TOKYO J. MATH.
VOL. 28, NO. 1, 2005

On the Density Function of an Invariant Measure under One-Dimensional Bernoulli Transformations

Yuko ICHIKAWA

Tokyo National College of Technology

(Communicated by Y. Maeda)

Abstract. The continuity of the density function of the invariant probability measure for piecewise C^2 Bernoulli transformations is proved.

1. Introduction

The purpose of this article is to show that the continuity of the density function of an invariant probability measure for piecewise C^2, expanding, and Bernoulli transformations of the unit interval [0, 1]. We study Markov (not necessarily Bernoulli) transformations approximating them by piecewise linear transformations. To deal piecewise linear transformations as symbolic dynamics, Mori defined Fredholm matrices in [2] and [3]. First using these matrices, we construct a recurrent formula between the eigenvectors of them, and show the existence of an eigenfunction of Perron-Frobenius operator of original transformation which is the density function of the invariant probability measure. Secondly, we show that the density function is continuous for Bernoulli transformations. For piecewise C^2 and expanding transformation, Lasota and York have shown the existence of the invariant measure in [1]. The first part of this paper gives another proof of the Lasota and York's result for restricted cases.

2. Notations and Results

Let $F : [0, 1] \to [0, 1]$ satisfy the following conditions.

(C1): piecewise C^2. There exists a partition $0 = p_0 < p_1 < \cdots < p_r = 1$ of $[0, 1]$ such that the restriction of F to $(p_{i-1}, p_i), i = 1, 2, \cdots, r$ is C^2 and monotone function which can be extended to $[p_{i-1}, p_i]$ as a C^2 function. We call a set $\mathcal{A} = \{a_1, a_2, \cdots, a_r\}$ 'alphabet' and (p_{i-1}, p_i) is labeled by $\langle a_i \rangle$. Here $\#\mathcal{A} = r$ is finite.

Received August 16, 2002; revised February 21, 2005

(C2): transitive. For any $x, y \in [0,1]$, and for any neighborhoods $V(x)$, $V(y)$ of x and y, respectively, there exists $n \in \mathbf{N}$ such that

$$F^n(V(x)) \cap V(y) \neq \emptyset ,$$

here F^n is the n-th iteration of F.

(C3): expanding.

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

F is called 'Markov' if for any $a \in \mathcal{A}$, there exist letters $b_1, b_2, \cdots, b_k$ $(b_i \in \mathcal{A})$ such that

$$\overline{F(\langle a\rangle)} = \bigcup_{i=1}^{k} \overline{\langle b_i\rangle} , \tag{1}$$

where $\bar{J}$ stands for the closure of J. F is called 'Bernoulli', if F for any $a \in \mathcal{A}$, $\overline{F(\langle a\rangle)} = [0,1]$. Throughout this paper, we assume that F is Markov, and in section 5 we assume that F is Bernoulli.

To express F as symbolic dynamics, we prepare several notation. We call a finite sequence of letters $w = b_1 b_2 \cdots b_n$ $(b_k \in \mathcal{A})$ a word, and we define

$$\begin{aligned}
&|w| = n \quad \text{(the length of a word)} ,\\
&w[k] = b_k \quad \text{for } 1 \le k \le |w| \quad \text{(n-th coordinate)} ,\\
&w[k,l] = b_k b_{k+1} \cdots b_l \quad \text{for } 1 \le k < l \le |w| ,\\
&\langle w\rangle = \bigcap_{i=1}^{n} F^{-i+1}(\langle w[i]\rangle) ,\\
&h(w) = b_1 \cdots b_{n-1} ,\\
&t(w) = b_2 \cdots b_n .
\end{aligned}$$

We say a word w F-admissible if $\langle w\rangle \neq \emptyset$, and define the sets of F-admissible words as follows:

$$\begin{aligned}
W_n &= \{w \in \mathcal{A}^n : \ |w| = n, w \text{ is } F\text{-admissible}\} ,\\
\tilde{W}_n &= \bigcup_{k=0}^{n} W_k = \{w : \ w \text{ is } F\text{-admissible}, |w| \le n\} ,\\
W_\infty &= \{w \in \mathcal{A}^{\mathbf{N}} : \ w[1,n] \in W_n \text{ for all } n\} .
\end{aligned}$$

It is well known that there exists a unique invariant probability measure μ under F and the dynamical system $([0,1], \mu, F)$ is mixing, therefore it is ergodic. From the condition that F is expanding, for any $\varepsilon > 0$ there exists N_0 such that for any $N \ge N_0$ and for any $w \in W_N$,

$$\text{Lebes}(\langle w\rangle) \le e^{-(\xi-\varepsilon)N} , \tag{2}$$

where Lebes$(\langle w\rangle)$ denote the Lebesgue measure of $\langle w\rangle$.

Let us introduce orders among admissible words.

DEFINITION 1. For two F-admissible words w and w', we define $w < w'$ if one of the following holds:

1. $|w| < |w'|$
2. $|w| = |w'|$ and $x < y$ holds for all $x \in \langle w \rangle$ and $y \in \langle w' \rangle$.

The orders in W_N and W_∞ are introduced by the above definition.

Let $\mathcal{P}_N = \{\langle w \rangle : w \in W_N\}$. Then $\mathcal{P}_N$ gives a partition of [0, 1]. For $M < N$, $\mathcal{P}_N$ is a refinement of $\mathcal{P}_M$. On $\langle w \rangle \in \mathcal{P}_N$, we define a piecewise linear transformation F_N, whose graph is the segment from $(p_w^-, \lim_{x \downarrow p_w^-} F(x))$ to $(p_w^+, \lim_{x \uparrow p_w^+} F(x))$, where p_w^- and p_w^+ are the left and the right end points of $\langle w \rangle$, respectively. We call F_N the N-th approximation of F. Let $\eta_w = |(F_N|_{\langle w \rangle})'|^{-1}$. Here, we note that for $w \in W_N$, $\bigcap_{i=1}^{|w|} F_N^{-i+1}(\langle w[i] \rangle) = \langle w \rangle$. That is, for $w \in \tilde{W}_N$, $\langle w \rangle$ is equal under F and F_N.

Let $P : L^1 \to L^1$ be the Perron-Frobenius operator associated with F, that is, for $f \in L^1$,

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

and P_N be the one associated with F_N.

