Ergod. Th. & Dynam. Sys. (1997), **17**, 1147–1181
Printed in the United Kingdom © *1997 Cambridge University Press*

Markov systems and transfer operators associated with cofinite Fuchsian groups

TAKEHIKO MORITA

Department of Mathematics, Tokyo Institute of Technology, Oh-okayama Meguro-ku, Tokyo 152, Japan

(*Received* 8 *July* 1995)

Abstract. In this paper we study a generalization of Mayer's result on the Selberg zeta function of $PSL(2,\mathbb{Z})$. Let Γ be a cofinite Fuchsian group. We construct a Markov system $\mathcal{T}_\Gamma$ by modifying the Bowen–Series construction of a Markov map T_Γ associated with Γ. The Markov system enables us to define transfer operators $L(s)$ for $\mathcal{T}_\Gamma$ so that they determine a meromorphic function taking values with nuclear operators on a nice function space. We show that the Selberg zeta function $Z(s)$ of Γ has a determinant representation $Z(s) = \mathrm{Det}(I - L(s))F(s)$, where $\mathrm{Det}(I - L(s))$ is the Fredholm determinant of $L(s)$ and $F(s)$ is a meromorphic function depending only on a finite number of hyperbolic conjugacy classes of Γ. Combining such a representation and the investigation of the spectral properties of $L(s)$, we can also obtain some analytic information of $Z(s)$.

1. *Introduction*

As is well known, the action of the group

$$PSL(2,\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc = 1 \right\} / \{I, -I\}$$

in the upper half plane $\mathbb{H} = \{z = x + \sqrt{-1}\, y : x \in \mathbb{R}, y > 0\}$ is given by

$$z \mapsto \frac{az + b}{cz + d}$$

and $PSL(2,\mathbb{R})$ is characterized as the isometry group with respect to the Poincaré metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

Let Γ be a finitely generated discrete subgroup of $PSL(2,\mathbb{R})$ whose action in $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ has a dense orbit. Such a group Γ is called a cofinite group since the quotient space $M = \mathbb{H}/\Gamma$ is a Riemann surface of finite area, possibly with a finite number of

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

ramification points. In particular, a group Γ with $M = \mathbb{H}/\Gamma$ being compact is called a cocompact group. The modular group

$$\Gamma(1) = PSL(2,\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathbb{Z}, ad-bc=1 \right\} /\{I,-I\}$$

is a typical example of cofinite group which is not cocompact.

For a cofinite group Γ we consider the following function defined by the formal Euler product:

$$Z(s) = \prod_{k=0}^{\infty} \prod_{c \in HC(\Gamma)} (1 - \exp(-(s+k)\ell(c))),$$

where $HC(\Gamma)$ denotes the totality of hyperbolic conjugacy classes, i.e. the conjugacy classes of primitive hyperbolic elements in Γ and the quantity $\ell(c)$ is given by

$$\ell(c) = 2\cosh^{-1}\frac{\operatorname{tr} h}{2}, \quad \text{for } h \in c.$$

The formally defined function $Z(s)$ is the so-called Selberg zeta function for Γ [**9**, **19**]. Recall that an element in $PSL(2,\mathbb{R})$ is said to be hyperbolic if it has exactly two fixed points on $\hat{\mathbb{R}}$. A hyperbolic element in Γ is said to be primitive if it cannot be written as an n-times composition of any hyperbolic element in Γ except for the case $n = \pm 1$. If $h \in \Gamma$ is a primitive hyperbolic element fixing $x, y \in \hat{\mathbb{R}}$, then it fixes the geodesic $\gamma(x,y)$ joining x and y, i.e. the semicircle perpendicular to the real axis at x and y. The projection of $\gamma(x,y)$ to the quotient space $M = \mathbb{H}/\Gamma$ is a closed geodesic and $\ell(c)$ with $c = (h)$ is the hyperbolic length of the closed geodesic. Moreover, there is a natural one-to-one correspondence between the set $HC(\Gamma)$ of the hyperbolic conjugacy classes and the set $CG(M)$ of oriented closed geodesics on M. Consider the Selberg L-function defined formally by

$$\zeta(s) = \frac{Z(s+1)}{Z(s)} = \prod_{c \in HC(\Gamma)} (1 - \exp(-s\ell(c)))^{-1}.$$

Clearly, $\zeta(s)$ is a direct analogue of a Riemann zeta function. Therefore we expect that the Selberg zeta function has fruitful information of the distribution of the length spectrum of the surface M.

Let us consider the modular group $\Gamma(1)$. If we define a map $T_{\Gamma(1)} : \hat{\mathbb{R}} \to \hat{\mathbb{R}}$ by

$$T_{\Gamma(1)}x = \begin{cases} x+1 & (x \le -1) \\ -\dfrac{1}{x} & (|x| < 1) \\ x-1 & (x \ge 1), \end{cases}$$

then it is easy to see that the action of $\Gamma(1)$ in $\hat{\mathbb{R}}$ is orbit equivalent to that of $T_{\Gamma(1)}$, i.e. for $x, y \in \hat{\mathbb{R}}$, $y = gx$ for some $g \in \Gamma(1)$ if and only if $T^m_{\Gamma(1)}x = T^n_{\Gamma(1)}y$ for some $m, n \ge 0$. In addition, the first return map of $T_{\Gamma(1)}$ to the unit interval $(0, 1)$ can be defined up to the set of rationals and coincides with the twice composition $T_G^2 = T_G \circ T_G$ of the well-known Gauss transformation

$$T_G x = \frac{1}{x} - \left[\frac{1}{x}\right],$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

where $[a]$ denotes the integral part of a. By a formal calculation, it is not hard to see that

$$Z(s) = \exp\left(-\sum_{k=0}^{\infty}\sum_{n=1}^{\infty}\frac{1}{n}Z_n(s)\right),$$

with

$$Z_n(s) = \sum_{T_G^{2n}x=x}(T_G^{2n})'(x)^{-s}.$$

Here $\sum_{Tx=x}$ denotes the sum taken over all fixed points of the map T. Thus we can see that the Selberg zeta function for the modular group has an expression similar to that of a dynamical zeta function appearing in the thermodynamic formalism [**14**, **16**–**18**] and so on.

In [**11**] and [**12**], Mayer considers a family of operators $L_s : \mathcal{A}_b(D) \to \mathcal{A}_b(D)$ defined by

$$L_s f(z) = \sum_{n=1}^{\infty}\left(\frac{1}{n+z}\right)^{2s} f\left(\frac{1}{n+z}\right),$$

where $\mathcal{A}_b(D)$ is the Banach space of bounded analytic functions on a suitably chosen bounded domain D in $\mathbb{C}$ containing the unit interval. He proves that L_s is analytic in $\{s : \Re s > 1/2\}$ and meromorphic in the entire s-plane as a function taking values in the space of nuclear operators of order zero. He also obtains a determinant representation

$$Z(s) = \mathrm{Det}(I - L_s)\,\mathrm{Det}(I + L_s)$$

and, without using the Selberg trace formula in [**9**], he gives an alternative proof of the fact that $Z(s)$ is extended to a meromorphic function in the entire s-plane with singularities, possibly removable, at $s = -k/2, k = -1, 0, 1, 2, \ldots$, where $\mathrm{Det}(I \pm L_s)$ is the Fredholm determinant of $\pm L_s$ [**8**]. We can see that the operator L_s has the form of the transfer operator with respect to the one-dimensional dynamical system T_G. In fact,

$$L_s f(x) = \sum_{T_G y = x}|T_G'(y)|^{-s} f(y)$$

holds up to a Lebesgue null set in [0, 1] if $\Re s > 1/2$. We must note that Efrat [**5**] obtains a remarkable result that relates the above factorization to the spectral theory of the Laplace operator on the modular surface.

On the other hand, there is a remarkable work by Bowen and Series [**3**]. It provides a 'canonical' way of constructing a single transformation T_Γ of $\hat{\mathbb{R}}$ whose action in $\hat{\mathbb{R}}$ is orbit equivalent to that of the group Γ. Therefore it is natural to expect that if we use the map T_Γ (or the first return map of T_Γ if necessary), we can represent the Selberg zeta function for a general cofinite group Γ as the Fredholm determinant of the 'transfer operator' with respect to the dynamical system associated with Γ. Mayer himself suggests such a generalization in [**12**, Introduction] and mentions the difficulties that we face when we try to carry it out. As noted in [**12**], we find that the cocompact case is treated in this direction by Pollicott [**16**].

In this paper, we study the Bowen–Series construction of Markov maps T_Γ associated with the cofinite group Γ and introduce the Markov system $\mathcal{T}_\Gamma$ associated with Γ by

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

modifying T_Γ. As stated later, the Markov system enables us to define the desired function space $\mathcal{A}_b(D(X))$ and a family of transfer operators $L(s)$ in such a natural way that we can show the following theorems announced essentially in [**13**].

THEOREM A. *$L(s)$ is an analytic function in the domain $\{s \in \mathbb{C} : \Re s > 1\}$ having a meromorphic extension to the entire s-plane with values in the space of nuclear operators of order zero on $\mathcal{A}_b(D(X))$. The candidates for its poles are the points $s = -k/2, k = -1, 0, 1, 2, \ldots$.*

THEOREM B. *For $s, \Re s > 1$, the Selberg zeta function has a determinant representation*

$$Z(s)\Xi(s)^2 = \mathrm{Det}(I - L(s)),$$

where $\Xi(s) = \prod_{k=0}^{\infty} \prod_{c \in HC_1(\Gamma)} (1 - \exp(-(s+k)\ell(c)))$ for a certain finite subset $HC_1(\Gamma)$ of $HC(\Gamma)$. In particular, $Z(s)$ is an analytic function without zeros in the domain $\{s \in \mathbb{C} : \Re s > 1\}$ having a meromorphic extension to the entire s-plane and its poles are located on the axes $\Re s = -k/2, k = -1, 0, 1, 2, \ldots$.

Through the investigation of the spectral properties of transfer operators, we can show that there is a domain containing $\{s \in \mathbb{C} : \Re s \geq 1\}$, where $s = 1$ is a unique zero of the analytic function $Z(s)$ and it is simple. Therefore the Wiener–Ikehara proof of the prime number theorem does work and we obtain the so called closed geodesic theorem (see [**14**]).

THEOREM C. *For any cofinite group, we have*

$$\sharp\{c : \exp(\ell(c)) \leq t\} \frac{\log t}{t} \to 1 \quad (t \to \infty).$$

The reader familiar with the Selberg zeta function for cocompact groups and the modular group notices that Theorem B is weaker than the results obtained by the Selberg trace formula even though it might be the most general result obtained by thermodynamic formalism without using the trace formula. For example, the author does not yet know how to prove the functional equation by this approach. Concerning the determinant representation of $Z(s)$ as the Fredholm determinant of the transfer operator, the author is informed by Professor Koyama of Keio University that Bogomolney and Carioli [**1**] obtain it from a physical point of view.

As in [**3**] we work on the unit disc $\mathbb{D}$ throughout the paper instead of the upper half plane $\mathbb{H}$ by conformal change of variables. Then the cofinite subgroups of $PSL(2, \mathbb{R})$ are transformed to those of

$$PSU(1,1) = \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} : |a|^2 - |b|^2 = 1 \right\} / \{I, -I\}.$$

This paper is organized as follows. In §2, starting with the Markov map T_Γ associated with a cofinite group Γ, we construct a triplet $\mathcal{T}_\Gamma = (\mathcal{R}, \mathcal{P}, T)$ which will be called the Markov system associated with Γ. Roughly speaking, T is the first return map of T_Γ to a suitably chosen subset X of the unit circle defined at least a.e., $\mathcal{R}$ is a finite partition of X into subarcs, and $\mathcal{P}$ is a refinement of $\mathcal{R}$ into, possibly, an infinite number of subarcs so that to each subarc $J \in \mathcal{P}$, an element T_J of Γ satisfying

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

$T|_{\text{int } J} = T_J|_{\text{int } J}$ is assigned. T, $\mathcal{R}$ and $\mathcal{P}$ in the general case play the roles of T_G^2, $\{[0, 1]\}$ and $\{[1/(n+1), 1/n] : n = 1, 2, \ldots\}$ in the case of the modular group, respectively.

In §3, some basic properties of the Markov system will be investigated. We introduce the notion of $\mathcal{T}_\Gamma$-periodic orbits and study the correspondence between them and hyperbolic conjugacy classes in §4. In §5 we define a family of transfer operators $L(s)$ with respect to the Markov system $\mathcal{T}_\Gamma$ and we show that it is realized as a meromorphic function taking values with nuclear operators on an appropriately chosen function space using the results in §3. §6 is devoted to the study of the ergodic or the spectral properties of transfer operators $L(s)$ with $\Re s > 1/2$. We prove the Ruelle–Perron–Frobenius type uniform ergodic theorem (cf. [**2**, **14**]). In the last section we give the results which imply Theorems A–C.

2. *Bowen–Series Markov maps and their modifications*

First of all we recall the Bowen–Series construction of the Markov maps associated with cofinite groups. As in [**3**] we work on the unit disc by means of the conformal transform. The group

$$PSU(1,1) = \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}; a, b \in \mathbb{C},\ a\bar{a} - b\bar{b} = 1 \right\} / \{-I, I\}$$

is identified with the group of linear fractional transformations

$$z \mapsto \frac{az + b}{\bar{b}z + \bar{a}}$$

acting in the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ as isometries with respect to the Poincaré or the hyperbolic metric

$$ds^2 = \frac{4\, dz\, d\bar{z}}{(1 - |z|^2)^2}.$$

Let Γ be a finitely generated Fuchsian group of the first kind. In other words, Γ is a discrete subgroup of $PSU(1, 1)$ for which the quotient space $M = \mathbb{D}/\Gamma$ is a Riemann surface of finite hyperbolic area, possibly, with a finite number of ramification points. Therefore such a group is sometimes called a cofinite group.

Every cofinite group Γ has a fundamental polygon R with a finite number of sides $\{s_i\}_{i=1}^N$, each of which extends to a circular arc $C(s_i)$ perpendicular to the unit circle S^1. For each side of R there is a corresponding element $g_i = g(s_i) \in \Gamma$ by which s_i is identified with another side $\tilde{s}_i$. Then $\Gamma_0 = \{g_i\}_{i=1}^N$ forms a symmetric set of generators for Γ. We denote by $\mathcal{N}$ the net consisting of all the Γ images of sides of R.

