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

We discussed in Chap. 7 the relationship between substitutive dynamical systems and shifts of finite type, in terms of Markov expanding maps on the interval. As an example, the Morse substitution and the expanding map $F(x) = 2x \pmod 1$ on $[0,1]$ have the same matrix $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (see Sec. 7.1.2).

Substitutions and expanding maps have other similarities. For instance, G. Rauzy gave a geometric realization of the Tribonacci substitution as an expanding dynamical system on a compact subset of the Euclidean plane, namely the Rauzy fractal (see Secs. 7.5 and 8.1.2). If β denotes the golden ratio, the map $F(x) = \beta x \pmod 1$ has the same symbolic dynamics on the Rauzy fractal as the action of the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

The aim of this chapter is to study ergodicity of expanding maps from a symbolic point of view, through the study of a few examples. Ergodic properties of dynamical systems were obtained in the preceding chapters by using the spectrum of the unitary operator U (let us recall that U is defined on (X, T, μ) by $U(f) = f \circ T$; see Sec. 1.4). In this chapter, we will focus on another operator, that is, the Perron-Frobenius operator, which has the same spectrum on the unit circle as the unitary operator. Thus, the spectrum of the Perron-Frobenius operator determines the ergodicity of the dynamical system. We will give an algorithm to compute this spectrum concretely in a few situations.

In [260], Lasota and Yorke focus their attention on expanding maps of the interval: they state a relationship between the ergodicity of these dynamical systems and the spectrum of their Perron–Frobenius operator. In Sec. 11.2.1 are summarized the definitions of the unitary operator and the Perron–Frobenius operator, and the relationship between the spectrum of these operators and the ergodicity of the dynamical systems. We do not claim here to give an exhaustive presentation of these topics. We rather chose to introduce these notions through the study of examples, in order to give an idea of these classical methods.

[1] This chapter has been written by M. Mori

Hofbauer and Keller [205] stated a relationship between the singularities of the Ruelle's zeta function and the spectrum of the Perron–Frobenius operator. Baladi and Ruelle [53] and Mori [300] proved similar results. However, their methods are quite different. In Sec. 11.2.2, along [300], is defined a *Fredholm determinant*, though the Perron–Frobenius operator is not compact. This definition involves generating functions and renewal equations. This Fredholm determinant allows us in Sec. 11.2.3 to determine explicitly the spectrum of the Perron–Frobenius operator as well as the Ruelle's zeta function. We deduce ergodic properties of the dynamical system such as ergodicity, mixing and rate of decay of correlations.

As an application, Sec. 11.4 is devoted to the computation of the Hausdorff dimension of some Cantor sets naturally associated with piecewise linear maps. For that purpose, we will define α–Fredholm matrices.

11.1 Definitions

Let F be a map from the unit interval $[0,1]$ into itself. Our main interest is to study the asymptotic behavior of orbits x, $F(x)$, $F^2(x)$, $\cdots$. Note that $F^n(x)$ denotes the n–th iteration and not the n–th power $(F(x))^n$ of F.

One–dimensional dynamical system. We focus on maps for which 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* $\varrho(x)$, denoted by $\varrho(x) = \frac{d\mu}{dx}(x)$, 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, that is, μ is an *invariant measure* with respect to the map F.

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

Class of 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).

We call a transformation F *piecewise linear* if the derivative F' is constant on each interval of monotonicity $\langle a\rangle$.

We call a transformation F *Markov* if the following holds:

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

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

We call a map F *expanding* if

$$\liminf_{n\to\infty} \frac{1}{n} \inf_{x\in[0,1]} \frac{1}{n} \log |(F^n)'(x)| > 0.$$

This coefficient of expansivity will be denoted in all that follows ξ.

Symbolic Dynamics. Let $\mathcal{A}$ be a finite set and S denotes the shift map on $\mathcal{A}^{\mathbb{N}}$. As defined in Chap. 1, a triple (X, S, ν) is called a *symbolic dynamical system* if X is a closed invariant subset of $\mathcal{A}^{\mathbb{N}}$ (i.e., $TX = X$) and ν an invariant probability measure.