Operating P_1 to the indicator function $1_{\langle a \rangle}$ $(a \in \mathcal{A})$, from the Markov condition (1) we obtain

$$P_1 1_{\langle a \rangle}(x) = \eta_a \sum_{b:ab \in W_2} 1_{\langle b \rangle}(x).$$

In general, for $w \in W_N$,

$$P_N 1_{\langle w \rangle}(x) = \eta_w \sum_{b:wb \in W_{N+1}} 1_{\langle t(wb) \rangle}(x). \tag{3}$$

Let Φ_N be the Fredholm matrix for F_N, that is, Φ_N is a $W_N \times W_N$ matrix:

$$(\Phi_N)_{w,w'} = \begin{cases} \eta_w & t(w) = h(w'), \\ 0 & \text{otherwise}. \end{cases}$$

For a partition $\mathcal{P}_N$, let $\boldsymbol{i}_N$ and $|\boldsymbol{i}|_N$ be the vectors corresponding to words $w \in W_N$, whose components are $(\boldsymbol{i}_N)_w = 1_{\langle w \rangle}$, and $(|\boldsymbol{i}|_N)_w = \text{Lebes}(\langle w \rangle)$, respectively. Then the equations (3) can be written

$$P_N 1_{\langle w \rangle} = (\Phi_N \boldsymbol{i}_N)_w. \tag{4}$$

EXAMPLE 1. Let

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

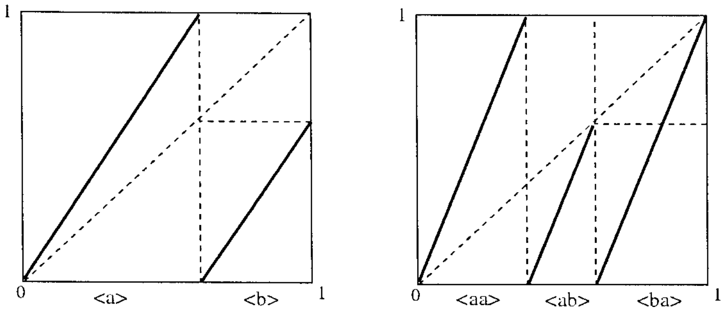

FIGURE 1. $F(x)$ and $F^2(x)$

Here, $\frac{1}{2} < \eta_a < 1$ and $1-\eta_a = \eta_a\eta_b$ holds. Fig. 1 shows the graphs of $F(x)$ and $F^2(x)$. Then $\mathcal{A} = \{a, b\}$, and $\langle a\rangle = (0, \eta_a)$, $\langle b\rangle = (\eta_a, 1)$. $W_2 = \{aa, ab, ba\}$. For this transformation, $\boldsymbol{i}_1 = \begin{pmatrix} 1_{\langle a\rangle} \\ 1_{\langle b\rangle} \end{pmatrix}$, $|\boldsymbol{i}|_1 = \begin{pmatrix} \eta_a \\ 1-\eta_a \end{pmatrix}$ and

$$\Phi_1 = \begin{pmatrix} \eta_a & \eta_a \\ \eta_b & 0 \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} \eta_{aa} & \eta_{aa} & 0 \\ 0 & 0 & \eta_{ab} \\ \eta_{ab} & \eta_{ab} & 0 \end{pmatrix}.$$

According to the property of Perron-Frobenius operator, it is well known that P_N is contractive, the eigenvalues of Φ_N are less than or equal to 1 in modules. A nonnegative eigenfunction of P associated with eigenvalue 1 is the density function of an invariant measure under F. Because Φ_N is nonnegative, then by the Perron-Frobenius' theorem, the maximal eigenvalue of Φ_N is simple, and its eigenvector can be taken that all the components are positive.

The Markov condition (1) can be expressed as $|\boldsymbol{i}|_N = \Phi_N|\boldsymbol{i}|_N$. This shows that Φ_N has an eigenvalue 1 and $|\boldsymbol{i}|_N$ is its eigenvector. Consequently, 1 is the maximal eigenvalue of Φ_N. Let $\boldsymbol{\rho}_N = (\rho_w)_{w\in W_N}$ be the eigenvector of Φ_N^* associated with eigenvalue 1 normalized in the sense $(\boldsymbol{\rho}_N, |\boldsymbol{i}|_N) = 1$. Here, A^* denotes the transpose matrix of A. We call the vector $|\boldsymbol{i}|_N$ *the interval vector of* F_N, and the vector $\boldsymbol{\rho}_N$ *the density vector of* F_N. Then we can express the density function of F_N-invariant measure with $\boldsymbol{\rho}_N$.

LEMMA 1. *Let $R_N(x) \equiv (\boldsymbol{\rho}_N, \boldsymbol{i}_N)(x) = \sum_{w\in W_N} \rho_w 1_{\langle w\rangle}(x)$. Then $R_N(x)$ is the density function of the invariant probability measure under F_N.*

PROOF. From (4)

$$\begin{aligned} P_N R_N(x) &= \sum_{w\in W_N} \rho_w P_N 1_{\langle w\rangle}(x) \\ &= (\Phi_N \boldsymbol{i}_N, \boldsymbol{\rho}_N)(x) \end{aligned}$$

$$= (\boldsymbol{i}_N, \Phi_N^* \boldsymbol{\rho}_N)(x)$$
$$= (\boldsymbol{i}_N, \boldsymbol{\rho}_N)(x) = R_N(x) .$$

This shows that $R_N(x)$ is an eigenfunction of P_N associated with eigenvalue 1. On the other hand, from the definition of ρ_N,

$$\int_{[0,1]} R_N(x)dx = \sum_{w \in W_N} \rho_w \int_{[0,1]} 1_{\langle w \rangle}(x)dx$$
$$= \sum_{w \in W_N} \rho_w (|\boldsymbol{i}|_N)_w$$
$$= (\boldsymbol{\rho}_N, |\boldsymbol{i}|_N) = 1 .$$

Thus the lemma is proved. □

The aim of this paper is to prove the following theorems.

THEOREM 1. *The limit function $R(x) = \lim_{N\to\infty} R_N(x)$ exists in L^1, and $R(x)$ is the density function of the F-invariant probability measure.*

THEOREM 2. *Suppose F is Bernoulli and $\xi > \frac{1}{2}\log r$, then $R(x)$ is continuous on* $[0, 1]$.

3. Framework

Before we proceed to the proof of Theorem 1, we need to examine several properties of F_N. Since the partition $\mathcal{P}_{N+1}$ is a refinement of $\mathcal{P}_N$, for an admissible word $w \in W_N$ $\langle w \rangle$ is the disjoint union of $\langle wa \rangle$, $wa \in W_{N+1}$. Then $1_{\langle w \rangle}(x) = \sum_{a: wa \in W_{N+1}} 1_{\langle wa \rangle}(x)$. For $wa \in W_N$, we get

$$P_N 1_{\langle wa \rangle}(x) = \eta_{h(wa)} \sum_{b \in \mathcal{A}} 1_{\langle t(wab) \rangle}(x)$$
$$= \eta_w \sum_{b \in \mathcal{A}} 1_{\langle t(wab) \rangle}(x) . \tag{5}$$

From (3) and (5), $P_{N+1} 1_{\langle w \rangle}$ $(w \in W_N)$ turns out to be

$$P_{N+1} 1_{\langle w \rangle}(x) = (P_N + (P_{N+1} - P_N)) 1_{\langle w \rangle}(x)$$
$$= P_N 1_{\langle w \rangle}(x) + (P_{N+1} - P_N) \sum_{a \in A} 1_{\langle wa \rangle}(x)$$
$$= \eta_w \sum_{a \in \mathcal{A}} 1_{\langle t(wa) \rangle}(x) + \sum_{a \in A} (\eta_{wa} - \eta_w) \sum_{b \in A} 1_{\langle t(wa)b \rangle}(x) .$$

Using this relation recursively, for $M < N$, and $w \in W_M$, we get

$$\begin{aligned} P_N 1_{\langle w \rangle}(x) &= \eta_w \sum_{b \in \mathcal{A}} 1_{\langle t(w) b \rangle}(x) \\ &\quad + \sum_{k=1}^{N-M} (\eta_{w b_1 b_2 \cdots b_k} - \eta_{w b_1 b_2 \cdots b_{k-1}}) \sum_{b \in A} 1_{t(w b_1 \cdots b_k) b}(x) \,. \end{aligned} \tag{6}$$

Let us rewrite this relation with a matrix, in the same way as (4).