Throughout the paper we impose the following conditions on R:

$C(s_i)$ is the isometric circle of g_i and lies completely in the net $\mathcal{N}$. $(*)$

R is not a triangle and does not have an elliptic vertex of order two. $(**)$

We have to note that even if R does not satisfy $(*)$ and $(**)$, we can construct the desired Markov map using the quasiconformal deformation of Fuchsian groups in the same way as [**3**, §4].

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

In what follows, our labelling of the sides is assumed to be given in anti-clockwise order along the unit circle. Let v_i denote the vertex made by the sides s_i and s_{i-1} ($s_0 = s_N$). The intersection points of $C(s_i)$ and S^1 are labelled by P_i and Q_{i+1} ($Q_{N+1} = Q_1$) so that the order of these points along $C(s_i)$ is $P_i, v_i, v_{i+1}, Q_{i+1}$. The Markov map T_Γ (denoted by f_Γ in [**3**]) is defined by

$$T_\Gamma x = g_i x \quad \text{for } x \in [P_i, P_{i+1}). \tag{2.1}$$

Here we use the notation $[P, Q]$, (P, Q), and $[P, Q)$, to denote a closed, an open, and a half-open arc lying between P and Q in anti-clockwise order from P to Q, respectively. The map $\tilde{T}_\Gamma$ is also defined symmetrically as

$$\tilde{T}_\Gamma x = g_{i-1} x \quad \text{for } x \in (Q_{i-1}, Q_i]. \tag{2.2}$$

To understand the structure of the map T_Γ in detail, we need more notation. For each vertex v_i, $\mathcal{N}(v_i)$ denotes the set of arcs in the net $\mathcal{N}$ passing through v_i. Let $W(v_i)$ be the set of points where arcs in $\mathcal{N}(v_i)$ meet S^1. Put $W = \cup_{i=1}^N W(v_i)$. Note that the arcs in $\mathcal{N}(v_i)$ are complete geodesics because of $(*)$.

If we put $2k_i = \sharp W(v_i)$, it is easy to see that $k_i = \infty$ if and only if v_i is a vertex on S^1. If $k_i < \infty$, we label the $2k_i$ arcs cut off from S^1 by the points in $W(v_i)$ as follows. First we label $[P_i, Q_i)$ as $L_{k_i}(v_i)$. Successively we label the arcs in anti-clockwise order along the unit circle as $L_{k_i-1}(v_i), L_{k_i-2}(v_i), \ldots$, so that the arc $[T_i, Q_{i+1})$ is labelled by $L_1(v_i)$, where T_i is the element of $W(v_i)$ immediately preceding Q_{i+1} in anti-clockwise order. Next we label the arc $[Q_{i+1}, P_{i-1})$ as $R_1(v_i)$ and continue the labelling anti-clockwise as $R_2(v_i), R_3(v_i), \ldots$ so that the arc $[P_{i-1}, S_{i-1})$ is labelled by $R_2(v_i)$, where S_{i-1} is the next element P_{i-1} of $W(v_i)$ in anti-clockwise order (see Figure 1). Clearly the preceding arc of $L_{k_i}(v_i)$ is labelled by $R_{k_i}(v_i)$.

If $k_i = \infty$, we label the arcs cut off by the points in $W(v_i)$ in the same way so that $[Q_{i+1}, P_{i-1})$ is labelled as $R_1(v_i)$ and the preceding arc in anti-clockwise order is labelled as $L_1(v_i)$. In this case $P_i = Q_i = v_i$.

The following rules play important roles in our later investigations:

$$\begin{aligned} T_\Gamma L_j(v_i) &= L_{j-1}(g_i v_i) \quad \text{for } 2 \le j \le k_i, \\ T_\Gamma R_j(v_i) &= R_{j-1}(g_{i-1} v_i) \quad \text{for } 3 \le j \le k_i. \end{aligned} \tag{2.3}$$

We pay attention to the vertex v_i on S^1. Let ν be the smallest positive integer such that $T_\Gamma^\nu(L_{\nu+1}(v_i)) = L_1(v_i)$. If we put $v_{i_0} = v_i$, $v_{i_k} = T_\Gamma^k v_i (k \ge 1)$. Then $i_{k+\nu} = i_k$ for any $k \ge 0$ and $h_i = g_{i_{\nu-1}} g_{i_{\nu-2}} \cdots g_{i_0}$ is a parabolic element in Γ fixing v_i and $h_i|_{L_{\nu+1}} = T_\Gamma^\nu|_{L_{\nu+1}}$. Moreover, we have

$$T_\Gamma^{k\nu+\ell} = h_i^k g_{i_{\nu-1}} \cdots g_{i_{\nu-\ell}} \quad \text{on } L_{k\nu+\ell+1}(v_{i_{\ell-\nu}}) \text{ for any } k \in \mathbb{N} \text{ and } 1 \le \ell \le \nu. \tag{2.4}$$

Since the rotations around the origin are elements in $SU(1,1)$, we may assume that $v_i = 1$ and the matrix representation of h_i^k has the form

$$h_i^k = \begin{pmatrix} 1 + \sqrt{-1}\, kb & -\sqrt{-1}\, kb \\ \sqrt{-1}\, kb & 1 - \sqrt{-1}\, kb \end{pmatrix} \quad \text{for some } b \in \mathbb{R}. \tag{2.5}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

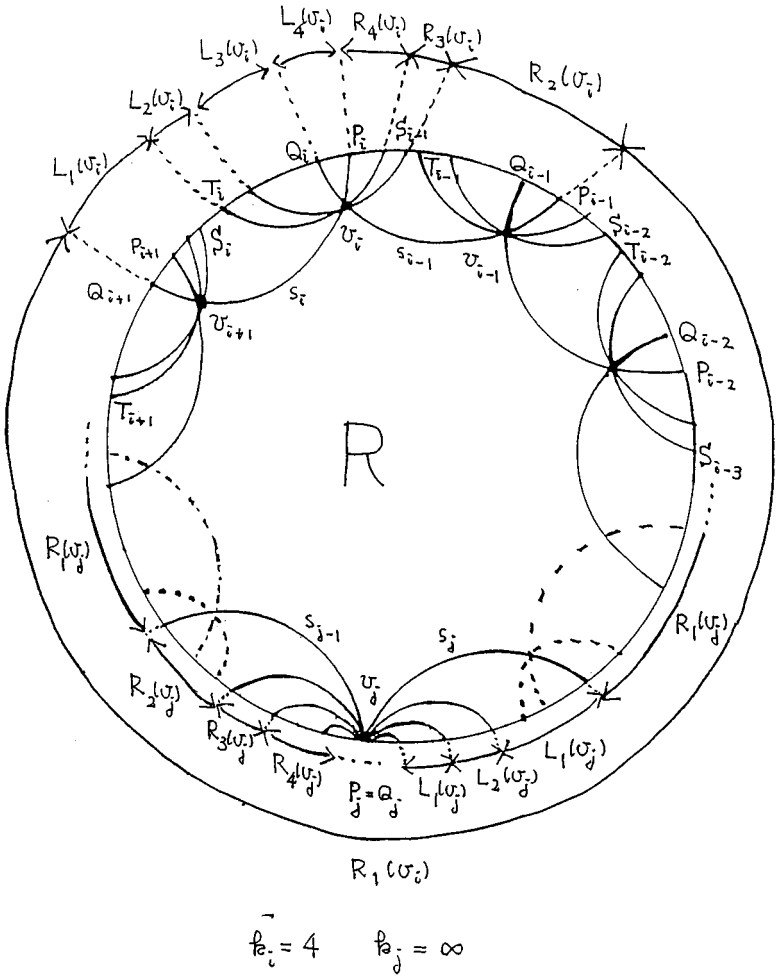


FIGURE 1.

In particular, we have

$$(h_i^k)'(z) = \frac{1}{(\sqrt{-1}\,kbz + 1 - \sqrt{-1}\,kb)^2}. \tag{2.6}$$

(2.6) yields the estimate

$$|L_n(v_i)| \le Cn^{-2} \tag{2.7}$$

for any $n \in \mathbb{N}$, where $|\cdot|$ denotes the arclength and C is a positive constant independent of n. The same is true for the $R_n(v_i)$'s.

In what follows we denote by $\mathcal{P}(W) = \{I_i\}_i$ the totality of the arcs cut off from S^1 by the points in W. We summarize the properties of T_Γ following directly from the definition.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

LEMMA 2.1. (See [**3**].)

(1) *(Orbit equivalence.) For any pair* $(x, y) \in S^1 \times S^1$, $y = gx$ *for some* $g \in \Gamma$ *if and only if* $T_\Gamma^m x = T_\Gamma^n y$ *for some* $m, n \geq 0$.

(2) *(Piecewise monotonicity.) For each* $I_i \in \mathcal{P}(W)$, $T_\Gamma|_{I_i} = g|_{I_i}$ *holds for some* $g \in \Gamma$. *In particular,* $T_\Gamma|_{I_i}$ *is strictly increasing and has an analytic extension.*

(3) $\sharp\mathcal{P}(W) < \infty$ *if and only if the Riemann surface* $\mathbb{D}/\Gamma$ *is compact.*

(4) *(Markov property.)* $T_\Gamma I_i \cap I_j \neq \emptyset$ *implies that* $T_\Gamma I_i \supset I_j$.

(5) *(Transitivity.)* $\cup_{n=0}^{\infty} T_\Gamma^n I_i \supset I_j$ *for any* $I_i, I_j \in \mathcal{P}(W)$.

Remark 2.2. (1) In the original statement of the orbit equivalence in [**3**], there are a finite number of exceptional pairs $(Q_i, g_{i-1}Q_i)$. But if one notices that $g_{i-1}Q_i = P_j$ for some j and $T_\Gamma|_{[P_j, P_{j+1})} = g_j = g_{i-1}^{-1}$, one can see that $T_\Gamma g_{i-1}Q_i = Q_i$. Thus the pair $(Q_i, g_{i-1}Q_i)$ is badly matched in the sense of [**3**] but not exceptional.

(2) Reading the proof of orbit equivalence in [**3**] carefully, one can notice that it tells us slightly stronger assertions, as follows.

> For any $x \in S^1$ and for any $g \in \Gamma$, let m and n be non-negative integers appearing in the orbit equivalence (1) in Lemma 2.1. Suppose that T_Γ^m and T_Γ^n are written as $g_{i_{m-1}} g_{i_{m-2}} \cdots g_{i_0}$ and $g_{j_{n-1}} g_{j_{n-2}} \cdots g_{j_0}$ at x and gx, respectively. Then there exists a small arc $[x, z)$ such that $T_\Gamma^m w = T_\Gamma^n gw$ holds for any $w \in [x, z)$ with $T_\Gamma^m w = g_{i_{m-1}} g_{i_{m-2}} \cdots g_{i_0} w$ and $T_\Gamma^n gw = g_{j_{n-1}} g_{j_{n-2}} \cdots g_{j_0} gw$. In particular, if the T_Γ-orbit of x does not intersect the set W of partition points, we can choose an open arc containing x instead of the arc $[x, z)$ in the above.

Now we are ready to modify the notion of Markov maps to be suitable for our investigations.

Put

$$\begin{aligned}
\mathcal{R}_i &= \{\overline{L_2(v_i)}, \overline{L_3(v_i)}, \ldots, \overline{L_{k_i}(v_i)}, \overline{R_3(v_i)}, \overline{R_4(v_i)}, \ldots, \overline{R_{k_i}(v_i)}, [T_i, S_i]\} \quad \text{for } v_i \in \mathbb{D}, \\
\mathcal{R}_i &= \{\overline{L_2(v_i)}, \overline{R_3(v_i)}, [T_i, S_i]\} \quad \text{for } v_i \in S^1, \\
\mathcal{R} &= \bigcup_{i=1}^{N} \mathcal{R}_i. \qquad (2.8)
\end{aligned}$$

We write $\mathcal{R} = \{A(1), A(2), \ldots, A(q)\}$ and often use the notation $A \in \mathcal{R}$ for the sake of simplicity.

Define a set of non-empty closed arcs in S^1 by

$$\begin{aligned}
\mathcal{P} = \{&J;\ J \text{ has the form } J = \overline{(\operatorname{int} A) \cap T_\Gamma^{-1}(\operatorname{int} I)} \text{ for some } A \in \mathcal{R} \\
&\text{and for some } I \in \mathcal{P}(W)\}. \qquad (2.9)
\end{aligned}$$

We write $\mathcal{P}$ as $\mathcal{P} = \{J(1), J(2), \ldots\}$.

Remark 2.3.

(1) If $J \in \mathcal{P}$, there are the following possibilities:

(a) $J = \overline{L_j(v_i)} \in \mathcal{R} (j \geq 3)$ for some $v_i \in \mathbb{D}$;

(b) $J = \overline{R_k(v_i)} \in \mathcal{R} (k \geq 4)$ for some $v_i \in \mathbb{D}$;

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

(c) J is a closed subarc in $[T_i, S_i]$, $\overline{L_2(v_i)}$, or $\overline{R_3(v_i)}$ for some $i = 1, 2, \ldots, N$.

(2) If we regard $\mathcal{R}$ as a partition (a.e.) of $K = \cup_{i=1}^{q} A(i)$, $\mathcal{P}$ can be regarded as a refinement of $\mathcal{R}$ (a.e.). Namely, we have:
   (a) $\cup_{i=1}^{q} A(i) \setminus \cup_{J\in\mathcal{P}} J$ is a Lebesgue null set;
   (b) for any $J \in \mathcal{P}$, there is a unique $A \in \mathcal{R}$ containing J.