The one–dimensional dynamical system 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 interval with respect to the partition into intervals of monotonicity. More precisely, let F denote a piecewise monotonic transformation. Let $\{\langle a\rangle\}_{a\in\mathcal{A}}$ be a partition of $[0,1]$ into subintervals of monotonicity, that is, 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 $x \in [0,1]$. The *expansion* $w^x = (w_n^x)_{n\in\mathbb{N}}$ of x is defined 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: $X := \overline{\{w^x | x \in [0,1]\}}$. Then the shift map S 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 $S(w^x)$.

Admissible words. Let $W = w_1 \cdots w_n$ be a finite word in $\mathcal{A}^\star$ and $|W| = n$ be the length of W. Let

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

be defined as the interval coded by the word W, i.e., $\forall x \in \langle W\rangle$, $x \in \langle w_1\rangle, \ldots, F^{n-1}(x) \in \langle w_n\rangle$. We say that a word W is *admissible* if $\langle W\rangle \neq \emptyset$.

Structure matrix. For a Markov map F, the following $\mathcal{A}\times\mathcal{A}$ matrix $\mathbf{M}$ plays an essential role, particularly when F is piecewise linear. Let for $a, b \in \mathcal{A}$, $\mathbf{M}_{a,b} = 1$ if $\overline{F(\langle a\rangle)} \supset \langle b\rangle$, $\mathbf{M}_{a,b} = 0$ otherwise. We call this matrix the *structure matrix* of F.

Functions of bounded variations. We denote by $\mathcal{L}^1$ and $\mathcal{L}^\infty$ 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 of *bounded variation* is one of the most important notions. We will denote it by BV.

Definition 11.1.1. *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 $(x_k)_{1\leq k\leq n}$ of $[0,1]$.*

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

11.2 Ergodic properties of piecewise linear Markov transformations

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

Let

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

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; second in statistical properties, for instance mixing. 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. 5 or [445]. The rate of convergence is called the *decay of correlations.*

Let us introduce some tools and notation, as the Perron–Frobenius operator, the generating functions or the Fredholm matrix.

11.2.1 The Perron–Frobenius operator

Definition 11.2.1. *The operator defined on* $\mathcal{L}^1$ *by:*

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

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

The name "Perron–Frobenius operator" is related to the fact that it has similar properties to matrices with positive entries. For more details on this operator, see for instance [93, 234, 335].

This operator corresponds to the change of variables $F(x) \to x$ in the following integration:

$$\int f(x)g(F(x))\,dx = \int Pf(x)g(x)\,dx \quad f \in \mathcal{L}^1,\ g \in \mathcal{L}^\infty.$$

However, restricting its domain to the set BV of bounded variation functions, the eigenvalues of this operator allow one to determine ergodic properties of the dynamical system [260]. There is no restriction here since eigenfunctions of expanding maps of the interval are shown to be BV [260, 301]. This implies that we can equivalently work here with $\mathcal{L}^1$ or with $\mathcal{L}^2$, as in the rest of this book. Indeed, the following properties are satisfied by the Perron–Frobenius operator

1. $P\colon \mathcal{L}^1 \to \mathcal{L}^1$ is positive,
2. the spectrum of P on the unit circle coincides with the spectrum of the unitary operator defined by $Uf = f \circ F$.

This implies the following:

1. The system admits 1 as an eigenvalue, and the eigenspace associated with the eigenvalue 1 is contained in BV. Moreover, there exists 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 [260];
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 strongly mixing [88], i.e., for $f \in \mathcal{L}^1$ and $g \in \mathcal{L}^\infty$
$$\int f(x) g(F^n(x)) \, dx \to \int f \, dx \int g \, d\mu,$$
where μ denotes the invariant probability measure absolutely continuous with respect to the Lebesgue measure. For more details on mixing, see Chap. 5.