DEFINITION 2. For words $w, w' \in \tilde{W}_N$ we say that w' is connectable to w if there exists an integer k $(0 < k < |w'|)$ such that $t(w) = h^k(w')$ and the connected word $w[1]w'$ is F-admissible.

Let $\tilde{\Phi}_N$ be a $\tilde{W}_N \times \tilde{W}_N$ matrix as

$$(\tilde{\Phi}_N)_{w,w'} = \begin{cases} \eta_w & t(w) = h(w') \,, \\ \eta_{w[1]h(w')} - \eta_{w[1]h^2(w')} & \text{if } |w| < |w'| \le N \text{ and } w' \text{ is connectable to } w \,, \\ 0 & \text{otherwise} \,, \end{cases}$$

and $\tilde{\boldsymbol{i}}_N$ be the vector of indicator functions similarly as $\boldsymbol{i}_N$, that is, $(\tilde{\boldsymbol{i}}_N)_w(x) = 1_{\langle w \rangle}(x)$, for $w \in \tilde{W}_N$. Take the example 1, we have

$$\tilde{\Phi}_2 = \begin{pmatrix} \eta_a & \eta_a & \eta_{aa} - \eta_a & \eta_{aa} - \eta_a & \eta_{ab} - \eta_a \\ \eta_b & 0 & \eta_{ba} - \eta_b & \eta_{ba} - \eta_b & 0 \\ 0 & 0 & \eta_{aa} & \eta_{aa} & 0 \\ 0 & 0 & 0 & 0 & \eta_{ab} \\ 0 & 0 & \eta_{ab} & \eta_{ab} & 0 \end{pmatrix}, \quad \tilde{\boldsymbol{i}}_2 = \begin{pmatrix} 1_{\langle a \rangle} \\ 1_{\langle b \rangle} \\ 1_{\langle aa \rangle} \\ 1_{\langle ab \rangle} \\ 1_{\langle ba \rangle} \end{pmatrix}.$$

Using $\tilde{\Phi}_N$, the equations (6) can be written by $P_N 1_{\langle w \rangle}(x) = (\tilde{\Phi}_N \tilde{\boldsymbol{i}}_N(x))_w$. The eigenvalues and eigenvectors of $\Phi_1, \Phi_2, \cdots, \Phi_N$ are related to the one of $\tilde{\Phi}_N$. Hence we shall be particularly interested in studying $\tilde{\Phi}_N$. To write this relation precisely, let us prepare the following matrices. For $k < l$, let $M_{k,l}$ be $W_k \times W_l$ matrix as

$$(M_{k,l})_{ww'} = \begin{cases} 1 & \text{if } w = h^{l-k}(w') \,, \\ 0 & \text{otherwise} \,. \end{cases}$$

$M_{k,l}$ expresses the Markov structure which is naturally induced from W_k to W_l. For $\boldsymbol{x}_k \in \mathbf{C}^{\# W_k}$ and $\boldsymbol{x}_l \in \mathbf{C}^{\# W_l}$, if $\boldsymbol{x}_k = M_{k,l} \boldsymbol{x}_l$ then $(\boldsymbol{x}_k)_w = \sum_{v : h^{l-k}(v) = w} (\boldsymbol{x}_l)_v$, and if $\boldsymbol{x}_l = M_{k,l}^* \boldsymbol{x}_k$ then $(\boldsymbol{x}_l)_w = (\boldsymbol{x}_k)_{h^{l-k}(w)}$. Let us divide $\tilde{\Phi}_N$ into the following blocks and define $D_{i,j}$ as

$W_i \times W_j$ matrix:

$$\tilde{\Phi}_N \quad = \quad \left(\begin{array}{c|c} & D_{1,N} \\ & D_{2,N} \\ \tilde{\Phi}_{N-1} & \vdots \\ & D_{N-1,N} \\ \hline 0 & \Phi_N \end{array}\right). \tag{7}$$

LEMMA 2.

$$D_{i,j} = \begin{cases} M_{i,j-1} D_{j-1,j} & \textit{if } 1 \le i < j-1\,, \\ M_{j-1,j}\Phi_j - \Phi_{j-1} M_{j-1,j} & \textit{if } i = j-1\,. \end{cases} \tag{8}$$

PROOF. By the definition of $\tilde{\Phi}_N$, we note that $(D_{j-1,j})_{w,w'} = (D_{j-2,j})_{h(w),w'} = (D_{i,j})_{h^{j-1-i}(w),w'}$. Then for $i < j-1$,

$$\begin{aligned}(M_{i,j-1} D_{j-1,j})_{w,w'} &= \sum_{v \in W_{j-1}} (M_{i,j-1})_{w,v} (D_{j-1,j})_{v,w'} \\ &= \sum_{v : w = h^{j-1-i}(v)} (D_{j-1,j})_{v,w'}\,.\end{aligned}$$

If w is connectable to w' then this value is equal to $(D_{j-1,j})_{w[1]h^2(w'),w'}$. In this case $h^{j-1-i}(w[1]h^2(w')) = w$. If w is not connectable to w' then it is equal to 0. For $i = j-1$, for $w \in W_{j-1}$ and $w' \in W_j$

$$\begin{aligned}&(M_{j-1,j}\Phi_j - \Phi_{j-1} M_{j-1,j})_{w,w'} \\ &\quad= \sum_{v \in W_j} (M_{j-1,j})_{w,v} (\Phi_j)_{v,w'} - \sum_{v \in W_{j-1}} (\Phi_{j-1})_{w,v} (M_{j-1,j})_{v,w'} \\ &\quad= \sum_{v : w = h(v)} (\Phi_j)_{v,w'} - \sum_{v : v = h(w')} (\Phi_{j-1})_{w,v} \\ &\quad= \sum_{v : w = h(v), t(v) = h(w')} \eta_{v[1]h(w')} - (\Phi_{j-1})_{w,h(w')} \\ &\quad= \begin{cases} \eta_{w[1]h(w')} - \eta_{w[1]h^2(w')} & \text{if } t(w) = h^2(w')\,, \\ 0 & \text{otherwise}\,. \end{cases}\end{aligned}$$

Therefore we get (8). □

In the remainder of this section, we will consider subspaces of $\mathbf{C}^{\#\tilde{W}_N}$ to which $\tilde{\Phi}_N$ operates. Let $\delta_{ww'}$ be

$$\delta_{ww'} = \begin{cases} 1 & w = w'\,, \\ 0 & w \neq w'\,, \end{cases}$$

and $\boldsymbol{e}_w = (\delta_{ww'})_{w' \in W_k}$ $(w \in W_k,\ k = 1, 2, \cdots, N)$. Then we easily see that

$$\boldsymbol{e}_{w[1,k]} = M_{k+1}\boldsymbol{e}_{w[1,k+1]} \quad (k = 1, 2, \cdots, |w| - 1)\,. \tag{9}$$

For $w \in W_k$, let $\tilde{\boldsymbol{e}}_w \in \mathbf{C}^{\#\tilde{W}_k}$ be a vector

$$\tilde{\boldsymbol{e}}_w = \begin{pmatrix} \boldsymbol{e}_{w[1]} \\ \boldsymbol{e}_{w[1,2]} \\ \vdots \\ \boldsymbol{e}_{w[1,k-1]} \\ \boldsymbol{e}_w \end{pmatrix}.$$

We identify $\tilde{\boldsymbol{e}}_w = \begin{pmatrix} \boldsymbol{e}_{w[1]} \\ \boldsymbol{e}_{w[1,2]} \\ \vdots \\ \boldsymbol{e}_{w[1,k-1]} \\ \boldsymbol{e}_w \end{pmatrix} \in \mathbf{C}^{\#\tilde{W}_k}$ with $\tilde{\boldsymbol{e}}_w = \begin{pmatrix} \boldsymbol{e}_{w[1]} \\ \boldsymbol{e}_{w[1,2]} \\ \vdots \\ \boldsymbol{e}_{w[1,k-1]} \\ \boldsymbol{e}_w \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \in \mathbf{C}^{\#\tilde{W}_N}$ for $k \leq N$.