Next we assign an element T_J of Γ to each arc J in $\mathcal{P}$ as follows. Assume that J is expressed as $J = (\operatorname{int} A) \cap \overline{T_\Gamma^{-1}(\operatorname{int} I)}$ for some $A \in \mathcal{R}$ and for some $I \in \mathcal{P}(W)$. In the case when $\operatorname{int} I \subset K = \cup_{i=1}^{q} A(i)$, T_J is an element $g \in \Gamma_0$ satisfying $T_\Gamma|_{\operatorname{int} J} = g|_{\operatorname{int} J}$. In the case when $\operatorname{int} I \not\subset K$, i.e. $\operatorname{int} I \cap K = \emptyset$, we have $T_\Gamma \operatorname{int} J = \operatorname{int} L_j(v_i)$ (respectively $R_k(v_i)$) for some $j \geq 3$ (respectively $k \geq 4$) with $v_i \in S^1$. Therefore we have $T_\Gamma^{j-1} \operatorname{int} J = \operatorname{int} L_2(T_\Gamma^{j-2} v_\ell)$ (respectively $T_\Gamma^{k-2} \operatorname{int} J = \operatorname{int} R_3(T_\Gamma^{k-3} v_\ell)$) for some ℓ in virtue of the relation (2.3). Thus we define T_J as the element g in Γ so that $g|_{\operatorname{int} J} = T_\Gamma^{j-1}|_{\operatorname{int} J}$ (respectively $g|_{\operatorname{int} J} = T_\Gamma^{k-2}|_{\operatorname{int} J}$).

Now we can define a map $T : \bigsqcup_{J\in\mathcal{P}} J \to X = \bigsqcup_{i=1}^{q} A(i)$ so that

$$T|_{\operatorname{int} J} = T_J|_{\operatorname{int} J} \quad \text{for each } J \in \mathcal{P}, \tag{2.10}$$

where $\bigsqcup$ denotes the disjoint union.

Remark 2.4. From the definition, T can be regarded as a map $T : X \to X$ (a.e.) and it plays the role of the first return map $T_K : K \to K$ (a.e.) of T_Γ to K.

Now we introduce the map which plays the same role for T_Γ^n as T does for T_Γ. Let $\mathcal{P}_n$ be the family of closed arcs with non-empty interior having the form

$$J = J(i_0) \cap T_{J(i_0)}^{-1} J(i_1) \cap \cdots \cap T_{J(i_0)}^{-1} T_{J(i_1)}^{-1} \cdots T_{J(i_{n-2})}^{-1} J(i_{n-1}) \tag{2.11}$$

for $J(i_0), J(i_1), \ldots, J(i_{n-1}) \in \mathcal{P}$. Such a J will be denoted by $J(i_0 i_1 \ldots i_{n-1})$ in what follows. To each $J = J(i_0 i_1 \ldots i_{n-1}) \in \mathcal{P}_n$, we assign an element

$$T_J^n = T_{J(i_{n-1})} T_{J(i_{n-2})} \cdots T_{J(i_0)} \tag{2.12}$$

of Γ. Then the n times iteration T^n of T is well defined on $\cup_{J\in\mathcal{P}_n} \operatorname{int} J$, therefore it determines an a.e. defined map $T^n : X \to X$ satisfying $T^n|_{\operatorname{int} J} = T_J^n|_{\operatorname{int} J}$ for each $J \in \mathcal{P}_n$.

Now we are ready to define a Markov system.

Definition 2.5. A triplet $\mathcal{T} = (\mathcal{R}, \mathcal{P}, T)$ satisfying the following conditions is called a Markov system (on S^1).

(1) $\mathcal{R} = \{A(1), \ldots, A(q)\}$ is a set of a finite number of closed arcs A with mutually disjoint non-empty interior.

(2) $\mathcal{P} = \{J\}$ is a set of closed arcs J with mutually disjoint non-empty interior such that for each $J \in \mathcal{P}$, there is an $A \in \mathcal{R}$ with $\operatorname{int} J \subset \operatorname{int} A$ and $|\cup_{J\in\mathcal{P}} J \setminus \cup_{A\in\mathcal{R}} A| = 0$ holds, where $|\cdot|$ denotes the Lebesgue measure on S^1.

(3) To each $J \in \mathcal{P}$, an arc $A \in \mathcal{R}$ and a homeomorphism T_J from J onto A are assigned.

(4) T is an a.e. defined transformation on $X = \bigsqcup_{A\in\mathcal{R}} A$ such that $T|_{\operatorname{int} J} = T_J|_{\operatorname{int} J}$.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

For a Markov system $\mathcal{T} = (\mathcal{R}, \mathcal{P}, T)$, we can define $\mathcal{P}_n$ and T_J^n in the same way as (2.11) and (2.12) for $n \geq 1$. Then we obtain a new Markov system $\mathcal{S} = (\mathcal{R}, \mathcal{P}_n, S)$, where S satisfies $S_J = T_J^n$ for each $J \in \mathcal{P}_n$. Clearly, $S = T^n$ a.e. Therefore such a Markov system $\mathcal{S}$ is called the n times iteration of $\mathcal{T}$ and denoted by $\mathcal{T}^n = (\mathcal{R}, \mathcal{P}_n, T^n)$ in what follows. In this paper, a Markov system associated with Γ means the Markov system determined by the triplet defined by (2.8)–(2.10) instead of the single maps T_Γ and T_K. We denote it by $\mathcal{T}_\Gamma$.

We close this section by introducing the following convention.

Convention 2.6. Let $\mathcal{T} = (\mathcal{R}, \mathcal{P}, T)$ be a Markov system with $\mathcal{R} = \{A(1), \ldots, A(q)\}$. For $J \in \mathcal{P}$, $\iota(J)$ denotes the index with $\operatorname{int} J \subset \operatorname{int} A(\iota(J))$ and $\tau(J)$ denotes the index with $T_J J = A(\tau(J))$.

3. *Auxiliary properties of the Markov systems*

In this section we prove some properties of the Markov system associated with a cofinite Fuchsian group that are necessary in the later investigation.

Consider the following mixing condition on a Markov system $\mathcal{T} = (\mathcal{R}, \mathcal{P}, T)$ with $\mathcal{R} = \{A(1), \ldots, A(q)\}$.

Condition 3.1. There exists a positive integer n_0 satisfying the following. For each $J \in \mathcal{P}$, there are elements $J_1, J_2, \ldots, J_q$ in $\mathcal{P}_{n_0}$ contained in J such that $T_{J_i}^{n_0} J_i = A(i)$ holds for each $j = 1, 2, \ldots, q$.

Although the Markov system $\mathcal{T}_\Gamma$ constructed in the previous section does not necessarily satisfy Condition 3.1, the next proposition enables us to assume its validity by employing the iteration $\mathcal{T}_\Gamma^d$ for some $d > 0$ if necessary.

PROPOSITION 3.2. *Let $\mathcal{T}_\Gamma = (\mathcal{R}, \mathcal{P}, T)$ be the Markov system associated with Γ defined by (2.8)–(2.10). Then for some natural number d, we can classify the elements in $\mathcal{R}$ into d classes $\mathcal{R}^{(j)}$, $j = 0, 1, \ldots, d-1$, such that the following are satisfied:*

(1) *if $J \in \mathcal{P}$ and $A(\iota(J)) \in \mathcal{R}^{(j)}$, then $T_J J = A(\tau(J)) \in \mathcal{R}^{(j+1)}$ mod d;*

(2) *If we set $X^{(j)} = \bigsqcup_{A \in \mathcal{R}^{(j)}} A$ and $\mathcal{P}_d^{(j)} = \{J \in \mathcal{P}_d : A(\iota(J)) \in \mathcal{R}^{(j)}\}$, then the Markov system $(\mathcal{R}^{(j)}, \mathcal{P}_d^{(j)}, T^d)$ satisfies Condition 3.1, where we use the same notation T^d for T^d itself and its restriction to $X^{(j)} = \bigsqcup_{A \in \mathcal{R}^{(j)}} A$.*

Proof. For $\mathcal{R} = \{A(1), A(2), \ldots, A(q)\}$, we define a $q \times q$ matrix $P = (P(i, j))_{1 \leq i, j \leq q}$ with 0-1 entries so that $P(i, j) = 1$ if and only if there is a $J \in \mathcal{P}$ with $\iota(J) = i$ and $\tau(J) = j$, i.e. $J \subset A(i)$ and $T_J J = A(j)$. From the transitivity of T_Γ of Lemma 2.1, it is easy to see that the matrix P is irreducible. Changing the labelling if necessary, we can classify the indicies $1, 2, \ldots, q$ as $I_0 = \{1, \ldots, i_0\}$, $I_1\{i_0 + 1, \ldots, i_1\}$, $\ldots$, $I_{d-1} = \{i_{d-2} + 1, \ldots, i_d = q\}$ for some integer d so that $P(i_k, i_\ell) = 1$ implies that $i_k \in I_j$ and $i_\ell \in I_{j+1}$ mod d for some $j \in \{0, \ldots, d-1\}$. Moreover, we can show that there exists a positive integer n_0 such that $P^n(i, j) > 0$ for any i, j belonging to the same class whenever $n \geq n_0$. Put $\mathcal{R}^{(j)} = \{A(i_{j-1}), \ldots, A(i_j)\}$ and $\mathcal{P}_d^{(j)} = \{J \in \mathcal{P}_d : J \subset A(i) \text{ for some } i \in I_j\}$. Using the Markov property of T_Γ, it is not hard to see that the triplet $(\mathcal{R}^{(j)}, \mathcal{P}_d^{(j)}, T^d)$ satisfies the desired condition. □

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

In §5 we introduce a kind of transfer operator; Proposition 3.2 will be used to simplify the investigation of the spectral properties of such operators. The following two lemmas play essential roles in our investigation. In particular, Lemma 3.4 guarantees that we can define the transfer operators with respect to $\mathcal{T}_\Gamma$ as nuclear operators acting on an appropriate space.

LEMMA 3.3. *The Markov system $\mathcal{T}_\Gamma$ satisfies the following conditions.*

(1) *(Lasota–Yorke condition.) There exits a positive integer k such that*

$$\inf_{J\in\mathcal{P}_k}\inf_{x\in J}|(T_J^k)'(x)|>1.$$

(2) *(Rényi condition.) For any integer $n\geq 1$ it holds that*

$$R(\mathcal{T}_\Gamma^n)=\sup_{J\in\mathcal{P}_n}\sup_{x\in J}\left|\frac{(T_J^n)''(x)}{(T_J^n)'(x)^2}\right|<\infty.$$

Moreover, we have

$$\sup_n\frac{R(\mathcal{T}_\Gamma^n)}{n}<\infty.$$

Proof. All the assertions except for the last one are proved essentially in [**3**]. But the last assertion is an easy consequence of assertion (1) and the fact that if S_1 is of class C^2 with non-zero derivatives in a neighborhood of x and so is S_2 in a neighborhood of S_1x, then we have

$$\left|\frac{(S_2\circ S_1)''(x)}{(S_2\circ S_1)'(x)^2}\right|\leq\left|\frac{(S_2)''(S_1x)}{(S_2)'(S_1x)^2}\right|+\left|\frac{(S_1)''(x)}{S_2'(S_1x)(S_1)'(x)^2}\right|.$$ □

LEMMA 3.4. *Let $\mathcal{T}_\Gamma=(\mathcal{R},\mathcal{P},T)$ with $\mathcal{R}=\{A(1),\dots,A(q)\}$ be the Markov system associated with a cofinite Fuchsian group Γ. Then for each $A(i)$ there exist simply connected domains $B(i)$ and $D(i)$ in $\mathbb{C}$ such that:*

(1) *$A(i)\subset B(i)\subset D(i)$;*

(2) *$T_J^{-1}D(\tau(J))\subset B(\iota(J))$ for any $J\in\mathcal{P}$; and*

(3)

$$\sum_{J\in\mathcal{P}}\sup_{z\in D(\tau(J))}(|T_J'(T_J^{-1}z)|^{-\delta})<\infty$$

for any $\delta>\frac{1}{2}$, where $\iota(J)$ and $\tau(J)$ are as in Convention 2.6.

Proof. In what follows g_I stands for the element in Γ_0 with $g|_I=T_\Gamma|_I$ for each $I\in\mathcal{P}(W)$, where $\mathcal{P}(W)$ is as in the paragraph just before Lemma 2.1. Set $\mathcal{P}_0(W)=\{I\in\mathcal{P}(W):\operatorname{int}I\subset K\setminus\cup_{i=1}^N L_{k_i}(v_i)\}$, where $K=\cup_{i=1}^q A(i)$.

By virtue of the condition (∗), there exists $\lambda>1$ and $\epsilon_0>0$ such that

$$\inf_{I\in\mathcal{P}_0(W)}\inf_{z\in U_\epsilon(I)}|g_I'(z)|\geq\lambda \tag{3.1}$$

holds for any $\epsilon<\epsilon_0$, where $U_r(S)$ denotes the r-neighborhood of the set S in $\mathbb{C}$. We show that we can choose $B(i)$ and $D(i)$ as follows:

$$B(i)=\begin{cases}U_{\frac{1+\lambda}{2}\epsilon}(A(i)) & \text{if } A(i)=\overline{L_{k_j}(v_j)} \text{ for some } v_j\in\mathbb{D},\\ U_{\frac{1+2\lambda}{3\lambda}\epsilon}(A(i)) & \text{otherwise,}\end{cases}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

and

$$D(i) = \begin{cases} U_{\frac{1+2\lambda}{3}\epsilon}(A(i)) & \text{if } A(i) = \overline{L_{k_j}(v_j)} \text{ for some } v_j \in \mathbb{D}, \\ U_\epsilon(A(i)) & \text{otherwise} \end{cases}$$

for ϵ sufficiently small.

We verify the above for all the possibilities of $J \in \mathcal{P}$ taking Remark 2.3 into consideration.

Case 1. $T_J J = \overline{L_j(v_i)}$, $v_i \in \mathbb{D}$ and $2 \le j \le k_i - 2$.
In this case J coincides with either $\overline{L_{j+1}(v_c)}$ or a subarc contained in $\overline{R_3(v_k)}$ or $[T_k, S_k]$ for some k. In any case T_J is an element of the symmetric set Γ_0 of generators of Γ. Thus (3.1) yields $T_J^{-1} U_\epsilon(\overline{L_j(v_i)}) \subset U_{\lambda^{-1}\epsilon}(J) \subset U_{\frac{1+2\lambda}{3\lambda}\epsilon}(A(\iota(J)))$ if $\epsilon < \epsilon_0$.