As an operator from $\mathcal{L}^1$ into itself, every $|z| < 1$ is an eigenvalue of P with infinite multiplicity [240]. 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. However, restricting P to BV, every $|z| < e^{-\xi}$, (where ξ corresponds to the expansivity coefficient, i.e., $\xi = \liminf_{n\to\infty} \inf_{x\in[0,1]} \frac{1}{n} \log |(F^n)'(x)|$) is an eigenvalue with infinite multiplicity. Hence, the eigenvalues of P restricted to BV in the annulus $e^{-\xi} < |z| \le 1$ play an essential role to determine ergodic properties such as ergodicity, mixing, the rate of decay of convergence and so on.

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

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

11.2.2 Generating functions and Fredholm matrix

To study ergodic properties of a dynamical system, we need to determine the eigenvalues of the Perron-Frobenius operator. Similarly as in the case of matrices, the Fredholm determinant of a nuclear operator P is defined to be $\det(I - zP)$. It is an entire function, the zeros of which are the reciprocals of the eigenvalues of P. Though the Perron-Frobenius operator is not nuclear, the Fredholm determinant can be defined by constructing renewal equations. However, in general, it is not an entire function and has for natural boundary $|z| = e^{\xi}$.

The *generating functions* of any $g \in \mathcal{L}^\infty$, are the formal power series defined by

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

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 invariant probability measure μ, absolutely continuous with respect to the Lebesgue measure such that the dynamical system $([0,1], \mu, F)$ is strongly mixing.

Lemma 11.2.2. *Let* $\Phi(z) = \begin{pmatrix} z\eta_a & z\eta_a \\ z\eta_b & 0 \end{pmatrix}$ *(called a* Fredholm matrix*). 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,

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

11.2.3 Ergodic properties

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

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

where

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

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.

More generally, for a word $W = w_1 \cdots w_m$, let η_W denote the product $\eta_W = \eta_{w_1} \cdots \eta_{w_m}$. Consider

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

Take

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

By dividing f into the sum $f = \sum_W C_W 1_{\langle W\rangle}$, we get

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

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 $|z| < e^{\xi}$.

Suppose now that $|z| < e^{\xi}$. 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

$$Lebes(\langle W\rangle) \le Ce^{-(\xi-\varepsilon)n} \quad \text{and} \quad \eta_W = \eta_{w_1}\cdots\eta_{w_n} \le Ce^{-(\xi-\varepsilon)n},$$

where $Lebes$ is the Lebesgue measure. Hence,

$$\begin{aligned}|\sum_W C_W \chi_g^W(z)| &= \left|\sum_{n=1}^{\infty}\sum_{|W|=n} |C_W|\chi_g^W(z)\right| \\ &\le \sum_{n=1}^{\infty}\sum_{|W|=n} |C_W| \sum_{m=0}^{n-1} |z|^m ||g||_\infty Ce^{-(\xi-\varepsilon)n},\end{aligned}$$

$$|\text{each component of } \sum_W \Phi^W(z)| \le \sum_{n=1}^{\infty}\sum_{|W|=n} |C_W||z|^n Ce^{-(\xi-\varepsilon)n}.$$

Therefore for $|z| < e^{\xi}$, $\sum_W C_W \chi_g^W(z)$ and all the 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,

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

The order of convergence is proved to be $\min\{(\eta_a - 1), e^{-\xi}\}$. Hence taking $f \in BV$ and $g \in \mathcal{L}^\infty$, 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 *strong mixing*. The *decay rate of correlations* is $\min\{(\eta_a - 1), e^{-\xi}\}$.

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 $\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. 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 $\operatorname{tr} \mathbf{M} = 1$ means $F(\langle a \rangle) \supset \langle a \rangle$, and this corresponds to the fixed point $x = 0$. At the same time, since $\mathbf{M}^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, its trace equals 3. This means there exists a fixed point $a \to a \to a$ and another two-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

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

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

Final result. 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 11.2.3. *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 exist 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 strongly mixing.

11.3 Non-Markov transformation: β–expansion

As an example of a non-Markov transformation, let us study the β–transformations.