PROPOSITION 1. 1. *The set* $\{\tilde{\boldsymbol{e}}_w :\ w \in \tilde{W}_N\}$ *forms a basis of* $\mathbf{C}^{\#\tilde{W}_N}$.

2. *Let* X_k *be a linear span of* $\{\tilde{\boldsymbol{e}}_w :\ w \in W_k\}$ *then* X_k *is* $\#W_k$ *dimensional subspace of* $\mathbf{C}^{\#W_k}$, *and* $\mathbf{C}^{\#\tilde{W}_N}$ *equals to the direct sum* $X_1 \oplus X_2 \oplus X_3 \oplus \cdots \oplus X_N$.

3. *For any* $\tilde{\boldsymbol{x}}_k \in X_k,\ k = 1, 2, \cdots, N$,

$$(\tilde{\boldsymbol{x}}_k)_w = \sum_{a:wa \in \tilde{W}_k} (\tilde{\boldsymbol{x}}_k)_{wa} \quad (|w| < k)\,. \tag{10}$$

4. X_k *is invariant under* $\tilde{\Phi}_N$. *The restriction* $\tilde{\Phi}_N|_{X_k}$ *is isomorphic to* Φ_N *on* $\mathbf{C}^{\#W_N}$.

PROOF. 1. The set of vectors $\{\tilde{\boldsymbol{\delta}}_w = (\delta_{ww'}) :\ w \in \tilde{W}_N\}$ becomes the natural basis of $\mathbf{C}^{\#\tilde{W}_N}$. The claim follows from $\tilde{\boldsymbol{\delta}}_w = \tilde{\boldsymbol{e}}_w - \tilde{\boldsymbol{e}}_{h(w)}$.

2. By (10) $\dim X_N$ is at most $\#W_N$. Moreover by the definition $\tilde{\boldsymbol{e}}_w$ $(w \in W_k)$ are linearly independent. Thus $\dim X_k = \#W_k$. Take $\boldsymbol{x} \in X_k \cap X_l$ $(k < l)$. Since $\boldsymbol{x} \in X_k$, $(\boldsymbol{x})_w$ is equal to 0 for $|w| > k$, particularly for $|w| = l$. On the other hand $\boldsymbol{x} \in W_l$, this leads to the conclusion $\boldsymbol{x} = \mathbf{0}$.

3. From the definition of $\tilde{\boldsymbol{e}}_w$, it is obvious.

4. From the definition of $\tilde{\Phi}_N$, $(\tilde{\Phi}_N)_{bh^l(w),h^k(w)} = \eta_{bh^{k+1}(w)} - \eta_{bh^{k+2}(w)}$, for $w \in W_N$ and $k < l < |w|$. Then

$$\tilde{\Phi}_N\tilde{\boldsymbol{e}}_w = \sum_{w:bh(w) \in W_N} \eta_{bh(w)}\tilde{\boldsymbol{e}}_{bh(w)}\,,$$

here $bh(w) \in W_N$, therefore $\tilde{\Phi}_N \tilde{\boldsymbol{e}}_w \in X_N$. □

Take $\boldsymbol{y}_k \in \mathbf{C}^{\# W_k}$ $(k = 1, 2, \cdots, N-1)$ arbitrary, and fix them. Let $\tilde{\boldsymbol{e}}_w^* = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 - M_{1,2}^* \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{e}_w - M_{N-1,N}^* \boldsymbol{y}_{N-1} \end{pmatrix}$ for $w \in W_N$. Then $\{\tilde{\boldsymbol{e}}_w^* : w \in W_N\}$ is the dual basis of $\{\tilde{\boldsymbol{e}}_w : w \in W_N\}$. Indeed, by (9),

$$
\begin{aligned}
(\tilde{\boldsymbol{e}}_u, \tilde{\boldsymbol{e}}_v^*) &= (\tilde{\boldsymbol{e}}_{u[1]}, \boldsymbol{y}_1) + \sum_{i=2}^{N-1} \{(\tilde{\boldsymbol{e}}_{u[1,i]}, \boldsymbol{y}_i) - (M_{i-1,i}^* \boldsymbol{y}_{i-1})\} \\
&\quad + (\boldsymbol{e}_u, \boldsymbol{e}_v) - (M_{N-1,N}^* \boldsymbol{y}_{N-1}) \\
&= (\boldsymbol{y}_1)_{u[1]} + \sum_{i=2}^{N-1} (\boldsymbol{y}_i - M_{i-1,i}^* \boldsymbol{y}_{i-1})_{u[1,i]} + (\boldsymbol{e}_v - M_{N-1,N}^* \boldsymbol{y}_{N-1})_u \\
&= (\boldsymbol{e}_v)_u = \delta_{uv}\,.
\end{aligned}
$$

Consequently X_N^*, the dual space of X_N, is the linear span of $\{\tilde{\boldsymbol{e}}_w^* : w \in W_N\}$. In the definition of $\tilde{\boldsymbol{e}}_w^*$, we can take $\boldsymbol{y}_1, \boldsymbol{y}_2 \cdots, \boldsymbol{y}_{N-1}$ arbitrarily. This means that $X_N^* \simeq \mathbf{C}^{\# \tilde{W}_N} \big/ \sim_N$. The relation $\sim_N$ is defined by

$$
\tilde{\boldsymbol{x}}_N = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_N \end{pmatrix} \sim_N \tilde{\boldsymbol{x}}_N' = \begin{pmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_N' \end{pmatrix} \Leftrightarrow \sum_{k=1}^{N} M_{k,N}^* \boldsymbol{x}_k = \sum_{k=1}^{N} M_{k,N}^* \boldsymbol{x}_k'\,.
$$

"If we rewrite $\tilde{\boldsymbol{x}}_N$ and $\tilde{\boldsymbol{x}}_N'$ to $\tilde{\boldsymbol{x}}_N = \begin{pmatrix} \xi_1 \\ \xi_2 - M_{1,2}^* \xi_1 \\ \vdots \\ \xi_N - M_{N-1,N} \xi_{N-1} \end{pmatrix}$, and $\tilde{\boldsymbol{x}}_N' = \begin{pmatrix} \xi_1' \\ \xi_2' - M_{1,2}^* \xi_1' \\ \vdots \\ \xi_N' - M_{N-1,N} \xi_{N-1}' \end{pmatrix}$, then this equivalent relation implies that $\xi_N = \xi_N'$.