Case 2. $T_J J = \overline{R_j(v_i)}$ for some $v_i \in \mathbb{D}$ and $3 \le j \le k_i$.
In this case J coincides with either $\overline{R_{j+1}(v_c)}$ with $v_c \in \mathbb{D}$ or a subarc contained in $\overline{L_2(v_k)}$ or $[T_k, S_k]$ for some k. Therefore we have, for $\epsilon < \epsilon_0$, $T_J^{-1} U_\epsilon(\overline{R_j(v_i)}) \subset U_{\lambda^{-1}\epsilon}(J) \subset U_{\frac{1+2\lambda}{3\lambda}\epsilon}(A(\iota(J)))$ in the same way as Case 1.

Case 3. $T_J J = \overline{L_{k_i}(v_i)}$ for $v_i \in \mathbb{D}$.
In this case J is a subarc of $\overline{L_2(v_j)}$, $\overline{R_3(v_j)}$ or $[T_j, S_j]$ for some j. Thus (3.1) yields $T_J^{-1} U_{\frac{1+2\lambda}{3}\epsilon}(\overline{L_{k_i}(v_i)}) \subset U_{\frac{1+2\lambda}{3\lambda}\epsilon}(A(\iota(J)))$.

Case 4. $T_J J = \overline{L_{k_i-1}(v_i)}$ for some $v_i \in \mathbb{D}$.
In this case J coincides with either $\overline{L_{k_i}(v_c)}$ for some $v_c \in \mathbb{D}$ or a subarc contained in $\overline{L_2(v_j)}$, $\overline{R_3(v_j)}$ or $[T_j, S_j]$ for some j. From the argument above, we have only to consider the first case.

Note that we can choose $\epsilon_1 > 0$ such that

$$T_{\overline{L_{k_i}(v_c)}}(U_{\frac{1+\lambda}{2}\epsilon}) \supset U_\epsilon(\overline{L_{k_i-1}(v_i)}) \tag{3.2}$$

holds for any $\epsilon < \epsilon_1$ and for any $v_i \in \mathbb{D}$. In fact, set $M = \sup_{1\le u\le N} \sup_{|z|\le 2} |g_i''|$. Since $|(T_{\overline{L_{k_i}(v_c)}})'(z)| = |g_c'(z)| \ge 1$ in virtue of condition $(*)$, we have $|g_c'(z)| \ge 1 - \epsilon M$ for any $z \in U_\epsilon(\overline{L_{k_i}(v_c)})$ if ϵ is sufficiently small. Therefore we have

$$g_c(U_{\frac{1+\lambda}{2}\epsilon})(\overline{L_{k_i}(v_c)}) \supset U_{(1-M\frac{1+\lambda}{2}\epsilon)(\frac{1+\lambda}{2}\epsilon)}(\overline{L_{k_i-1}(v_i)}).$$

Thus (3.2) is valid.

Case 5. $T_J J = [T_i, S_i]$ for some i.
In this case, J is a subarc of $\overline{L_2(v_j)}$ or $\overline{R_3(v_j)}$ or $[T_j, S_j]$ for some j. Thus the assertion is obvious as in the above.

Case 6. $T_J J = \overline{L_2(v_i)}$ or $\overline{R_3(v_i)}$ for some $v_i \in S^1$.
In this case, J is again a subarc of $\overline{L_2(v_j)}$, $\overline{R_3(v_j)}$ or $[T_j, S_j]$ for some j. Consider only the case when $J \subset [T_j, S_j]$ since the treatment of the other cases is quite similar. Note that $[T_i, S_i)$ is itself an element in $\mathcal{P}(W)$. Set $I(i_0) = [T_i, S_i)$. By virtue of (2.3) and (2.4), there are integers $\nu > 0$, $n > 1$, $1 \le k \le \nu - 1$ and a finite sequence $I(i_1), \ldots, I(i_\nu) \in \mathcal{P}(W)$ such that T_J can be written as

$$\begin{aligned} T_J &= (g_{I(i_{k+\nu})} g_{I(i_{k+\nu-1})} \cdots g_{I(i_{k+1})})^n g_{I(i_k)} \cdots g_{I(i_1)} g_{I(i_0)} \\ &= h^n g_{I(i_k)} \cdots g_{I(i_1)} g_{I(i_0)}. \end{aligned} \tag{3.3}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Here h is a parabolic element in Γ fixing a vertex of the fundamental polygon R on S^1. Using a rotation if necessary, we may assume that the fixed vertex is 1. Since the form of h is as in (2.5), it is not hard to see that there exists $\epsilon_2 > 0$ such that $g_{I(i_0)}T_J^{-1}U_\epsilon(T_J J) \subset U_\epsilon(g_{I(i_0)}J)$ for any $\epsilon < \epsilon_2$. Thus if $\epsilon < \min(\epsilon_1, \epsilon_2)$, we have $T_J^{-1}U_\epsilon(T_J J) \subset U_{\lambda^{-1}\epsilon}(A(\iota(J))) \subset U_{\frac{1+2\lambda}{3\lambda}\epsilon}(A(\iota(J)))$. Hence the assertions (1) and (2) are established.

Recalling the typical form of the parabolic element in $PSU(1,1)$ and that of its derivative (cf. (2.5)–(2.7)), it is easy to see the validity of assertion (3) for sufficiently small ϵ. □

Remark 3.5. In a similar way to [**3**], we can construct a Markov system $\mathcal{T}_\Gamma$ associated with a cofinite group Γ satisfying all the results in §2 and §3, even if Γ does not have a fundamental polygon satisfying (∗) and (∗∗). In fact, we can construct a cofinite Fuchsian group Γ^* such that Γ^* has a fundamental polygon satisfying (∗) and (∗∗) and the signature of the Riemann surface $M^* = \mathbb{D}/\Gamma^*$ coincides with that of $M = \mathbb{D}/\Gamma$. By virtue of the theory of quasiconformal deformation, we can find a quasiconformal mapping $w : \mathbb{C} \to \mathbb{C}$ such that $j : \Gamma^* \to \Gamma$ with $j(g^*) = wg^*w^{-1}$ gives an isomorphism and $h = w|_{S^1}$ is a homeomorphism with $h(g^*x) = j(g^*)h(x)$. Then the Bowen–Series Markov map T_Γ is obtained by $T_\Gamma = hT_{\Gamma^*}h^{-1}$. Let $\mathcal{T}_{\Gamma^*} = (\mathcal{R}^*, \mathcal{P}^*, T^*)$ be the Markov system associated with Γ^* so that $\mathcal{R}^* = \{A^*\}$, $\mathcal{P}^* = \{J^*\}$ and $T^*|_{\operatorname{int} J^*} = T_{J^*}|_{\operatorname{int} J^*}$ for each $J^* \in \mathcal{P}^*$. Set $\mathcal{R} = \{hA^*\}$, $\mathcal{P} = \{hJ^*\}$ and $T_{hJ^*} = wT_{J^*}w^{-1}$. Then the triplet $\mathcal{T}_\Gamma = (\mathcal{R}, \mathcal{P}, T)$ with $T|_{\operatorname{int} hJ^*} = T_{hJ^*}|_{\operatorname{int} hJ^*}$ turns out to be the desired Markov system.

4. *$\mathcal{T}_\Gamma$-periodic orbits and hyperbolic conjugacy classes of Γ*

In the case of the modular group $PSL(2,\mathbb{Z})$, it is well known that there is a natural one-to-one correspondence between the set of its hyperbolic conjugacy classes and that of periodic orbits of the twice iteration T_G^2 of the Gauss transformation T_G acting on the unit interval. This fact enables us to study the Selberg zeta function in terms of the thermodynamic formalism for one-dimensional maps (see [**11**] and [**12**]). The aim of this section is to establish the analogous fact to this for a general cofinite group Γ and our Markov system $\mathcal{T}_\Gamma$ following the idea of Series [**20**].

Recall that for the Markov system $\mathcal{T}_\Gamma = (\mathcal{R}, \mathcal{P}, T)$, T is an a.e. defined map on the disjoint union $X = \bigsqcup_{i=1}^q A(i)$. Therefore if $A(i_1) = [a, b]$ and $A(i_2) = [b, c]$, it is necessary to distinguish the former b and the latter as points in X. Hence, it will be better to write these arcs as $A(i_1) = [a, b-]$ and $A(i_2) = [b+, a]$. In the case when there is no possibility of confusion, we may omit the sign $+$ and $-$.

Now we write $\mathcal{P} = \{J(j)\}$ for convenience. A sequence $\xi = (\xi_0\xi_1\ldots)$ of integers is said to be *admissible* if $T_{J(\xi_k)}J(\xi_k) \supset J(\xi_{k+1})$ for each $k \geq 0$. For a point $x \in X$, if there is an admissible sequence $\xi = (\xi_0\xi_1\ldots)$ with $x \in J(\xi_0)$ and $T_{J(\xi_k)} \cdots T_{J(\xi_0)}x \in J(\xi_{k+1})(k \geq 0)$, we call ξ the *$\mathcal{T}_\Gamma$-itinerary of x* (simply, the itinerary of x). We denote it as $\xi(x) = (\xi_0(x)\xi_1(x)\ldots)$ if it is unique. We note that the $\mathcal{T}_\Gamma$-itinerary of x is unique if either $T_{J(\xi_k)} \cdots T_{J(\xi_0)}x \in \operatorname{int} J(\xi_{k+1})$ holds for any $k \geq 0$ or x is an endpoint of some arc in $\mathcal{R}$. The uniqueness of the latter case follows from the fact that we define X as the disjoint union of the arcs in $\mathcal{R}$.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

In what follows the word period always means the least period. A point $x \in X$ with periodic $\mathcal{T}_\Gamma$-itinerary is called a *$\mathcal{T}_\Gamma$-periodic point*. A set of p distinct points $\tau = \{x_0, \ldots, x_{p-1}\}$ is called a *$\mathcal{T}_\Gamma$-periodic orbit of period p* if a $\mathcal{T}_\Gamma$-itinerary $\xi(x_i)$ exists for each i and if by an appropriate change of indexing of the x_i's, we can make $\xi_{k+1}(x_i) = \xi_k(x_{i+1})$ mod p hold for any $k \geq 0$. Since $\mathcal{T}_\Gamma$ satisfies the Lasota–Yorke condition, it is easy to see that for each admissible periodic sequence ξ with period p, there exists a unique $\mathcal{T}_\Gamma$-periodic point $x = x(\xi)$ whose $\mathcal{T}_\Gamma$-itinerary $\xi(x)$ coincides with ξ. Moreover, we can see that the set $\tau(x) = \{x, T_{J(\xi_0)}x, \ldots, T_{J(\xi_{p-2})}T_{J(\xi_{p-3})} \cdots T_{J(\xi_0)}x\}$ forms a $\mathcal{T}_\Gamma$-periodic orbit of period p, since each T_J is monotone increasing on J. Therefore we can define *the period of* $\mathcal{T}_\Gamma$-periodic point as the period of its $\mathcal{T}_\Gamma$-itinerary. We denote by $PO(\mathcal{T}_\Gamma)$ the totality of $\mathcal{T}_\Gamma$-periodic orbits. Put

$$\begin{aligned} PO(\mathcal{T}_\Gamma)_0 &= \{\tau \in PO(\mathcal{T}_\Gamma) : x \in \tau \Longrightarrow x \in \operatorname{int} J(\xi_0(x))\}, \\ PO(\mathcal{T}_\Gamma)_{P-} &= \{\tau \in PO(\mathcal{T}_\Gamma) : P_i- \in \tau \text{ for some } i\}, \\ PO(\mathcal{T}_\Gamma)_{Q+} &= \{\tau \in PO(\mathcal{T}_\Gamma) : Q_i+ \in \tau \text{ for some } i\}, \\ PO(\mathcal{T}_\Gamma)_{Q-} &= \{\tau \in PO(\mathcal{T}_\Gamma) : Q_i- \in \tau \text{ for some } i\}. \end{aligned}$$

Remark 4.1. The following observations using the identity (2.3) and Remark 2.3 are convenient for our investigation of $\mathcal{T}_\Gamma$-periodic orbits.

(1) Assume that Q_i- is a $\mathcal{T}_\Gamma$-periodic point with itinerary $(\xi_0\xi_1\ldots)$. We write J_j instead of $J(\xi_j)$ for simplicity. Then $J_0 = \overline{L_{k_i}(v_i)}$ for some $v_i \in \mathbb{D}$. Successively, we have $T_{J_{j-1}}J_{j-1} = \overline{L_{k_i-j}(v(j))}$, $T_{J_{j-1}}|_{\operatorname{int} J_{j-1}} = T_\Gamma|_{\operatorname{int} J_{j-1}}$ and $v(j) = T_{J_{j-1}}v(j-1)$ with $v(0) = v_i$ for $1 \leq j \leq k_i - 2$. In particular, $T_{J_{j-1}}J_{j-1} = J_j \in \mathcal{P}$ for $1 \leq j \leq k_i - 3$. J_{k_i-2} is the subarc of $\overline{L_2(v(k_i-2))}$ containing the point T_k as an endpoint, where the index k satisfies $v_k = v(k_i - 2)$. Therefore, we have seen that $T_{J_{k_i-2}}T_{J_{k_i-3}} \cdots T_{J_0}Q_i- = T_{J_{k_i-2}}T_k- = Q_\ell-$ for some ℓ with $v_{\ell-1} = v(k_i - 1)$ (see Figure 2(a)).

(2) Assume that Q_i+ is a $\mathcal{T}_\Gamma$-periodic point with itinerary $(\xi_0\xi_1\ldots)$. We use the same notation as above. Then we have $J_0 = \overline{L_{k_i-1}(v_i)}$ for some $v_i \in \mathbb{D}$. This time, we have $T_{J_{j-1}}J_{j-1} = \overline{L_{k_i-j-1}(v(j))}$, $T_{J_{j-1}}|_{\operatorname{int} J_{j-1}} = T_\Gamma|_{\operatorname{int} J_{j-1}}$ and $v(j) = T_{J_{j-1}}v(j-1)$ with $v(0) = v_i$ for $1 \leq j \leq k_i - 3$. J_{k_i-3} is the subarc of $\overline{L_2(v(k_i-3))}$ containing the point $T_{J_{k_i-3}}T_{J_{k_i-2}} \cdots T_{J_0}Q_i+$ as an endpoint and J_{k_i-2} is the subarc of $[T_k, S_k]$ containing the point T_k as an endpoint, where k satisfies $v_k = v(k_i - 2)$. Thus we have $T_{J_{k_i-2}}T_{J_{k_i-3}} \cdots T_{J_0}Q_i+ = Q_\ell+$ for some ℓ with $v_{\ell-1} = v(k_i - 1)$ (see Figure 2(b)).

