambda_c)^\alpha + (\eta_c\lambda_a)^\alpha = (1,1)\Phi_\alpha(1)\begin{pmatrix} \lambda_a^\alpha \\ \lambda_c^\alpha \end{pmatrix}.$$

Generally, the n–th approximation becomes

$$(1,1)\Phi_\alpha(1)^{n-1}\begin{pmatrix} \lambda_a^\alpha \\ \lambda_c^\alpha \end{pmatrix}.$$

If all the eigenvalues of the α–Fredholm matrix is less than 1 in modulus, $(1,1)\Phi_\alpha(1)^{n-1}\begin{pmatrix} \lambda_a^\alpha \\ \lambda_c^\alpha \end{pmatrix}$ converges to 0 as $n \to \infty$. This says that α is greater than the Hausdorff dimension of $\mathcal{C}$. On the contrary, if one of the eigenvalue is greater than 1 in modulus, then $(1,1)\Phi_\alpha(1)^{n-1}\begin{pmatrix} \lambda_a^\alpha \\ \lambda_c^\alpha \end{pmatrix}$ will diverge. Note that by Perron–Frobenius' theorem the maximal eigenvalue of $\Phi_\alpha(1)$ is positive. Thus, when α equals the Hausdorff dimension α_0 of $\mathcal{C}$, $\Phi_{\alpha_0}(1)$ has eigenvalue 1. Therefore α_0 will be the solution of $\det(I - \Phi_\alpha(1)) = 0$. Namely, α_0 will satisfy

$$\eta_a^{\alpha_0} + (\eta_a\eta_c)^{\alpha_0} = 1.$$

Take $\begin{pmatrix} \lambda_a^* \\ \lambda_c^* \end{pmatrix}$ as an eigenvector of $\Phi_{\alpha_0}(1)$ associated with the eigenvalue 1 with $\lambda_a^* + \lambda_c^* = 1$. To be more precise, $\lambda_a^* = \eta_a^{\alpha_0}$ and $\lambda_c^* = (\eta_a\eta_c)^{\alpha_0} = 1 - \eta_a^{\alpha_0}$. The other eigenvalue equals $(\eta_a^{\alpha_0} - 1)$ and $\begin{pmatrix} \eta_a^{\alpha_0} \\ -1 \end{pmatrix}$ is an eigenvector associated with it. Let us denote

$$\begin{pmatrix} \lambda_a^{\alpha_0} \\ \lambda_c^{\alpha_0} \end{pmatrix} = C_1\begin{pmatrix} \lambda_a^* \\ \lambda_c^* \end{pmatrix} + C_2\begin{pmatrix} \eta_a^{\alpha_0} \\ -1 \end{pmatrix},$$

where

$$C_1 = \frac{(1-\eta_a-\eta_c)^{\alpha_0}+1}{(1-\eta_a)^{\alpha_0}+1}$$
$$C_2 = \frac{(1-\eta_a)^{\alpha_0}-(1-\eta_a-\eta_c)^{\alpha_0}}{(1-\eta_a)^{\alpha_0}+1}.$$

This shows $0 < C_1 < 1$.

For a word $W = w_1 \cdots w_n$, $(w_i \in \{a, c\})$, its Hausdorff measure will be

$$\begin{aligned}\nu(\langle W\rangle) &= \lim_{m\to\infty} \eta_{w_1}^{\alpha_0}\cdots\eta_{w_{n-1}}^{\alpha_0} v(w_n)\Phi_{\alpha_0}(1)^m \begin{pmatrix}\lambda_a^{\alpha_0}\\ \lambda_c^{\alpha_0}\end{pmatrix}\\ &= C_1\eta_{w_1}^{\alpha_0}\cdots\eta_{w_{n-1}}^{\alpha_0}\lambda_{w_n}^{*},\end{aligned} \tag{11.1}$$

where $v(a) = (1, 0)$ and $v(c) = (0, 1)$. Hence, the total measure $0 < \nu(\mathcal{C}) = C_1 < 1$. The corresponding generating function is

$$s_g(z) = (I - \Phi_{\alpha_0}(z))^{-1}\chi_g = \frac{1}{1-\eta_a^{\alpha_0}z-(\eta_a\eta_c)^{\alpha_0}z^2}\begin{pmatrix}1 & z\eta_a^{\alpha_0}\\ z\eta_c^{\alpha_0} & 1-z\eta_a^{\alpha_0}\end{pmatrix}\chi_g,$$

where $\chi_g = \begin{pmatrix}\chi_g^a\\ \chi_g^c\end{pmatrix}$ and $\chi_g^d = \int_{\langle d\rangle} g\,d\nu$ $(d = a, c)$. Therefore,