PROPOSITION 2.

$$\tilde{\Phi}_N^* \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 - M_{1,2}^* \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_N - M_{N-1,N}^* \boldsymbol{y}_{N-1} \end{pmatrix} = \begin{pmatrix} \Phi_1^* \boldsymbol{y}_1 \\ \Phi_2^* \boldsymbol{y}_2 - M_{1,2}^* \Phi_1^* \boldsymbol{y}_1 \\ \vdots \\ \Phi_N^* \boldsymbol{y}_N - M_{N-1,N}^* \Phi_{N-1}^* \boldsymbol{y}_{N-1} \end{pmatrix} .$$

Especially,

$$\tilde{\boldsymbol{\rho}}_N = \begin{pmatrix} \boldsymbol{\rho}_1 \\ \boldsymbol{\rho}_2 - M_{1,2}^* \boldsymbol{\rho}_1 \\ \boldsymbol{\rho}_3 - M_{2,3}^* \boldsymbol{\rho}_2 \\ \vdots \\ \boldsymbol{\rho}_N - M_{N-1,N}^* \boldsymbol{\rho}_{N-1} \end{pmatrix}$$

is the eigenvector of $\tilde{\Phi}_N^*$ *associated with eigenvalue* 1, *where* $\boldsymbol{\rho}_k$ *is the density vector for* F_k $(k = 1, 2, \cdots, N)$.

PROOF. We get the proof by induction. Since $\Phi_1 = \tilde{\Phi}_1$, the claim is true for $N = 1$. Let

$$\tilde{\boldsymbol{y}}_k = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 - M_{1,2}^* \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_k - M_{k-1,k}^* \boldsymbol{y}_{k-1} \end{pmatrix} , \tag{11}$$

and assume that the claim is true for $k - 1$. Then for $|w| < k$, by the formula of (7), we get

$$(\tilde{\Phi}_k^* \tilde{\boldsymbol{y}}_k)_w = (\tilde{\Phi}_{k-1}^* \tilde{\boldsymbol{y}}_{k-1})_w .$$

By Lemma 2,

$$\begin{aligned}
(\tilde{\Phi}_k^* \tilde{\boldsymbol{y}}_k)_w &= ((D_{1,k}^* \cdots D_{k-1,k}^*) \tilde{\boldsymbol{y}}_{k-1} + \Phi_k^* (\boldsymbol{y}_k - M_{k-1,k}^* \boldsymbol{y}_{k-1}))_w \\
&= (D_{1,k}^* \boldsymbol{y}_1 + D_{2,k}^* (\boldsymbol{y}_2 - M_{1,2}^* \boldsymbol{y}_1) + \cdots \\
&\quad + (D_{k-1,k} (\boldsymbol{y}_{k-1} - M_{k-2,k-1} \boldsymbol{y}_{k-2}) + \Phi_k^* (\boldsymbol{y}_k - M_{k-1,k}^* \boldsymbol{y}_{k-1}))_w \\
&= (D_{k-1,k}^* \boldsymbol{y}_{k-1} + \Phi_k^* (\boldsymbol{y}_k - M_{k-1,k}^* \boldsymbol{y}_{k-1}))_w \qquad (12) \\
&= ((\Phi_k^* M_{k-1,k}^* - M_{k-1,k}^* \Phi_{k-1}^*) \boldsymbol{y}_{k-1} + \Phi_k^* (\boldsymbol{y}_k - M_{k-1,k}^* \boldsymbol{y}_{k-1}))_w \\
&= (\Phi_k^* \boldsymbol{y}_k - M_{k-1,k}^* \Phi_{k-1}^* \boldsymbol{y}_{k-1})_w . \qquad (13)
\end{aligned}$$

Hence when $|w| = k$, we get the conclusion. □

In the above calculation, we get the following equation:

LEMMA 3.

$$(E_N - \Phi_N^*)(\boldsymbol{\rho}_N - M_{N-1,N}^*\boldsymbol{\rho}_{N-1}) = D_{N-1,N}^*\boldsymbol{\rho}_{N-1}\,,$$

where E_N is the $\# W_N$ dimensional identity matrix.

PROOF. In (12) and (13), for $k = N$, we substitute $\boldsymbol{\rho}_{N-1}$ and $\boldsymbol{\rho}_N$ for $\boldsymbol{y}_{N-1}$ and $\boldsymbol{y}_N$, respectively. This leads to the conclusion. □

By Proposition 2, to take $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \cdots, \boldsymbol{\rho}_{N-1}$ as $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{N-1}$, in (11) we can identify $\tilde{\Phi}_N$ on X_N^* with Φ_N on $\mathbf{C}^{\# W_N}$.

Now, we define norms of $\boldsymbol{x}_N = (x_w)_{w\in\tilde{W}_N} \in X_N$ as follows:

$$\|\boldsymbol{x}_N\|_N \equiv \sup_{1\le k\le N} \sum_{w\in W_k} |x_w| = \sum_{w\in W_N} |x_w|\,.$$

Then norms of $\boldsymbol{x}_N^* = (x_w^*) \in X_N^*$ are induced by

$$\|\boldsymbol{x}_N^*\|_N^* = \sup_{z_N\in X_N, \|z\|_N=1} |(z_N, \boldsymbol{x}_N^*)| = \sup_{w\in W_N} |x_w^*|\,.$$

4. Proof of Theorem 1

We can now proceed to the proof of Theorem 1. This will require some additional preliminary lemmas. Let us decompose $\mathbf{C}^{\# W_N}$ into the generalized eigenspace of Φ_N^*. Let $\lambda_i\,(i = 1, 2, \cdots, s)$ be the eigenvalues of Φ_N^*. Since 1 is an eigenvalue of Φ_N^*, we set $\lambda_1 = 1$. Set $G_N^i = \{\boldsymbol{x} \in \mathbf{C}^{\# W_N} : (\Phi_N^* - \lambda_i E_N)^{k_i}\boldsymbol{x} = \boldsymbol{0}\}$, then $\mathbf{C}^{\# W_N} = G_N^1 \oplus G_N^2 \oplus \cdots \oplus G_N^s$, here k_i is the index of λ_i, and E_N is the $\# W_N$ dimensional identity matrix. Note that G_N^1 is the eigenspace associated with eigenvalue 1. Since 1 is simple, $\dim G_N^1 = 1$. Let us denote $G_N^2 \oplus G_N^3 \oplus \cdots \oplus G_N^s$ by $\bar{G}_N$, that is, $\mathbf{C}^{\# W_N} = G_N^1 \oplus \bar{G}_N$.

LEMMA 4. $\boldsymbol{\rho}_N - M_{N-1,N}^*\boldsymbol{\rho}_{N-1}$ *belongs to* $\bar{G}_N$.