Let $1 < \eta_a + \eta_b \le 2$, $\mathcal{A} = \{a, b\}$, $\langle a\rangle = [0, \eta_a)$ and $\langle b\rangle = [\eta_a, 1]$. Let

$$F(x) = \begin{cases} x/\eta_a & x \in \langle a\rangle \\ (x-\eta_a)/\eta_b & x \in \langle b\rangle. \end{cases}$$

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

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

we get

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

where $w^1 = (w_n^1)_{n\in\mathbb{N}}$ is the expansion of 1 (following the notation of Sec. 11.1) and

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

with $J_0 = \langle b\rangle$. Note that all the 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 to that of the Markov case, 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 strongly 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 (see [300]). 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 11.2.3, we can determine the ergodicity of the dynamical system.

11.4 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.

11.4.1 Hausdorff Dimension

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

Definition 11.4.1. *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 11.4.2. *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 11.4.2, 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 $\mathcal{C}$.

11.4.2 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

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

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 do 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 are 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 tends to ∞. 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

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

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

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

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,

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

Now put

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

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

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

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

Let us 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 the 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 $\Phi_{\alpha_0}(z)$.

11.4.3 Rigorous Results

Let us state now the above results more precisely. We can extend these results even to general piecewise linear cases by using signed symbolic dynamics. We restrict ourselves to the Markov case in this chapter.

Let $F : [0,1] \to [0,1]$ be a piecewise linear Markov transformation. We are going to consider a subset $\mathcal{A}_1 \subset \mathcal{A}$ and put

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

The aim of this section is to study the Hausdorff dimension of $\mathcal{C}$ and the ergodic properties of this dynamical system.

The α–Fredholm matrix is defined by

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

Moreover we assume that F is expanding and the α–Fredholm matrix irreducible ($z \neq 0$).

Theorem 11.4.3. *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, let us state elementary lemmas.

Lemma 11.4.4. *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 11.4.5. *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 11.4.3, 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 the infimum is taken over all coverings by words $\{W_i\}$ such that $\mu(\langle W_i \rangle) < \delta$. The difference between ν_α and μ_α lies in two points:

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

Put

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

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

Lemma 11.4.6. *The Hausdorff dimension of* $\mathcal{C}$ *equals* $\dim_{Lebes}(\mathcal{C})$.

Proof. It is clear that the Hausdorff dimension is less than or equal to $\dim_{Lebes}(\mathcal{C})$. We will prove 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 $W_1, \ldots, W_k$ such that $|W_j| = n_i - 1$ and $\langle W_j \rangle$ intersects J_i, we will divide J_i into k subintervals $J_i \cap \langle W_1 \rangle, \ldots, J_i \cap \langle W_k \rangle$. We denote this new covering by words again by $\{J_i\}$. Therefore we can assume that J_i is contained in some $\langle W_i \rangle$ with $|W_i| = n_i - 1$, and contains at least one $\langle W_i' \rangle$ with $|W_i'| = n_i$. Then by Lemma 11.4.4 and the assumption,

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

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 $w_{i,1}, \ldots, W_{i,j_1}$ with length n_i such that $\langle W_{i,j}\rangle \subset J_i$. Note that $j_1 \le \mathrm{Card}\ \mathcal{A}$. Next we will choose words $W_{i,j_1+1}, \ldots,$ 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 $2\,\mathrm{Card}\ \mathcal{A}$, and by Lemma 11.4.5,

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

Then

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

where

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

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 $\dim_{Lebes}(\mathcal{C})$. ■

A proof of the following theorem can be found in [80] (Theorem 14.1).

Theorem 11.4.7 (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 $x \in \langle w^x[1,n]\rangle$. Then

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

In the examples studied in the former sections, the mapping $G\colon [0,1] \to [0,1]$ was built from an eigenvector of $\Phi_{\alpha_0}(1)$. We are going to 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 $x \in \langle a^* \rangle$

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

Then we can construct a mapping G which has the same symbolic dynamics as that of 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 that the Hausdorff measure on $\mathcal{C}$ is absolutely continuous with respect to μ_G, and is nonzero and finite [204]. 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 strongly mixing. We denote by $\varrho(x) = \frac{d\mu}{dx}(x)$ the density of μ with respect to the Lebesgue measure. The induced measure $\hat{\mu}$ on $\mathcal{C}$ (deduced from this measure 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 have respectively the same expansion with respect to G and F. This shows that the dynamical system equipped with this measure is also strongly mixing.