(3) Assume P_i- is a $\mathcal{T}_\Gamma$-periodic point with itinerary $(\xi_0\xi_1\ldots)$. In this case $J_0 = J_{\xi_0} = \overline{R_{k_i}(v_i)}$ for some $v_i \in \mathbb{D}$. Repeatedly, we have $T_{J_{j-1}}J_{j-1} = \overline{R_{k_i-j}(v(j))}$, $T_{J_{j-1}}|_{\operatorname{int} J_{j-1}} = T_\Gamma|_{\operatorname{int} J_{j-1}}$ and $v(j) = T_{J_{j-1}}v(j-1)$ with $v(0) = v_i$ for $1 \leq j \leq k_i - 3$. In particular, we have $T_{J_{j-1}}J_{j-1} = J_j \in \mathcal{P}$ for $1 \leq j \leq k_i - 4$. J_{k_i-3} is the subarc of $\overline{R_3(v(k_i-3))}$ containing the point $T_{J_{k_i-4}}T_{J_{k_i-5}} \cdots T_{J_0}P_i-$ as an endpoint and J_{k_i-2} is a subarc of $[T_{k-1}, S_{k-1}]$ containing the point S_{k-1} as an endpoint, where k satisfies $v_k = v(k_i - 2)$. Thus we have $T_{J_{k_i-2}}T_{J_{k_i-3}} \cdots T_{J_0}P_i- = T_{J_{k_i-2}}S_{k-1}- = P_\ell-$ for some ℓ with $v_{\ell+1} = v(k_i - 1)$ (see Figure 2(c)).

(4) Since $T_{J_0}P_i+ = Q_\ell+$ for some ℓ, observation (2) implies that P_i+ cannot be a $\mathcal{T}_\Gamma$-periodic point for any i.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

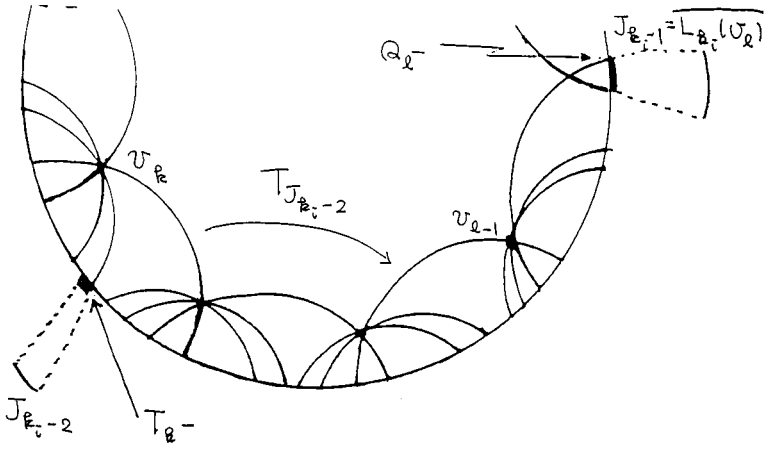


FIGURE 2(a).

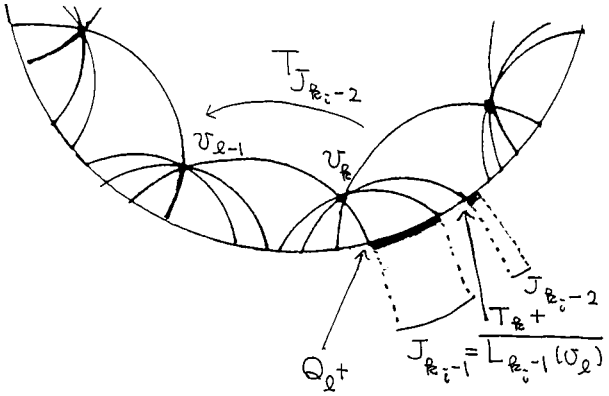


FIGURE 2(b).

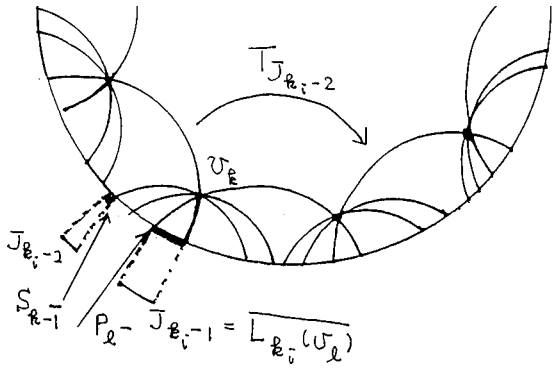


FIGURE 2(c).

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Now we state the main result in this section as the following lemma. It will enable us to obtain a determinant representation of the Selberg zeta function. Recall that $HC(\Gamma)$ denotes the totality of hyperbolic conjugacy classes, i.e. each $c \in HC(\Gamma)$ is written as $c = (h) = \{ghg^{-1} : g \in \Gamma\}$ for some primitive hyperbolic element $h \in \Gamma$.

LEMMA 4.2.

(1) *For* $\tau = \{x_0, \dots, x_{p-1}\} \in PO(\mathcal{T}_\Gamma)$ *with* $\xi(x) = (\dot{j}_0 \cdots \dot{j}_{p-1})$, $T^p_{J(j_0 j_1 \dots j_{p-1})} = T_{J(j_{p-1})} T_{J(j_{p-2})} \cdots T_{J(j_0)}$ *is a primitive hyperbolic element in* Γ.

(2) *We define a map* $\Phi : PO(\mathcal{T}_\Gamma) \to HC(\Gamma)$ *by*

$$\Phi(\tau) = (T^p_{J(j_0 j_1 \dots j_{p-1})}),$$

if τ *is written as in* (1). *Then we have:*

(2a) $\Phi(PO(\mathcal{T}_\Gamma)_0) \cup \Phi(PO(\mathcal{T}_\Gamma)_{Q+}) = HC(\Gamma)$ *(disjoint union);*

(2b) *the restrictions* $\Phi|_{PO(\mathcal{T}_\Gamma)_0}$, $\Phi|_{PO(\mathcal{T}_\Gamma)_{Q+}}$, $\Phi|_{PO(\mathcal{T}_\Gamma)_{Q-}}$ *and* $\Phi|_{PO(\mathcal{T}_\Gamma)_{P-}}$ *are injective;*

(2c) $\Phi(PO(\mathcal{T}_\Gamma)_{Q+}) = \Phi(PO(\mathcal{T}_\Gamma)_{Q-}) = \Phi(PO(\mathcal{T}_\Gamma)_{P-})$;

(2d) $\exp(\ell(\Phi(\tau))) = (T^p_{J(j_0 j_1 \dots j_{p-1})})'(x_0)$.

The rest of this section is devoted to the proof of Lemma 4.2. We divide it into several steps.

Step 1. In the following we make use of the original Markov map T_Γ and $\tilde{T}_\Gamma$ (see (2.1) and (2.2)).

First of all we can consider the following two possibilities for a primitive hyperbolic element having a repelling fixed point x:

(i) $T^n_\Gamma x \notin W$ for any $n \geq 0$;

(ii) $T^n_\Gamma x \in W$ for some $n \geq 0$.

Choose any hyperbolic conjugacy class c. We show that if (ii) holds for some $h \in c$, then it holds for any element in c. In fact, for any $g \in \Gamma$ the repelling fixed point of ghg^{-1} is clearly gx. Therefore the orbit equivalence Lemma 2.1(1) implies that we can find non-negative integers n_1 and n_2 such that $T^{n_1}_\Gamma gx = T^{n_2}_\Gamma x$. Thus we have $T^{n_1+n}_\Gamma gx = T^{n_2+n}_\Gamma x \in W$ for some n.

Step 2. Now we can prove that $c \in \Phi(PO(\mathcal{T}_\Gamma)_0)$ or $c \in \Phi(PO(\mathcal{T}_\Gamma)_{Q+})$ occurs according to (i) or (ii) in Step 1. This implies assertion (2a) in Lemma 4.2.

Choose any $h \in c$. Suppose (i) occurs. Then we can find non-negative integers m, n and a neighborhood of U of x such that $T^m_\Gamma hy = T^n_\Gamma y$ holds for any $y \in U$ by virtue of (2) in Remark 2.2. Following [**19**], we write $e_k^{-1} = g_{I(i_k)}$ if $T^k_\Gamma y \in I(i_k)$ (recall the notation at the beginning of the proof of Lemma 3.4). Then using an elementary property of linear fractional transformations, we conclude that

$$e_{m-1}^{-1} e_{m-2}^{-1} \cdots e_0^{-1} h = e_{n-1}^{-1} e_{n-2}^{-1} \cdots e_0^{-1}.$$

We have $n > m$ as follows. Since h is not the identity, n and m do not coincide. If $m > n$ holds, we have $h = (e_0 e_1 \cdots e_{n-1})(e_n e_{n+1} \cdots e_{m-1})(e_0 e_1 \cdots e_{n-1})^{-1}$. Thus we have $h'(x) = (e_n e_{n+1} \cdots e_{m-1})'((e_0 e_1 \cdots e_{n-1})^{-1}(x)) > 1$. On the other hand, $e_{n+k}^{-1} = g_{I(i_{n+k})}$ on $I(i_{n+k})$ is expanding by the condition $(*)$ for $k = 0, 1, \dots, m-n-1$. Consequently, we have $(e_n e_{n+1} \cdots e_{m-1})'((e_0 e_1 \cdots e_{n-1})^{-1}(x)) < 1$. Now we reach

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

a contradiction. Therefore $n > m$ must hold. Hence we have $e_{m-1}^{-1}e_{m-2}^{-1}\cdots e_0^{-1}h = e_{n-1}^{-1}\cdots e_m^{-1}e_{m-1}^{-1}\cdots e_0^{-1}$ and $T_\Gamma^m x = T_\Gamma^{n-m}T_\Gamma^m x$.

From the definition of $\mathcal{T}_\Gamma$, we can find an integer $0 \le k < n-m$ such that $T_\Gamma^{m+k}x$ is $\mathcal{T}_\Gamma$-periodic point and $e_{m+k-1}^{-1}\cdots e_m^{-1}e_{n-1}^{-1}\cdots e_{m+k}^{-1}$ coincides with $T^p_{J(j_0\dots j_{p-1})}$ for some p and $J(j_0\dots j_{p-1}) \in \mathcal{P}_p$. Since h is primitive, p is the $\mathcal{T}_\Gamma$-period of $T_\Gamma^{m+k}x$. In particular, we have $(h) = (T^p_{J(j_0\dots j_{p-1})})$.

Next suppose that (ii) occurs. From rule (2.3) and Remark 4.1, we may assume that $x = Q_i$ for some $v_i \in \mathbb{D}$. Then there exist non-negative integers m, n and a half-open arc $[Q_i, z)$ such that $T_\Gamma^m hy = T_\Gamma^n y$ holds for any $y \in [Q_i, z)$. Using (2) of Remark 4.1, we can find a $\mathcal{T}_\Gamma$-periodic orbit τ containing Q_i+ such that $\Phi(\tau) = (h)$ in the same way as in the case when (i) occurs.

Step 3. The injectivity of $\Phi|_{PO(\mathcal{T}_\Gamma)_0}$ and $\Phi|_{PO(\mathcal{T}_\Gamma)_{Q+}}$ in (2b) will be proved as follows. Let τ_1 and τ_2 be $\mathcal{T}_\Gamma$-periodic orbits with $\Phi(\tau_1) = \Phi(\tau_2)$. Choose any $x_1 \in \tau_1$ and $x_2 \in \tau_2$. If (i) holds for x_1, then it holds for x_2. From the construction of $\mathcal{T}_\Gamma$, we can find primitive hyperbolic elements h_1 and h_2 so that $h_1x_1 = x_1|h_1'(x_1)| > 1$, $h_2x_2 = x_2|h_2'(x_2)| > 1$ and $h_1 = gh_2g^{-1}$ for some $g \in \Gamma$ hold. Clearly we have $gx_2 = x_1$. Therefore, by virtue of the orbit equivalence, we have $T_\Gamma^{n_1}x_1 = T_\Gamma^{n_2}x_2$ for some non-negative integers n_1 and n_2. Thus x_1 and x_2 must be in the same T_Γ-orbit. In particular, (i) implies that they are in the same $\mathcal{T}_\Gamma$-orbit. Hence we have $\tau_1 = \tau_2$.

If (ii) occurs for x_1, we may assume that $Q_i+ \in \tau_1$ and $Q_j+ \in \tau_2$. Using observation (2) in Remark 4.1 we conclude that $\tau_1 = \tau_2$ in the same way as above.

Step 4. Here we are in the position to be able to show the validity of assertion (1) in Lemma 4.2. The hyperbolicity of $T^p_{J(j_0j_1\dots j_{p-1})} = T_{J(j_{p-1})}T_{J(j_{p-2})}\cdots T_{J(j_0)}$ is an easy consequence of the definition of $\mathcal{T}_\Gamma$ and (1) of Lemma 3.3 since an element in $SU(1,1)$ fixing a point in the unit circle is hyperbolic or parabolic and the modulus of the derivative at the fixed point must be one in the latter case. Therefore it remains to prove the primitivity. But in Step 2 we find a $T^p_{J(j_0j_1\dots j_{p-1})} = T_{J(j_{p-1})}T_{J(j_{p-2})}\cdots T_{J(j_0)}$ for each $c \in HC(\Gamma)$ so that it is primitive and $c = (T^p_{J(j_0j_1\dots j_{p-1})})$. Combining this with the injectivity in Step 3, we can easily see that assertion (1) is true.