PROOF. Let us decompose the vector $\boldsymbol{\rho}_N - M_{N-1,N}^*\boldsymbol{\rho}_{N-1} = x\boldsymbol{\rho}_N + \boldsymbol{v}$, where $\boldsymbol{v} \in \bar{G}_N$. Then

$$\begin{aligned}
&(|\boldsymbol{i}|_N, \boldsymbol{\rho}_N - M_{N-1,N}^*\boldsymbol{\rho}_{N-1})\\
&\quad= (\Phi_N^j|\boldsymbol{i}|_N, \boldsymbol{\rho}_N - M_{N-1,N}^*\boldsymbol{\rho}_{N-1})\\
&\quad= (\Phi_N^j|\boldsymbol{i}|_N, x\boldsymbol{\rho}_N + \boldsymbol{v}) = (|\boldsymbol{i}|_N, (\Phi_N^*)^j(x\boldsymbol{\rho}_N + \boldsymbol{v}))\\
&\quad= (|\boldsymbol{i}|_N, (\Phi_N^*)^j x\boldsymbol{\rho}) + (|\boldsymbol{i}|_N, (\Phi_N^*)^j\boldsymbol{v}) = x(|\boldsymbol{i}|_N, \boldsymbol{\rho}) + (|\boldsymbol{i}|_N, (\Phi_N^*)^j\boldsymbol{v})\,.
\end{aligned}$$

Since $\bar{G}_N$, Φ_N^* is strictly contractive on $\bar{G}_N$, $(\Phi_N^*)^j \boldsymbol{v}$ converges to $\mathbf{0}$ as $j \to \infty$. On the other hand, by the definition of $\boldsymbol{\rho}_k$, $(\boldsymbol{\rho}, |\boldsymbol{i}|_k) = 1$, then

$$
\begin{aligned}
(|\boldsymbol{i}|_k, \boldsymbol{\rho}_k - M_{k-1,k}^* \boldsymbol{\rho}_{k-1}) &= \sum_{wa \in W_k} (\rho_{wa} - \rho_w) |\langle wa \rangle| \\
&= \sum_{wa \in W_k} \rho_{wa} |\langle wa \rangle| - \sum_{w \in W_{k-1}} \rho_w \sum_{a : wa \in W_k} |\langle wa \rangle| \\
&= \sum_{wa \in W_k} \rho_{wa} |\langle wa \rangle| - \sum_{w \in W_{k-1}} \rho_w |\langle w \rangle| \\
&= (\boldsymbol{\rho}_k, |\boldsymbol{i}|_k) - (\boldsymbol{\rho}_{k-1}, |\boldsymbol{i}|_{k-1}) \\
&= 0 \,.
\end{aligned}
$$

Consequently, $x = 0$ therefore $\boldsymbol{\rho}_N - M_{N-1,N}^* \boldsymbol{\rho}_{N-1}$ belongs to $\bar{G}_N$. □

The next lemma has a crucial role in the proof of Theorem 1.

LEMMA 5. *For $w \in W_\infty$, the sequence $\{(\boldsymbol{\rho}_N)_{w[1,N]}\}$ converges uniformly in W_∞ as $N \to \infty$.*

PROOF. For simplicity, we write $(\boldsymbol{\rho}_N)_w$ instead of $(\boldsymbol{\rho}_N)_{w[1,N]}$. By Lemma 4, $\boldsymbol{\rho}_N - M_{N-1,N}^* \boldsymbol{\rho}_{N-1}$ belongs to $\bar{G}_N$, so $E_N - \Phi_N^*$ is invertible on $\bar{G}_N$. Put $\Psi_N = (E_N - \Phi_N^*)|_{\bar{G}_N}^{-1}$. Then by Lemma 3

$$
\boldsymbol{\rho}_N - M_{N-1,N}^* \boldsymbol{\rho}_{N-1} = \Psi_N D_{N-1,N}^* \boldsymbol{\rho}_{N-1} \,. \tag{14}
$$

Therefore,

$$
\begin{aligned}
\boldsymbol{\rho}_N &= (M_{N-1,N}^* + \Psi_N D_{N-1,N}^*) \boldsymbol{\rho}_{N-1} \\
&= (M_{N-1,N}^* + \Psi_N D_{N-1,N}^*)(M_{N-2,N-1}^* + \Psi_{N-1} D_{N-2,N-1}^*) \boldsymbol{\rho}_{N-2} \\
&= (M_{N-1,N}^* + \Psi_N D_{N-1,N}^*) \cdots (M_{1,2}^* + \Psi_2 D_{1,2}^*) \boldsymbol{\rho}_1 \,.
\end{aligned} \tag{15}
$$

On the other hand, operator norm of $D_{N,N-1}^*$ is evaluated as follows:

$$
\begin{aligned}
\|D_{N-1,N}^*\| &= \sup_{\boldsymbol{x}^* \in X_N^*, \|\boldsymbol{x}^*\|_N = 1} \|D_{N,N-1}^* \boldsymbol{x}^*\|_N^* \\
&= \sup_{w \in W_{N-1}} \sum_{b \in \mathcal{A}, bw \in W_N} |\eta_{bw} - \eta_{h(bw)}| \\
&\le r \max_{w \in W_N} |\eta_w - \eta_{h(w)}| \\
&\le r \max_{x,y \in \langle w \rangle} \left| \frac{1}{|F'(x)|} - \frac{1}{|F'(y)|} \right| \\
&\le r \max_{c \in \langle w \rangle} \mathrm{Lebes}(\langle w \rangle) \left| \frac{F''(c)}{(F'(c))^2} \right| \,.
\end{aligned}
$$

Here recall $r = \#\mathcal{A} < \infty$. Therefore from (2), for $N \geq N_0$, we get

$$\|D^*_{N-1,N}\| \leq K_0 e^{-(\xi-\varepsilon)N} ,$$

where $K_0 = r \cdot \max_{x\in[0,1]} \left|\frac{F''(x)}{(F'(x))^2}\right|$. On $\bar{G}_N$, Φ^*_N is strictly contractive and

$$(E_N - \Phi^*_N)|^{-1}_{\bar{G}_N} = \sum_{n\geq 0} (\Phi^*_N|_{\bar{G}_N})^n ,$$

and, since $\Psi_N : \bar{G}_N \to \bar{G}_N$,

$$\|\Psi_N\| = \|(E_N - \Phi_N)|^{-1}_{\bar{G}_N}\| \leq \frac{1}{1 - \|\Phi^*_N|_{\bar{G}_N}\|} .$$

Note that the eigenvalues of Φ_N converge to the eigenvalues of the Perron-Frobenius operator P restricted to the set of functions with bounded variation ([4]). This says that there exists $\delta > 0$ such that for sufficiently large N

$$\|\Psi_N\| \leq \frac{1}{1-\delta} < \infty .$$

Moreover, there is just one 1 on each column of $M_{N-1,N}$, so

$$\|M^*_{N-1,N}\| = \sup_{\|\boldsymbol{x}^*\|^*_N = 1} |M^*_{N-1,N}\boldsymbol{x}^*| = 1 .$$