$$\begin{aligned}s_g(z) &= \frac{1}{(1-z)(1-(1-\eta_a^{\alpha_0})z)}\\ &\quad\times\begin{pmatrix}\int_{\langle a\rangle} g\,d\nu + \eta_a^{\alpha_0}z\int_{\langle c\rangle} g\,d\nu\\ \eta_c^{\alpha_0}z\int_{\langle a\rangle} g\,d\nu + (1-\eta_a^{\alpha_0}z)\int_{\langle c\rangle} g\,d\nu\end{pmatrix}\\ &= \frac{1}{1-z}\left[\frac{1}{(2-\eta_a^{\alpha_0})\eta_a^{\alpha_0}}\int_{\langle a\rangle} g\,d\nu + \frac{1}{2-\eta_0^{\alpha_0}}\int_{\langle c\rangle} g\,d\nu\right]\\ &\quad\times\begin{pmatrix}\eta_a^{\alpha_0} + \text{small order}\\ 1-\eta_a^{\alpha_0} + \text{small order}\end{pmatrix}.\end{aligned} \tag{11.2}$$

Now put

$$\frac{d\mu}{d\nu}(x) = \begin{cases}\frac{1}{C_1(2-\eta_a^{\alpha_0})\eta_a^{\alpha_0}} & x \in \langle a\rangle,\\ \frac{1}{C_1(2-\eta_a^{\alpha_0})} & x \in \langle c\rangle.\end{cases}$$

Then, μ becomes an invariant probability measure absolutely continuous with respect to the Hausdorff measure. By Equation (11.2),

$$\int 1_{\langle a\rangle}(x)g(F^n(x))\,d\nu - \nu(\langle a\rangle)\int g\,d\mu \sim (1-\eta_a^{\alpha_0})^n.$$

The similar equation also holds for c. Therefore, this shows that the dynamical system is mixing.

We will construct another map $G : [0,1] \to [0,1]$. Let the lengths of $\langle a^* \rangle$ equal $\lambda_a^* = \eta_a^{\alpha_0}$, and the length of $\langle c^* \rangle$ equal $\lambda_c^* = 1 - \eta_a^{\alpha_0}$, corresponding to the eigenvector of $\Phi_{\alpha_0}(1)$ associated with eigenvalue 1, respectively. Take

$$G(x) = \begin{cases} \eta_a^{-\alpha_0} x & \text{if } x \in \langle a^* \rangle, \\ \eta_c^{-\alpha_0}(x - \lambda_a^*) & \text{if } x \in \langle c^* \rangle. \end{cases}$$

Note that $G(1) = \lambda_a^*$. Also in this case, the generating function equals

$$s_g(z\colon G) = (I - \Phi_{\alpha_0}(z))^{-1} \chi_g(G).$$

One may easily understand that the invariant probability measure derived from the above equation becomes Markov, and at the same time, the Fredholm matrix associated with G equals $\Phi_{\alpha_0}(z)$.

11.3.3 Rigorous Results

We will study now the above results rigorously. We can extend these results even for general piecewise linear cases by using signed symbolic dynamics. We will study in this book, only for Markov cases. Let $F : [0,1] \to [0,1]$ be a piecewise linear Markov transformation. We will consider a subset $\mathcal{A}_1 \subset \mathcal{A}$ and put

$$\mathcal{C} = \{x \in [0,1] \colon w_i^x \in \mathcal{A}_1 \text{ for all } i\}.$$

We want to study the Hausdorff dimension of $\mathcal{C}$ and the ergodic properties of this dynamical system.

The α–Fredholm matrix will be defined by

$$\Phi_\alpha(z)_{a,b} = \begin{cases} z\eta_a^\alpha & \text{if } \overline{F(\langle a \rangle)} \supset \langle b \rangle, \\ 0 & \text{otherwise.} \end{cases}$$

We assume moreover$F$ is expanding and α–Fredholm matrix is irreducible ($z \neq 0$).

Theorem 11.3.1. *Let α_0 be the maximal solution of* $\det(I - \Phi_\alpha(1)) = 0$. *Then α_0 is the Hausdorff dimension of $\mathcal{C}$.*

First, we will state trivial lemmas.

Lemma 11.3.2. *Let $0 < \alpha < 1$, and $x_i > 0$ $(1 \leq i \leq k)$. Then,*

$$k^{\alpha-1}(\sum_{i=1}^k x_i^\alpha) \leq (\sum_{i=1}^k x_i)^\alpha \leq \sum_{i=1}^k x_i^\alpha.$$

Lemma 11.3.3. *There exists a constant $K > 1$ such that*

$$\frac{1}{K} < \frac{Lebes(\langle Wa \rangle)}{Lebes(\langle Wb \rangle)} < K,$$

for any word W and any symbols $a, b \in \mathcal{A}$ such that $\langle Wa \rangle, \langle Wb \rangle \neq \emptyset$.