Step 5. We show the validity of the other assertions. By virtue of observations (1) and (2) in Remark 4.1, it is not hard to see that there is a natural one-to-one correspondence between $PO(\mathcal{T}_\Gamma)_{Q+}$ and $PO(\mathcal{T}_\Gamma)_{Q-}$ so that $\Phi(PO(\mathcal{T}_\Gamma)_{Q+}) = \Phi(PO(\mathcal{T}_\Gamma)_{Q-})$. Next we construct $\tilde{\mathcal{T}}_\Gamma = (\mathcal{R}, \mathcal{P}, \tilde{T})$ starting with $\tilde{T}_\Gamma$ and define $\tilde{\Phi}$ in the same way as $\mathcal{T}_\Gamma$ and Φ. In this case $PO(\tilde{\mathcal{T}}_\Gamma)_0$, $PO(\tilde{\mathcal{T}}_\Gamma)_{P+}$, $PO(\tilde{\mathcal{T}}_\Gamma)_{P-}$, and $PO(\tilde{\mathcal{T}}_\Gamma)_{Q+}$ play the roles of $PO(\mathcal{T}_\Gamma)_0$, $PO(\mathcal{T}_\Gamma)_{Q-}$, $PO(\mathcal{T}_\Gamma)_{Q+}$, and $PO(\mathcal{T}_\Gamma)_{P-}$, respectively. Moreover, observation (3) in Remark 4.1 implies that $PO(\mathcal{T}_\Gamma)_{P-}$ coincides with $PO(\tilde{\mathcal{T}}_\Gamma)_{P-}$ as a set since $T_\Gamma|_{(Q_i.P_{i+1})} = \tilde{T}_\Gamma|_{(Q_i.P_{i+1})}$ holds for each i. Combining these with the fact proved in Step 3, we complete the proof of (2b) and (2c).

If $c = (h)$ and x is a repelling fixed point of h, it is well known in hyperbolic geometry that $h'(x) = \exp(\ell(c))$. Hence assertion (2d) is true. Now the proof of Lemma 4.2 is complete. □

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

5. *Transfer operators with respect to* $\mathcal{T}_\Gamma$

As mentioned in the Introduction, Mayer investigates the family of transfer operators denoted by L_s in his paper [**12**] and establishes a determinant representation of the Selberg zeta function for $PSL(2,\mathbb{Z})$. The results in [**12**] are based on the fact that the family L_s can be realized as meromorphic functions in s taking values with nuclear operators of order zero. The aim of the present section is to define an appropriate family of transfer operators $L(s)$ with respect to the Markov system $\mathcal{T}_\Gamma$ which plays the role of L_s in our investigation.

To begin with we prepare the function spaces on which our transfer operator will act. For a bounded domain D in $\mathbb{C}$, $\mathcal{A}(D)$ and $\mathcal{A}_b(D)$ denote the totality of holomorphic functions and its subspace of bounded holomorphic functions, respectively. $\mathcal{A}(D)$ is a typical example of a nuclear Fréchet space with semi-norms $p_K(f) = \|f|_K\|_\infty = \sup_{z\in K}|f(z)|$, where K is any compact subset of D. $\mathcal{A}_b(D)$ is a Banach space with supremum norm $\|f\|_\infty = \sup_{z\in D}|f(z)|$. For a compact set K, $C(K)$ denotes the Banach space of continuous functions on K with supremum norm.

For a finite number of bounded domains $D_1, D_2, \ldots, D_k$ in $\mathbb{C}$, a function defined on the disjoint union $\bigsqcup_{i=1}^k D_i$ is said to be an element of $\mathcal{A}(\bigsqcup_{i=1}^k D_i)$ or of $\mathcal{A}_b(\bigsqcup_{i=1}^k D_i)$ if $f|_{D_i} \in \mathcal{A}(D_i)$ for each i or $f|_{D_i} \in \mathcal{A}_b(D_i)$ for each i, respectively. It is clear that $\mathcal{A}(\bigsqcup_{i=1}^k D_i)$ (respectively $\mathcal{A}_b(\bigsqcup_{i=1}^k D_i)$) is naturally identified with $\bigoplus_{i=1}^k \mathcal{A}(D_i)$ (respectively $\bigoplus_{i=1}^k \mathcal{A}_b(D_i)$) by $\mathcal{A}(\bigsqcup_{i=1}^k D_i) \ni f \mapsto \oplus_{i=1}^k f_i \in \bigoplus_{i=1}^k \mathcal{A}(D_i)$, where $f_i = f|_{D_i}$. For $f = \oplus_{i=1}^k f_i \in \mathcal{A}_b(\bigsqcup_{i=1}^k D_i)$, the norm of f is given by $\|f\|_\infty = \max_{1\le i\le k}\|f_i\|_\infty$. In the same way, for a finite number of compact sets $K_1, K_2, \ldots, K_k$, $C(\bigsqcup_{i=1}^k K_i)$ denotes the totality of functions defined on $\bigsqcup_{i=1}^k K_i$ with $f|_{K_i} \in C(K_i)$ for all i and $C(\bigsqcup_{i=1}^k K_i)$ is identified with $\bigoplus_{i=1}^k C(K_i)$.

Consider the Markov system $\mathcal{T}_\Gamma = (\mathcal{R}, \mathcal{P}, T)$. Recall that $\mathcal{R} = \{A(1), A(2), \ldots, A(q)\}$ is a family of a finite number of closed arcs in S^1 such that $\sharp(A(i) \cap A(j)) \le 1$ if $i \ne j$. Let $\mathcal{B} = \{B(1), B(2), \ldots, B(q)\}$ and $\mathcal{D} = \{D(1), D(2), \ldots, D(q)\}$ be the family of domains in $\mathbb{C}$ constructed in Lemma 3.4. Set $X = \bigsqcup_{i=1}^q A(i)$, $B(X) = \bigsqcup_{i=1}^q B(i)$ and $D(X) = \bigsqcup_{i=1}^q D(i)$. For $J \in \mathcal{P}$ and complex numbers w and s, we set

$$G_J(s)(w) = \left(T_J'(w)\frac{w}{T_J w}\right)^{-s}.$$

Since it is easy to see that $T_J'(w)(w/T_J w) = |T_J'(w)|$ for $w \in S^1$, the right-hand side of the above is well defined and gives a holomorphic function in $(w,s) \in B(\iota(J)) \times \mathbb{C}$ if we choose the $B(j)$'s sufficiently small. For an element $f = \oplus_{i=1}^q f_i$ in $\mathcal{A}(D(X))$ or $C(X)$, we define formally a family of operators as

$$(L(s)f)_i(z) = \sum_{J\in\mathcal{P}:\tau(J)=i} G_J(s)(T_J^{-1}z) f_{\iota(J)}(T_J^{-1}z).$$

In terms of the a.e. defined map $T : X \to X$, the operator $L(s)$ has the form of the transfer operator

$$\text{`}\mathcal{L}(s)f(x) = \sum_{Ty=x} |T'y|^{-s} f(y)\text{'}.$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

We call the formally defined operators $L(s)$ *the transfer operators with respect to the Markov system* $\mathcal{T}_\Gamma$.

Next we will express the iteration of $L(s)$. For $J = J(i_0 1_1 \dots i_{n-1}) \in \mathcal{P}_n$, we set

$$G_{n,J}(s)(w) = \left((T_J^n)'(w) \frac{w}{T_J^n w} \right)^{-s}. \tag{5.1}$$

From the definition (2.12) of T_J^n we see that

$$\begin{aligned} G_{n,J}(s)(w) &= G_{J(i_0)}(s)(w) G_{J(i_1)}(s)(T_{J(i_0)}w) G_{J(i_2)}(s)(T^2_{J(i_0 i_1)}w) \\ &\quad \cdots G_{J(i_{n-1})}(s)(T^{n-1}_{J(i_0 i_1 \dots i_{n-2})}w). \end{aligned} \tag{5.2}$$

Therefore $L(s)^n$ is expressed as

$$(L(s)^n f)_i(z) = \sum_{J \in \mathcal{P}_n : \tau(J) = i} G_{n,J}(s)(T_J^{-n} z) f_{\iota(J)}(T_J^{-n} z), \tag{5.3}$$

for $f = \oplus_{i=1}^q f_i$, where T_J^{-n} denotes the inverse of T_J^n.

For $s \in \mathbb{C}$ with $\Re s > 1/2$, the formally defined operators $L(s)$ can be realized as elements in $\mathcal{L}(\mathcal{A}_b(B(X)) \to \mathcal{A}_b(D(X)))$, $\mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$, and $\mathcal{L}(C(X) \to C(X))$ by virtue of Lemma 3.4, where $\mathcal{L}(\mathcal{X} \to \mathcal{Y})$ is the space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ for linear topological spaces $\mathcal{X}$ and $\mathcal{Y}$. We distinguish these realizations and use the notation $L_0(s)$, $L(s)$, and $L_X(s)$ for the realizations in $\mathcal{L}(\mathcal{A}_b(B(X)) \to \mathcal{A}_b(D(X)))$, $\mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$, and $\mathcal{L}(C(X) \to C(X))$, respectively. $L_0(s)$ and $L(s)$ are related by $L(s) = L_0(s) \circ r$, where $r : \mathcal{A}_b(D(X)) \to \mathcal{A}_b(B(X))$ is the restriction operator.

Note that r is a nuclear operator of order zero in the sense of Grothendieck [**7**]. In fact, r is expressed as

$$r = \sum_{n=1}^{\infty} u_n \otimes v_n,$$

where $u_n \in \mathcal{A}_b(D(X))'$ and $v_n \in \mathcal{A}_b(B(X))$ are such that the norm $\|u_n\|_\infty$ as a bounded linear functional and the norm $\|u_n\|_\infty$ are rapidly decreasing as n tends to infinity. More precisely, if $\Psi_i : D(i) \to \mathbb{D}$ is the Riemann mapping and $0 < \rho_1 < \rho_2 < 1$ are chosen so that $|\Psi(z)| < \rho_1$ for any $z \in \overline{B(i)}$, then we can write

$$r(f)_i(z) = \sum_{k=0}^{\infty} \frac{1}{2\pi\sqrt{-1}} \int_{|\zeta|=\rho_2} \frac{\rho_1^k f_i(\Psi_i^{-1}(\zeta))}{\zeta^{k+1}} \, d\zeta \left(\frac{\Psi_i(z)}{\rho_1} \right)^k$$

for $f = \oplus_{i=1}^q f_i$. Then we can write

$$L(s) = \sum_{n=1}^{\infty} u_n \otimes L_0(s) v_n. \tag{5.4}$$

The main result in this section is the following theorem.

THEOREM 5.1. *In the domain* $\{s \in \mathbb{C} : \Re s > 1/2\}$, $L_0(s)$ *is an* $\mathcal{L}(\mathcal{A}_b(B(X)) \to \mathcal{A}_b(D(X)))$*-valued analytic function having a meromorphic extension to the entire s-plane. Moreover, there exists a sequence* $\{L_{0,k}\}_{k=-1,0,1,2,\dots}$ *of elements in* $\mathcal{L}(\mathcal{A}_b(B(X)) \to$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

$\mathcal{A}_b(D(X)))$ of rank at most q^2 such that

$$L_0(s) - \left(\frac{L_{0,-1}}{s - \frac{1}{2}} + \frac{L_{0,0}}{s} + \frac{L_{0,1}}{s + \frac{1}{2}} + \cdots + \frac{L_{0,k}}{s + \frac{k}{2}} \right)$$

is an $\mathcal{L}(\mathcal{A}_b(B(X)) \to \mathcal{A}_b(D(X)))$-valued analytic function in the domain $\{s \in \mathbb{C} : \Re s > -(k+1)/2\}$ for each $k \geq -1$.

Before proving the theorem, we state a corollary.

COROLLARY 5.2. *An analytic function $\{s \in \mathbb{C} : \Re s > 1/2\} \ni s \mapsto L(s) \in \mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$ has a meromorphic extension to the entire s-plane which takes values in the space of nuclear operators of order zero. The candidates for poles are the points $s = -k/2$, $k = -1, 0, 1, 2, \ldots$. In particular, the Fredholm determinant $\mathrm{Det}(I - L(s))$ extends to a meromorphic function on the entire s-plane and the candidates for its poles are the same as those of $L(s)$, possibly with different orders.*

Proof of Corollary 5.2. First we recall the facts stated in [**8**] (see also [**15**, Chapter 4]). Suppose that an operator L on a Banach space $\mathcal{X}$ has the form

$$L = \sum_{i=1}^{\infty} x_i \otimes y_i,$$

where x_i and y_i are elements in $\mathcal{X}'$ and $\mathcal{X}$ respectively with rapidly decreasing norms as $i \to \infty$. Then $\mathrm{Tr} \wedge^n L$ satisfies the Hadamard inequality

$$|\mathrm{Tr} \wedge^n L| \leq \frac{n^{n/2}}{n!} \|L\|_1^n,$$

where $\wedge^n L$ is the n-fold exterior product of L acting on $\bigwedge^n \mathcal{X}$ and $\|L\|_1$ is the norm defined as the infimum of $\sum_{n=1}^{\infty} \|u_n\| \|v_n\|$ taken over all $\{u_n\}$ with $u_n \in \mathcal{X}'$ and $\{v_n\}$ with $v_n \in \mathcal{X}$ satisfying $L = \sum_{n=1}^{\infty} u_n \otimes v_n$ and $\sum_{n=1}^{\infty} \|u_n\| \|v_n\| < \infty$. We note that $\mathrm{Tr}\, L = \sum_{n=1}^{\infty} u_n(v_n)$ does not depend on the choice of $\{u_n\}$ and $\{v_n\}$ if the Banach space $\mathcal{X}$ has the approximation property. With the help of the Hadamard inequality we can express the Fredholm determinant $\mathrm{Det}(I - L)$ as

$$\mathrm{Det}(I - L) = \sum_{n=0}^{\infty} (-1)^n \mathrm{Tr} \wedge^n L. \tag{5.5}$$

Now apply the above result to $L(s)$. Since $L(s)$ has the representation (5.4), the trace $\mathrm{Tr} \wedge^n L(s)$ of an n-fold exterior product is analytic in $\{s : \Re s > 1/2\}$. Combining this fact, the Hadamard inequality and formula (5.5), we can see that $\mathrm{Det}(I - L(s))$ is analytic in $\{s : \Re s > 1/2\}$. Next let $\tilde{s}$ be a pole of $L_0(s)$. By virtue of Theorem 5.1, $L(s)$ can be written as $L(s) = (s - \tilde{s})^{-1} \tilde{L} + M(s)$ in a neighborhood U of $\tilde{s}$ in $\mathbb{C}$, where $\tilde{L}$ is an element in $\mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$ of rank at most q^2 and $U \ni s \mapsto M(s) \in \mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$ is an analytic function taking values with nuclear operators of order zero. Thus by the Hadamard inequality, we have

$$|\mathrm{Tr} \wedge^n L(s)| \leq \frac{n^{n/2}}{n!} \sum_{j=0}^{q^2} \binom{n}{j} |s - \tilde{s}|^j \|\tilde{L}\|_1^j \|M(s)\|_1^{n-j} \tag{5.6}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

for $n \geq q^2$ in $U \setminus \{\tilde{s}\}$. Therefore combining (5.5) with (5.6) we can conclude that $(s-\tilde{s})^{q^2}\,\mathrm{Det}(I-L(s))$ is an analytic function in U. Hence we obtain the corollary. □

In the rest of this section we give a proof of Theorem 5.1.