Then by (15) and (2),

$$\begin{aligned}
\|\boldsymbol{\rho}_N\|_N &\leq \|M^*_{N-1,N} + \Psi_N D^*_{N-1,N}\|\|M^*_{N-2,N-1} + \Psi_{N-1} D^*_{N-2,N-1}\| \\
&\qquad \cdots \|M^*_{1,2} + \Psi_2 D^*_{1,2}\|\|\boldsymbol{\rho}_1\|_1 \\
&\leq (\|M^*_{N-1,N}\| + \|\Psi_N\|\|D^*_{N-1,N}\|)(\|M^*_{N-2,N-1}\| + \|\Psi_{N-1}\|\|D^*_{N-2,N-1}\|) \\
&\qquad \cdots (\|M^*_{1,2}\| + \|\Psi_2\|\|D^*_{1,2}\|)\|\boldsymbol{\rho}_1\|_1 \\
&\leq (1 + K_1\|\|D^*_{N-1,N}\|)(1 + K_1\|D^*_{N-2,N-1}\|) \cdots (1 + K_1\|D^*_{1,2}\|)\|\boldsymbol{\rho}_1\|_1 \\
&\leq K_2 \prod_{j=N_0}^{N} (1 + K_1 e^{-(\xi-\varepsilon)j}) ,
\end{aligned}$$

where $K_1 = \frac{1}{1-\delta}$, and $K_2 = (\prod_{j=1}^{N_0-1}(1 + K_1\|D_j\|))\|\boldsymbol{\rho}_1\|_1$. We can take ε such that $0 < \varepsilon < \xi$, then $\sum_{j=N_0}^{\infty} K_1 e^{-(\xi-\varepsilon)j} < \infty$. By the convergence of the infinite product, $\|\boldsymbol{\rho}_N\|_N$ is bounded. Then for $m > n > N_0$, using (14) again,

$$|(\boldsymbol{\rho}_m)_w - (\boldsymbol{\rho}_n)_w| \leq \sum_{k=n}^{m-1} |((\boldsymbol{\rho}_{k+1})_w - (\boldsymbol{\rho}_k)_w)|$$

$$
\begin{aligned}
&\le \sum_{k=n}^{m-1} \|\boldsymbol{\rho}_{k+1} - M^*_{k,k+1}\boldsymbol{\rho}_k\|_{k+1} \\
&= \sum_{k=n}^{m-1} \|\Psi_{k+1} D^*_{k,k+1}\boldsymbol{\rho}_k\|_{k+1} \\
&\le K_3 \sum_{k=n}^{m-1} e^{-(\xi-\varepsilon)k} \\
&= K_3 e^{-(\xi-\varepsilon)n} \sum_{k=0}^{m-n-1} e^{-(\xi-\varepsilon)k}\,,
\end{aligned}
$$

where $K_3 = \frac{1}{1-\delta}\sup_N \|\boldsymbol{\rho}_N\|_N$. We can take this term arbitrarily small for large enough m and n. So the sequence $\{(\boldsymbol{\rho}_N)_{w[1,N]}\}$ is a Cauchy sequence and converges uniformly on W_∞. □

To prove Theorem 1, we need one more lemma which is proved in [4].

LEMMA 6. *$\|P - P_N\| \to 0$ in $L^1[0, 1]$.*

PROOF OF THE THEOREM 1. First we will show that $\{R_N(x)\}_{N\in\mathbf{N}}$ is a Cauchy sequence in $L^1[0, 1]$. By Lemma 1, for $M > N$,

$$
\begin{aligned}
\|R_M - R_N\|_{L_1} &= \sum_{w\in W_M} \int_{\langle w\rangle} |(\boldsymbol{\rho}_M)_w - (\boldsymbol{\rho}_N)_{h(w)}|dx \\
&= \sum_{w\in W_M} \mathrm{Lebes}(\langle w\rangle)|(\boldsymbol{\rho}_M)_w - (\boldsymbol{\rho}_N)_{h(w)}| \\
&\le \max_{w\in W_M} |(\boldsymbol{\rho}_M)_w - (\boldsymbol{\rho}_N)_{h^{M-N}(w)}|\,.
\end{aligned}
$$

By Lemma 5, this converges to 0 as $M, N \to \infty$. So let $R(x)$ be the limit function of $\{R_N(x)\}$. Now, we show that $R(x)$ is an eigenfunction of P associated with the eigenvalue 1.

$$
\|PR - P_N R_N\| \le \|P\|\|R - R_N\| + \|P - P_N\|\|R_N\|\,. \tag{16}
$$

Here by the definition of $R(x)$ and by Lemma 6, if N is large enough, then we can make the right hand side of (16) arbitrarily small. Thus for any $\varepsilon > 0$, we can take large enough N such that

$$
\begin{aligned}
\|PR - R\| &\le \|PR - P_N R_N\| + \|P_N R_N - R_N\| + \|R_N - R\| \\
&< \varepsilon\,.
\end{aligned}
$$

Therefore, $R(x)$ is an eigenfunction of P associated with eigenvalue 1. □

5. Proof of Theorem 2—Bernoulli case

We will give now the proof of Theorem 2 with a direct estimation of the value $R_N(x)$. First we prepare the next lemma.

LEMMA 7. *Let $\boldsymbol{\rho}_N$ be the density vector for F_N. For $w, w' \in W_N$, if $h(w) = h(w')$ then $(\boldsymbol{\rho}_N)_w = (\boldsymbol{\rho}_N)_{w'}$.*

PROOF. Denote $(\Phi_N)_w$ the w'th column of Φ_N. From the definition of Φ_N, if $h(w) = h(w')$ then $(\Phi_N)_w = (\Phi_N)_{w'}$. Since $\Phi_N^* \boldsymbol{\rho}_N = \boldsymbol{\rho}_N$,

$$(\boldsymbol{\rho}_N)_w = ((\Phi_N^*)_w, \boldsymbol{\rho}_N) = ((\Phi_N^*)_{w'}, \boldsymbol{\rho}_N) = (\boldsymbol{\rho}_N)_{w'} . \qquad \square$$

From now, we assume that F is Bernoulli, then all the words that belong to $\mathcal{A}^N$ are F-admissible.

Now we are ready to prove Theorem 2.

PROOF OF THE THEOREM 2. Let $R_N(x)$ be the density function of the F_N-invariant probability measure. Then by Lemma 1 $R_N(x) = (\boldsymbol{\rho}_N, \boldsymbol{i}_N)(x)$. Note that $R_N(x)$ is constant on $\langle w \rangle$ for $w \in W_N$. For simplicity, we denote $(\boldsymbol{\rho}_N)_w = \rho_w$ for $w \in W_N$. As is well known, the Lebesgue measure is invariant with respect to F_1, so $R_1(x) \equiv 1$. Now we take F_2. Then for words $b_1 b_2$ and $c_1 c_2$, if $b_1 = c_1$ then $\rho_{b_1 b_2} = \rho_{c_1 c_2}$ by Lemma 7. Therefore, the points where the discontinuity of $R_2(x)$ might happen are restricted to the dividing points of the partition $\mathcal{P}_1$. Suppose $R_2(x)$ is not continuous at the point x_0. Since F is Bernoulli, the left and the right intervals of x_0 are of the form $\langle a_\nu a_r \rangle$ and $\langle a_{\nu+1} a_1 \rangle$ $(\nu = 1, 2, \cdots, r-1)$, respectively. From $\boldsymbol{\rho}_2 = \Phi_2^* \boldsymbol{\rho}_2$, for any $b, c \in \mathcal{A}$, we get

$$\rho_{bc} = \sum_{k=1}^{r} \eta_{a_k b} \rho_{a_k b} .$$

Then,

$$\rho_{a_\nu a_r} - \rho_{a_{\nu+1} a_1} = \sum_{k=1}^{r} (\eta_{a_k a_\nu} - \eta_{a_k a_{\nu+1}}) \rho_{a_k a_\nu} .$$