To prove Theorem 11.3.1, we need to define another Hausdorff dimension. Let μ be a probability measure on $[0, 1]$. Define

$$\mu_\alpha(\mathcal{C}, \delta) = \inf \sum_i \mu(\langle W_i \rangle)^\alpha,$$

where infimum is taken over all covering by words $\{W_i\}$ such that $\mu(\langle W_i \rangle) < \delta$. The difference between ν_α and μ_α lies at two points:

1. μ_α uses the probability measure μ, and ν_α uses the Lebesgue measure,
2. we only consider coverings by words to define μ, and we consider any coverings by intervals to define ν_α.

Put

$$\mu_\alpha(\mathcal{C}) = \lim_{\delta \to 0} \mu_\alpha(\mathcal{C}, \delta),$$

and denote the critical point whether $\mu_\alpha(\mathcal{C})$ converges or diverges by $\dim_\mu(\mathcal{C})$.

Lemma 11.3.4. *The Hausdorff dimension of* $\mathcal{C}$ *equals* $\dim_{Lebes}(\mathcal{C})$.

Proof. It is trivial that the Hausdorff dimension is less than or equal to $\dim_{Lebes}(\mathcal{C})$. We will show another inequality. Let $\{J_i\}$ be a covering by words such that $\sum (Lebes(J_i))^\alpha < M < \infty$. For each J_i, put

$$n_i = \min\{n \colon |W| = n, \langle W \rangle \subset J_i\}.$$

If there exist words $\langle W_1 \rangle, \dots, \langle W_k \rangle$ such that $|W_j| = n_i - 1$ and intersect with J_i, we will divide J_i into k subintervals $J_i \cap \langle W_1 \rangle, \dots, J_i \cap \langle W_k \rangle$. We denote this new covering by words again by $\{J_i\}$. Therefore we can assume that J_i is contained in some $\langle V_i \rangle$ with $|V_i| = n_i - 1$, and contains at least one $\langle V_i' \rangle$ with $|V_i'| = n_i$. Then by Lemma 11.3.2 and the assumption,

$$\sum_i (Lebes(J_i))^\alpha \leq \#\mathcal{A}^{1-\alpha} M < \#\mathcal{A} \cdot M.$$

Now we will choose a family of words $\{W_{i,j}\}$ such that $\cup_j \langle W_{i,j} \rangle = J_i$ as follows. First we will choose words $W_{i,1}, \dots, W_{i,j_1}$ with length n_i such that $\langle W_{i,j} \rangle \subset J_i$. Note that $j_1 \leq \#\mathcal{A}$. Next we will choose words $W_{i,j_1+1}, \dots,$ W_{i,j_2+1} with length $n_i + 1$ and $\langle W_{i,j} \rangle \subset J_i \setminus \cup_{j=1}^{j_1} \langle W_{i,j} \rangle$, and repeat this procedure. It can be noted that the number of words $W_{i,j}$ with length $n_i + k$ is at most $2\#\mathcal{A}$, and by Lemma 11.3.3,

$$\begin{aligned} Lebes(\langle W_{i,j} \rangle) &\leq e^{-k\xi} Lebes(\langle V_i \rangle) \\ &\leq e^{-k\xi} K \#\mathcal{A} \cdot \max\{Lebes(\langle W_{i,j} \rangle) \colon 1 \leq j \leq j_1\} \\ &\leq e^{-k\xi} K \#\mathcal{A} \cdot Lebes(J_i). \end{aligned}$$

Then

$$\begin{aligned}\sum_i \sum_j (Lebes(\langle W_{i,j}\rangle)^\alpha &\le \sum_i \sum_k \sum_{j:\ |W_{i,j}|=n_i+k} (Lebes(\langle W_{i,j}\rangle)^\alpha \\ &\le \sum_i \sum_k 2\#\mathcal{A}(e^{-k\xi}K\#\mathcal{A}\cdot Lebes(J_i))^\alpha \\ &\le \frac{2K^\alpha(\#\mathcal{A})^{1+\alpha}}{1-e^{-\xi\alpha}} \sum_i (Lebes(J_i))^\alpha \\ &\le K'M,\end{aligned}$$