Proof of Theorem 5.1. Using the identification $f \in \mathcal{A}_b(D(X))$ with $\oplus_{j=1}^q f_j \in \bigoplus_{j=1}^q \mathcal{A}_b(D(j))$, $f_j = f|_{D(j)}$ introduced above, $L_0(s)$ can be written as

$$L_0(s)f = \bigoplus_{j=1}^{q}\left(\sum_{i=1}^{q} M_{ij}(s)f_i\right),$$

where $M_{ij}(s) : \mathcal{A}_b(B(i)) \to \mathcal{A}_b(D(j))$ is the operator defined by

$$M_{ij}(s)g(z) = \sum_{J\in\mathcal{P}:\iota(J)=i,\tau(J)=j} G_J(s)(T_J^{-1}z)g(T_J^{-1}z)$$

for $g \in \mathcal{A}_b(B(i))$.

Put $m_{ij} = \sharp\{J \in \mathcal{P} : \iota(J) = i \text{ and } \tau(J) = j\}$. If $m_{ij} < \infty$, $M_{ij}(s)$ is an analytic function with values in $\mathcal{L}(\mathcal{A}_b(B(i)) \to \mathcal{A}_b(D(j)))$. Therefore our goal is to show that $M_{ij}(s)$ has meromorphic extension to the entire s-plane if $m_{ij} = \infty$.

Consider a pair of indices i and j with $m_{ij} = \infty$. By (2.4), (3.3) and the definition of the Markov system $\mathcal{T}_\Gamma$, $A(j)$ must be $\overline{L_2(v_k)}$ or $\overline{R_3(v_k)}$ for some parabolic vertex v_k in S^1 and $M_{ij}(s)$ has the form

$$M_{ij}(s) = \sum_{n=0}^{\infty}\left(\frac{(G^{-1}H^{-n})'(z)z}{G^{-1}H^{-n}z}\right)^s f(G^{-1}H^{-n}z) \tag{5.7}$$

for $f \in \mathcal{A}_b(B(i))$, where G is an element in Γ and H is a parabolic element in Γ fixing v_k (see Figure 3). Since our arguments will be carried out fixing such indices, we may drop the indices i and j, i.e. we write B instead of $B(i)$ and D instead of $D(j)$. Moreover, we may assume that H fixes 1, $v_k = 1$. In addition, we assume that B and D are chosen to be small enough so that $\overline{D}$ does not contain 1 and 0. Such a choice of B and D is guaranteed by Lemma 3.4.

Recalling the matrix representation (2.5), we can write

$$G^{-1} = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \text{ and } H^{-1} = \begin{pmatrix} 1+\sqrt{-1}\,b & -\sqrt{-1}\,b \\ \sqrt{-1}\,b & 1-\sqrt{-1}\,b \end{pmatrix}$$

for some $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$ and $b \in \mathbb{R}$. Therefore $G^{-1}H^{-n}$ has the form

$$G^{-1}H^{-n} = \begin{pmatrix} \alpha(n) & \beta(n) \\ \overline{\beta(n)} & \overline{\alpha(n)} \end{pmatrix},$$

where

$$\alpha(n) = \alpha + \sqrt{-1}\,nb(\alpha+\beta), \quad \beta(n) = \beta - \sqrt{-1}\,nb(\alpha+\beta) \tag{5.8}$$

We write, for $z \in D$,

$$\begin{aligned} \frac{G^{-1}H^{-n}z}{(G^{-1}H^{-n})'(z)\cdot z} &= (\alpha(n)z+\beta(n))(\overline{\beta(n)}z+\overline{\alpha(n)})z^{-1} \\ &= A(z)n^2 + B(z)n + C(z), \end{aligned}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

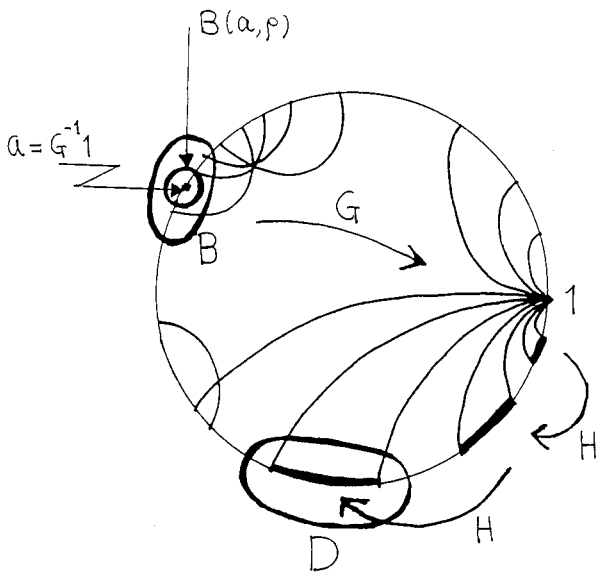


FIGURE 3.

where

$$\begin{aligned} A(z) &= b^2|\alpha+\beta|^2(2-(z+z^{-1})) \\ B(z) &= 2b(2\Im(\alpha\bar{\beta})+\sqrt{-1}\,(\alpha\bar{\beta}z-\bar{\alpha}\beta z^{-1})+\sqrt{-1}\,(|\alpha|^2+|\beta|^2)(z-z^{-1})) \\ C(z) &= |\alpha|^2+|\beta|^2+\alpha\bar{\beta}z+\bar{\alpha}\beta z^{-1}. \end{aligned} \tag{5.9}$$

Note that

$$A(z)>0, \quad C(z)>0, \quad \text{and} \quad A(z)n^2+B(z)n+C(z)>0, \quad \text{for } n\geq 0 \text{ and for any } z\in S^1\cap D.$$

Therefore, if D is small enough, it is easy to see that there is a positive constant C_1 independent of $z\in D$ such that

$$\left|\frac{(G^{-1}H^{-n})'(z)\cdot z}{G^{-1}H^{-n}z}\right| = |(A(z)n^2+B(z)n+C(z))^{-1}| \leq C_1 n^{-2} \tag{5.10}$$

holds for any non-negative integer n.

Put $a=G^{-1}1(=G^{-1}H^{-1}1\in B)$. Choose $\rho>0$ so small that the open ball $B(a,\rho)$ of radius ρ centered at a is contained in B. It follows from (5.8) that there is a positive integer C_2 satisfying

$$|G^{-1}H^{-n}z-a| = |G^{-1}H^{-n}z-G^{-1}H^{-n}1| \leq C_2 n^{-1} \tag{5.11}$$

for all $z\in D$. Consequently, we can find a positive integer $n(\rho)$ such that $G^{-1}H^{-n}D\subset B(a,\rho)\subset B$ holds for any $n\geq n(\rho)$. Therefore, if $f\in\mathcal{A}_b(B)$ and $n\geq n(\rho)$, we obtain

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

the Taylor expansion

$$f(G^{-1}H^{-n}z) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(G^{-1}H^{-n}z - a)^k. \tag{5.12}$$

Now we return to the operator $M(s) = M_{ij}(s) : \mathcal{A}_b(B) \to \mathcal{A}_b(D)$. Choose any positive integer ℓ. If $\Re s > 1/2$, the sum in (5.7) is convergent absolutely and uniformly in $z \in D$ by the inequality (5.10). Applying (5.9) and (5.12) to (5.7), we can write

$$\begin{aligned} M(s)f(z) &= \sum_{n=0}^{n(\rho)-1} (A(z)n^2 + B(z)n + C(z))^{-s} f(G^{-1}H^{-n}z) \\ &\quad + \sum_{n=n(\rho)}^{\infty} (A(z)n^2 + B(z)n + C(z))^{-s} f(G^{-1}H^{-n}z) \\ &= \sum_{n=0}^{n(\rho)-1} (A(z)n^2 + B(z)n + C(z))^{-s} f(G^{-1}H^{-n}z) \\ &\quad + \sum_{k=0}^{\ell} \frac{1}{k!} f^{(k)}(a) \sum_{n=n(\rho)}^{\infty} (A(z)n^2 + B(z)n + C(z))^{-s} (G^{-1}H^{-n}z - a)^k \\ &\quad + \sum_{k=\ell+1}^{\infty} \frac{1}{k!} f^{(k)}(a) \sum_{n=n(\rho)}^{\infty} (A(z)n^2 + B(z)n + C(z))^{-s} (G^{-1}H^{-n}z - a)^k \\ &= I_\ell(s)f(z) + II_\ell(s)f(z) + III_\ell(s)f(z). \end{aligned}$$

Clearly, $I_\ell(s)$ is an $\mathcal{L}(\mathcal{A}_b(B) \to \mathcal{A}_b(D))$-valued entire function in s. By virtue of the estimates (5.10) and (5.11), it follows that

$$\sum_{n=n(\rho)}^{\infty} |A(z)n^2 + B(z)n + C(z)|^{-\Re s} |G^{-1}H^{-n}z - a|^{\ell+1} \le C_1^{2\Re s} C_2^{\ell+1} \sum_{n=n(\rho)}^{\infty} n^{-2\Re s - \ell - 1}.$$

Thus we conclude that $III_\ell(s)$ is an $\mathcal{L}(\mathcal{A}_b(B) \to \mathcal{A}_b(D))$-valued analytic function in the domain $\{s \in \mathbb{C} : \Re s > \ell/2\}$.

It remains to investigate $II_\ell(s)$. To this end we set, for a non-negative integer m,

$$P_m(s, z) = \sum_{n=0}^{\infty} (A(z)n^2 + B(z)n + C(z))^{-s} \left(\frac{\alpha(n)z + \beta(n)}{\overline{\beta(n)}z + \overline{\alpha(n)}} \right)^m. \tag{5.13}$$

Looking at the form of $II_\ell(s)$ and comparing with (5.7), it is not hard to see that the next lemma completes the proof of Theorem 5.1.

LEMMA 5.3. *Let $P_m(s, z)$ be as in* (5.13)*. Then there exists a sequence $\{C_{m,k}\}_{k=-1,0,1,\dots}$ of functions in $\mathcal{A}_b(D)$ such that*

$$P_m(s, z) - \left(\frac{C_{m,-1}(z)}{s - \frac{1}{2}} + \frac{C_{m,0}(z)}{s} + \frac{C_{m,1}(z)}{s + \frac{1}{2}} + \cdots + \frac{C_{m,k}(z)}{s + \frac{k}{2}} \right)$$

is analytic in the domain $\{(s, z) \in \mathbb{C} \times \mathbb{C} : \Re s > -(k + 1)/2, z \in D\}$.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Proof. First of all we note that if the domain D is sufficiently small, then we can find positive numbers C_3 and C_4 satisfying

$$\sup_{z\in D}\sup_{x\in\mathbb{R}}\left|\frac{\alpha(x)z+\beta(x)}{\overline{\beta(x)}z+\overline{\alpha(x)}}\right|\le C_3 \quad\text{and}\quad \inf_{z\in D}\inf_{x\in\mathbb{R}}|\overline{\beta(x)}z+\overline{\alpha(x)}|\ge C_4, \tag{5.14}$$

where $\alpha(x)$ and $\beta(x)$ are quantities obtained by replacing n by x. Therefore we see that $P_m(s,z)$ is analytic in $(s,z)\in\{s:\Re s>1/2\}\times D$ by virtue of the inequalities (5.10) and (5.14).

Suppose $\Re s>1/2$ and multiply the Gamma function $\Gamma(s)=\int_0^\infty e^{-t}t^{s-1}\,dt$ by $P_m(s,z)$ as follows:

$$\begin{aligned}\Gamma(s)P_m(s,z) &= \int_0^\infty\sum_{n=0}^\infty e^{-(A(z)n^2+B(z)n+C(z))t}\left(\frac{\alpha(n)z+\beta(n)}{\overline{\beta(n)}z+\overline{\alpha(n)}}\right)^m t^{s-1}\,dt\\ &= \left(\int_0^1+\int_1^\infty\right)\sum_{n=0}^\infty e^{-(A(z)n^2+B(z)n+C(z))t}\left(\frac{\alpha(n)z+\beta(n)}{\overline{\beta(n)}z+\overline{\alpha(n)}}\right)^m t^{s-1}\,dt\\ &= I(s,z)+II(s,z).\end{aligned} \tag{5.15}$$

Then applying integration by parts to $II(s,z)$ repeatedly, we can easily show that $II(s,z)$ is analytic in $\mathbb{C}\times D$.

Before we investigate $I(s,z)$ in (5.15), we prepare the following fact. Let u be a function of bounded variation on the real line. Assume that there is a Lebesgue integrable function v such that v is decreasing on $[0,\infty)$ and $v(x)=v(-x)$ and $|u(x)|\le v(x)$ holds for any $x\in\mathbb{R}$. Then the Poisson summation formula

$$\sum_{n\in\mathbb{Z}}\bar u(x)=\lim_{n\to\infty}\sum_{|k|\le n}\hat u(k)\,e^{2\pi\sqrt{-1}kx} \tag{5.16}$$

holds for $x\in\mathbb{R}$, where

$$\bar u(x)=\frac{u(x-)+u(x+)}{2}\quad\text{and}\quad \hat u(x)=\int_{-\infty}^\infty u(y)\,e^{-2\pi\sqrt{-1}xy}\,dy.$$

We just note that this is easily proved by combining the usual proof of the Poisson summation formula with a theorem of Jordan concerning the convergence of Fourier series.