Similarly, the points that $R_N(x)$ is not continuous are the dividing points of the partition $\mathcal{P}_{N-1}$. Therefore, for $w \in W_N$,

$$\begin{aligned}
\rho_w &= \sum_{k=1}^{r} \eta_{a_k h(w)} \rho_{a_k h(w)} \\
&= \sum_{k=1}^{r} \eta_{a_k h(w)} \sum_{j=1}^{r} \eta_{a_j a_k h^2(w)} \rho_{a_j a_k h^2(w)} \\
&= \sum_{v : |v| = N-1} \rho_{v w[1]} \prod_{k=0}^{N-1} \eta_{t^k(v) h^{N-k}(w)} .
\end{aligned} \tag{17}$$

Denote the word $a\underbrace{bb\cdots b}_{n}$ by $ab(n)$. For the partition $\mathcal{P}_N$, the left and the right intervals of above x_0 are of the form $\langle a_\nu \underbrace{a_r\cdots a_r}_{N-1}\rangle$ and $\langle a_{\nu+1}\underbrace{a_1\cdots a_1}_{N-1}\rangle$, that is, $\langle a_\nu a_r(N-1)\rangle$ and $\langle a_{\nu+1}a_1(N-1)\rangle$.

Since $\rho_{va_\nu} = \rho_{va_{\nu+1}}$, we get

$$\begin{aligned}
&\rho_{a_\nu a_r(N-1)} - \rho_{a_{\nu+1}a_1(N-1)} \\
&\quad= \sum_{v:|v|=N-1} \rho_{va_\nu}\left(\prod_{k=0}^{N-1}\eta_{t^k(v)a_\nu h^{N-k-1}(a_r(N-1))} - \prod_{k=0}^{N-1}\eta_{t^k(v)a_{\nu+1}h^{N-k-1}(a_1(N-1))}\right) \\
&\quad= \sum_{v:|v|=N-1} \rho_{va_\nu}\left(\prod_{k=0}^{N-1}\eta_{t^k(v)a_\nu a_r(k)} - \prod_{k=0}^{N-1}\eta_{t^k(v)a_{\nu+1}a_1(k)}\right). \qquad (18)
\end{aligned}$$

For fixed $v \in W_{N-1}$,

$$\begin{aligned}
&\prod_{k=0}^{N-1}\eta_{t^k(v)a_\nu a_r(k)} - \prod_{k=0}^{N-1}\eta_{t^k(v)a_{\nu+1}a_1(k)} \\
&= \sum_{j=1}^{N-1}(\eta_{t^j(v)a_\nu a_r(j-1)} - \eta_{t^j(v)a_{\nu+1}a_1(j-1)})\left(\prod_{k=0}^{j-1}\eta_{t^k(v)a_\nu a_r(k)}\prod_{k=j+1}^{N-1}\eta_{t^k(v)a_{\nu+1}a_1(k)}\right) \\
&\le \max_{w\in W_N}\mathrm{Lebes}(\langle w\rangle)\max_{x\in[0,1]}\frac{|F''(x)|}{|F'(x)|}\sum_{j=1}^{N-1}\left(\prod_{k=0}^{j-1}\eta_{t^k(v)a_\nu a_r(k)}\prod_{k=j+1}^{N-1}\eta_{t^k(v)a_{\nu+1}a_1(k)}\right) \\
&\le Ke^{-(\xi-\varepsilon)N}\left(\sum_{j=0}^{N_0-1} + \sum_{j=N_0}^{N-N_0} + \sum_{j=N-N_0+1}^{N-1}\right)\left(\prod_{k=0}^{j-1}\eta_{t^k(v)a_\nu a_r(k)}\prod_{k=j+1}^{N-1}\eta_{t^k(v)a_{\nu+1}a_1(k)}\right). \qquad (19)
\end{aligned}$$

According to the note in (2), for $N > 2N_0$

$$\begin{aligned}
(19) &\le Ke^{-(\xi-\varepsilon)N}(2N_0c^{N_0}e^{-(\xi-\varepsilon)(N-N_0)} + (N-2N_0)e^{-(\xi-\varepsilon)N}) \\
&\le Ke^{-2(\xi-\varepsilon)N}\{N + 2N_0(c^{N_0}e^{(\xi-\varepsilon)N_0}) + 1\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
(18) &\le Ke^{-2(\xi-\varepsilon)N}(N+K')\sum_{v:|v|=N-1}\rho_v \\
&\le Ke^{-2(\xi-\varepsilon)N}(N+K')r^{N-1}\|\boldsymbol{\rho}_N\|_N \\
&\le Kr^{N-1}e^{-2(\xi-\varepsilon)N}(N+K'),
\end{aligned}$$

where $K' = 2N_0(c^{N_0}e^{\xi-\varepsilon})^{N_0} + 1$. By the assumption $\xi \ge \frac{1}{2}\log r$ and $\|\boldsymbol{\rho}_N\|_N < \infty$, this converges to 0 as $N \to \infty$.

The other discontinuities of $R_N(x)$ are between $\langle wa_\nu a_r(m)\rangle$ and $\langle wa_{\nu+1}a_1(m)\rangle$ for $w \in W_{N-m-1}$, $m = 1, 2, \cdots, N-2$. All the discontinuity of $R_N(x)$ is of this form. Similarly to (17), we get

$$\rho_{wa_\nu a_1(m)} = \sum_{v:|v|=m} \rho_{vwa_\nu} \prod_{k=0}^{m} \eta_{t^k(v)wa_\nu a_r(k)} .$$

Therefore,

$$\begin{aligned}
&\rho_{wa_\nu a_r(m)} - \rho_{wa_{\nu+1}a_1(m)} \\
&\quad = \sum_{v:|v|=m} \rho_{vwa_\nu} \prod_{k=0}^{m} \eta_{t^k(v)wa_\nu a_r(k)} - \sum_{v:|v|=m} \rho_{vwa_{\nu+1}} \prod_{k=0}^{m} \eta_{t^k(v)wa_\nu a_1(k)} \\
&\quad = \sum_{v:|v|=m} \rho_{vwa_\nu} \left(\prod_{k=0}^{m} \eta_{t^k(v)wa_\nu a_r(k)} - \prod_{k=0}^{m} \eta_{t^k(v)wa_\nu a_1(k)} \right).
\end{aligned}$$

For the fixed discontinuity, $m \to \infty$ as $N \to \infty$. Then by the similar calculation as (18), this difference converges to 0. Thus the theorem is proved. $\square$

References

[1] A. LASOTA and J. YORKE, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. **186** (1973), 481–488.

[2] M. MORI, *Fredholm determinant for piecewise linear transformations*, Osaka J. Math. **27** (1990), 481–116.

[3] M. MORI, *Fredholm determinant for piecewise monotonic transformations*, Osaka J. Math. **29** (1992), 479–529.

[4] M. MORI, *On the Convergence of the Spectrum of Perron-Frobenius Operators*, Tokyo J. Math. **17** (1994), 1–19.

Present Address:
DEPARTMENT OF NATURAL, SCIENCES, TOKYO NATIONAL COLLEGE OF TECHNOLOGY,
KUNUGIDA-MACHI, HACHIOJI-SHI, TOKYO 193–0997, JAPAN.