where

$$K' = \frac{2K^\alpha(\#\mathcal{A})^{2+\alpha}}{1-e^{-\xi}}.$$

Now let us take any α which is greater than the Hausdorff dimension of $\mathcal{C}$. Then for any $\varepsilon > 0$, there exists a covering by intervals $\{J_i\}$ such that $\sum_i (Lebes(J_i))^\alpha < \varepsilon$. Then we can choose a covering by words $\{\langle W_{i,j}\rangle\}$ such that

$$\sum_i \sum_j (Lebes(\langle W_{i,j}\rangle))^\alpha < K'\varepsilon.$$

This shows that $\dim_{Lebes}(\mathcal{C}) < \alpha$. Thus $\dim_{Lebes}(\mathcal{C})$ is less than or equal to the Hausdorff dimension of $\mathcal{C}$. Namely, the Hausdorff dimension of $\mathcal{C}$ equals $\dim_{Lebes}(\mathcal{C})$. ∎

Now we will apply the following theorem.

Theorem 11.3.2 (Billingsley). *Let μ_1, μ_2 be probability measures. Assume that*

$$\mathcal{C} \subset \left\{x\colon \lim_{n\to\infty} \frac{\log \mu_1(\langle W^x[1,n]\rangle)}{\log \mu_2(\langle W^x[1,n]\rangle)} = \alpha\right\}$$

holds for some $0 \le \alpha \le \infty$, where $W^x[1,n]$ is a word with length n and $\langle W^x[1,n]\rangle \ni x$. Then

$$\dim_{\mu_2}(\mathcal{C}) = \alpha \dim_{\mu_1}(\mathcal{C}).$$

We will not prove this theorem here. This theorem can be found in [67] Theorem 14.1.

We will use the mapping $G\colon [0,1] \to [0,1]$. For the examples, we constructed it from an eigenvector of $\Phi_{\alpha_0}(1)$. We will construct it for general piecewise linear Markov transformations. Since $\Phi_\alpha(1)$ is irreducible, and $\Phi_{\alpha_0}(1)$ is a non–negative matrix, by Perron–Frobenius' theorem, the maximal eigenvalue is non–negative (according to the assumption, this equals 1), and we can choose a non–negative eigenvector associated with the maximal eigenvalue 1 for which the sum of its components equals 1. Let the measure of a^* $(a \in \mathcal{A}_1)$ be the component of the eigenvector. We arrange a^* as in the natural order of a, and the slope of G for $x \in \langle a^*\rangle$

$$G'(x) = \begin{cases} +\eta_a^{-\alpha_0} & \text{if } F'(y) > 0 \\ -\eta_a^{-\alpha_0} & \text{if } F'(y) < 0 \end{cases} \quad \text{for } y \in \langle a\rangle.$$

Then we can construct a mapping G which has the same symbolic dynamics with F restricted to $\mathcal{C}$. Hence the Fredholm matrix of G equals $\Phi_{\alpha_0}(z)$.

We denote by μ_G the induced measure from the Lebesgue measure over $[0,1]$ where G acts over $[0,1]$ where F acts. Then

$$\lim_{n\to\infty} \frac{\log Lebes(\langle W^x[1,n]\rangle)}{\log \mu_G(\langle W^x[1,n]\rangle)} = \frac{1}{\alpha_0}$$

holds for every $x \in \mathcal{C}$. Of course, $\mu_G(\mathcal{C}) = Lebes([0,1]) = 1$, i.e., $\dim_{\mu_G}(\mathcal{C}) = 1$. Therefore,

$$\dim_{Lebes}(\mathcal{C}) = \alpha_0 \dim_{\mu_G}(\mathcal{C}) = \alpha_0.$$

This shows that the Hausdorff dimension of $\mathcal{C}$ equals α_0.

The measure μ_G is called a *conformal measure* and using this we can prove the Hausdorff measure on $\mathcal{C}$ is absolutely continuous to μ_G, and nonzero finite[170]. Moreover, according to the the assumption made, G is expanding. Therefore it has an invariant probability measure μ absolutely continuous with respect to the Lebesgue measure, and the dynamical system is mixing. We denote by $\varrho(x) = \frac{d\mu}{dx}(x)$ the density of μ with respect to the Lebesgue measure. The induced measure $\hat{\mu}$ over $\mathcal{C}$ from this μ is absolutely continuous with respect to the Hausdorff measure, and the density function $\hat{\varrho}(y)$ of $\hat{\mu}$ with respect to the Hausdorff measure ν equals $\varrho(x)/\nu(\mathcal{C})$, where x and y has the same expansion with respect to G and F, respectively. This shows that the dynamical system with this measure is also mixing.

Acknowledgements. The author would like to thank Valérie Berthé for his careful reading of the original manuscript, and for his modifications to make it suitable as a chapter of this book.