Set

$$Q_m(t,z)=\sum_{n=0}^\infty e^{-(A(z)n^2+B(z)n+C(z))t}\left(\frac{\alpha(n)z+\beta(n)}{\overline{\beta(n)}z+\overline{\alpha(n)}}\right)^m.$$

Applying the Poisson summation formula (5.16) at $x=0$ to the function

$$u(x)=\begin{cases} e^{-(A(z)x^2+B(z)x+C(z))t}\left(\dfrac{\alpha(x)z+\beta(x)}{\overline{\beta(x)}z+\overline{\alpha(x)}}\right)^m & (x>0)\\ 0 & (x\le 0),\end{cases}$$

we have

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

$$\begin{aligned} Q_m(t,z) - \frac{1}{2}\, e^{-C(z)t} \left(\frac{\alpha z+\beta}{\overline{\beta}z+\overline{\alpha}}\right)^m \\ = \int_0^\infty e^{-(A(z)x^2+B(z)x+C(z))t} \left(\frac{\alpha(x)z+\beta(x)}{\overline{\beta(x)}z+\overline{\alpha(x)}}\right)^m dx \\ +2\sum_{n=1}^{\infty}\int_0^\infty e^{-(A(z)x^2+B(z)x+C(z))t} \left(\frac{\alpha(x)z+\beta(x)}{\overline{\beta(x)}z+\overline{\alpha(x)}}\right)^m \cos 2\pi nx\, dx. \end{aligned} \tag{5.17}$$

Using (5.8), we set

$$\frac{\alpha(x)z+\beta(x)}{\overline{\beta(x)}z+\overline{\alpha(x)}} = \frac{a(z)x+b(z)}{c(z)x+d(z)}$$

with

$$\begin{aligned} a(z) &= \sqrt{-1}\, b(\alpha+\beta)(z-1), \quad b(z) = \alpha z+\beta, \\ c(z) &= \sqrt{-1}\, b(\bar{\alpha}+\bar{\beta})(z-1), \quad d(z) = \bar{\beta}z+\bar{\alpha}. \end{aligned}$$

Since $1 \notin \overline{D}$, $c(z) \neq 0$ in D. Put

$$A(z)x^2+B(z)x+C(z) = U(z)(c(z)x+d(z))^2 + V(z)(c(z)x+d(z)) + W(z)$$

with

$$\begin{aligned} U(z) &= \frac{A(z)}{c(z)^2}, \quad V(z) = \frac{c(z)B(z)-2d(z)A(z)}{c(z)^2}, \\ W(z) &= \frac{c(z)^2C(z)-c(z)d(z)B(z)+2d(z)^2A(z)}{c(z)^2}. \end{aligned}$$

Now we consider

$$F(k,n,t,z) = \int_0^\infty e^{-(U(z)(c(z)x+d(z))^2+V(z)(c(z)x+d(z)))t}(c(z)x+d(z))^{-k} \cos 2\pi nx\, dx$$

for $k = -1, 0, 1, 2, \ldots$ and $n = 0, 1, 2, \ldots$. In what follows we drop the letter z for the sake of simplicity. Differentiating $F(k,n,t)$ at $t > 0$, we obtain the relation

$$\frac{dF(k,n,t)}{dt} = -UF(k-2,n,t) - VF(k-1,n,t). \tag{5.18}$$

Therefore in order to obtain the asymptotic expansion of $F(k,n,t)(k \geq -1)$ for small $t > 0$, we can reduce our problem to that of $F(0,n,t)$ and $F(-1,n,t)$. By changing the variable x to $x/\sqrt{t}$, we have

$$F(-1,n,t) = \frac{1}{t}\int_0^\infty e^{-U(cx+\sqrt{t}d)^2-V(cx\sqrt{t}+dt)}(cx+\sqrt{t}d)\cos\frac{2\pi nx}{\sqrt{t}}\, dx \tag{5.19}$$

and

$$F(0,n,t) = \frac{1}{\sqrt{t}}\int_0^\infty e^{-U(cx+\sqrt{t}d)^2-V(cx\sqrt{t}+dt)}\cos\frac{2\pi nx}{\sqrt{t}}\, dx. \tag{5.20}$$

First we consider the case when $n \geq 1$. Both $tF(-1,n,t)$ and $\sqrt{t}F(0,n,t)$ have the form

$$\int_0^\infty R(x,t)\, e^{-S(x,t)} \cos\frac{2\pi nx}{\sqrt{t}}\, dx,$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

where $R(x,t)$ and $S(x,t)$ are polynomials of x and $\sqrt{t}$ with coefficients in $\mathcal{A}_b(D)$. If we denote by $R_k(x,t)\,e^{-S(x,t)}$ the kth derivative $(\partial^k/\partial x^k)(R(x,t)\,e^{-S(x,t)})$, then $R_k(x,t)$ is also a polynomial of x and $\sqrt{t}$ with coefficients in $\mathcal{A}_b(D)$. By using integration by parts repeatedly, it is easy to show that

$$\begin{aligned}\int_0^\infty R(x,t)\,e^{-S(x,t)}\cos\frac{2\pi nx}{\sqrt{t}}\,dx &= \sum_{j=1}^{\ell}\frac{(-1)^j t^j}{(2\pi n)^{2j}}R_{2j-1}(0,t)\,e^{-S(0,t)} \\ &\quad+\frac{(-1)^\ell t^\ell}{(2\pi n)^{2\ell}}\int_0^\infty R_{2\ell}(x,t)\,e^{-S(x,t)}\cos\frac{2\pi nx}{\sqrt{t}}\,dx. \qquad (5.21)\end{aligned}$$

Combining (5.19)–(5.21), we can find sequences $\{G(-1,n,j)\}_{j=0,1,2,\dots}$ and $\{G(0,n,j)\}_{j=0,1,2,\dots}$ of functions in $\mathcal{A}_b(D)$ and a sequence $\{C_{(j)}\}_{j=0,1,2,\dots}$ of positive numbers such that

$$\sup_{n\ge1}\sup_{0<t\le1}\sup_{z\in D}\left|F(-1,n,t,z)-\sum_{i=0}^{j}G(-1,n,i,z)t^{i/2}\right|n^2t^{-(j+1)/2}\le C_{(j)} \qquad (5.22)$$

and

$$\sup_{n\ge1}\sup_{0<t\le1}\sup_{z\in D}\left|F(0,n,t,z)-\sum_{i=0}^{j}G(0,n,i,z)t^{i/2}\right|n^2t^{-(j+1)/2}\le C_{(j)} \qquad (5.23)$$

hold.

Next consider the case of $n=0$ in (5.19) and (5.20). Using complex integration, we have

$$F(-1,0,t)=\frac{1}{ct}e^{-V^2t/4U}\left(\int_0^\infty\left(y-\frac{V\sqrt{t}}{2U}\right)e^{-Uy^2}\,dy-\int_{C(t)}\left(y-\frac{V\sqrt{t}}{2U}\right)e^{-Uy^2}\,dy\right)$$

and

$$F(0,0,t)=\frac{1}{c\sqrt{t}}e^{-V^2t/4U}\left(\int_0^\infty e^{-Uy^2}\,dy-\int_{C(t)}e^{-Uy^2}\,dy\right),$$

where $\int_{C(t)}$ means the integral from the origin to the point $(d+(V/2U))\sqrt{t}$ along the line segment joining them. Therefore we can conclude that there exist sequences $\{G(-1,0,j)\}_{j=-2,-1,0,\dots}$ and $\{G(0,0,j)\}_{j=-1,0,1,\dots}$ of functions in $\mathcal{A}_b(D)$ and a sequence $\{C'_{(j)}\}_{j=-1,0,1,\dots}$ of positive numbers such that

$$\sup_{0<t\le1}\sup_{z\in D}\left|F(-1,0,t,z)-\sum_{i=-2}^{j}G(-1,0,i,z)t^{i/2}\right|t^{-(j+1)/2}\le C'_{(j)} \qquad (5.24)$$

and

$$\sup_{0<t\le1}\sup_{z\in D}\left|F(0,0,t,z)-\sum_{i=-1}^{j}G(0,0,i,z)t^{i/2}\right|t^{-(j+1)/2}\le C'_{(j)} \qquad (5.25)$$

hold. Thus combining the formulae (5.17), (5.18) and (5.22)–(5.25), we conclude that there exist sequences $\{Q_{m,j}\}_{j=0,1,2,\dots}$ of functions in $\mathcal{A}_b(D)$ and a sequence $\{C''_{(j)}\}_{j=1,2,\dots}$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

of positive numbers such that

$$\sup_{0<t\le 1}\sup_{z\in D}\left|Q_m(t,z)-Q_{m,0}\log t-\sum_{i=1}^{j}Q_{m,i}t^{i/2}\right|t^{-(j+1)/2}\le C''_{(j)}.$$

Recalling the definition of $I(s,z)$ in formula (5.15) and combining it with the above expansion, we can find a sequence $\{P_{m,j}\}_{j=-2,-1,0,1,\dots}$ of functions in $\mathcal{A}_b(D)$ such that

$$\Gamma(s)P_m(s,z)-\frac{P_{m,-2}(z)}{s^2}-\left(\sum_{j=-1}^{k}\frac{P_{m,j}(z)}{s+(j/2)}\right)$$

is analytic in the domain $\{s\in\mathbb{C}:\Re s>-(k+1)/2\}\times D$. Since $\Gamma(s)$ is a meromorphic function without zero having only simple poles $0,-1,-2,\dots$, we obtain the desired result. □

6. *Spectral properties of $L(s)$ with $\Re s>1/2$*

In this section we study the spectral properties of the transfer operators $L(s)$ with $\Re s>1/2$. The main result is a Ruelle–Perron–Frobenius type uniform ergodic theorem (Theorem 6.4). In order to simplify our investigation, we impose Condition 3.1 introduced in §3 on the Markov system $\mathcal{T}_\Gamma$. By virtue of Proposition 3.2, this is not an essential restriction since we can consider the Markov system $(\mathcal{R}^{(0)},\mathcal{P}_d^{(0)},T^d)$ for some d instead of $\mathcal{T}_\Gamma$.

Consider the subset

$$\Lambda=\{f\in\mathcal{A}(D(X)):f|_X\ge 0,\|f\|_\infty\le 1\}$$

of the Fréchet space $\mathcal{A}(D(X))$. Then we have the following lemma.

LEMMA 6.1. *Assume that the Markov system $\mathcal{T}_\Gamma$ satisfies Condition 3.1. Then for each $\sigma>1/2$, the operator $L(\sigma)$ acting on the Banach space $\mathcal{A}_b(D(X))$ has a maximal eigenvalue $\lambda(\sigma)$ in the sense that $\lambda(\sigma)$ is positive and the modulus of the other eigenvalue is less than or equal to $|\lambda(\sigma)|$. In addition, we can choose an eigenfunction f belonging to the maximal eigenvalue $\lambda(\sigma)$ so that it is positive valued on X.*

Proof. Since $T_J^{-1}D(\tau(J))\subset B(\iota(J))\subset D(\iota(J))$ holds for any $J\in\mathcal{P}$ by Lemma 3.4, it follows that $L(s)$ with $\Re s>1/2$ can be regarded as a bounded operator from the Fréchet space $\mathcal{A}(D(X))$ into the Banach space $\mathcal{A}_b(D(X))$. Thus $L(\sigma)|_\Lambda:\Lambda\to\mathcal{A}_b(D(X))$ is continuous. For each n, we define a continuous map $\mathcal{F}_n:\Lambda\to\Lambda$ by

$$\mathcal{F}_n(f)=\frac{L(\sigma)(f+(1/n))}{\|L(\sigma)(f+(1/n))\|_\infty}.$$

Since Λ is clearly a compact and convex set in the Fréchet space $\mathcal{A}(D(X))$ containing the constant function 1, the Schauder–Tychonoff Theorem [**4**, V. 10] implies that there exists a function $f_n\in\Lambda$ with $\mathcal{F}_n(f_n)=f_n$, i.e.

$$L(\sigma)\left(f_n+\frac{1}{n}\right)=\lambda_n f_n,\quad \lambda_n=\left\|L(\sigma)\left(f_n+\frac{1}{n}\right)\right\|_\infty. \tag{6.1}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Let z be a point with $f_n(z) = \min_{x\in X} f_n(x)$. Then from (6.1) we have $\alpha(\sigma)(f_n(z) + (1/n)) \leq \lambda_n f_n(z)$, where $\alpha(\sigma) = \min_{x\in X} L(\sigma)1$. Therefore we obtain $f_n(z) > 0$ and $\lambda_n > 0$. Since the sequence $\{L(\sigma)f_n\}_{n=1,2,\dots}$ is relatively compact in $\mathcal{A}_b(D(X))$ by virtue of Corollary 5.2, so is the sequence $\{f_n\}_{1,2,\dots}$ from (6.1). Choosing a subsequence if necessary, we may assume that f_n converges to a function $f \in \mathcal{A}_b(D(X))$ uniformly in $D(X)$ and λ_n converges to a number λ as n tends to infinity. Since $\|L(\sigma)f\|_\infty = \lambda \geq \alpha(\sigma)$ from (6.1), f is a non-trivial solution of $L(\sigma)f = \lambda f$.

Next we show that $f|_X$ is positive everywhere on X. Making use of Condition 3.1, it is not hard to see that there is a positive integer n_0 such that for any n and for any $J \in \mathcal{P}_n$, we can find $J'_1, J'_2, \dots, J'_q \in \mathcal{P}_{n+n_o}$ satisfying $J'_i \subset J$ and $T_{J'_i}^{n+n_0} J'_i = A(i)$ for $i = 1, 2, \dots, q$. Therefore if $f(x_0) = 0$ could occur for some $x_0 \in A(i)$, then the equation

$$\begin{aligned} f(x_0) &= \lambda^{-(n+n_0)} L(\sigma)^{n+n_0} f(x_0) \\ &= \lambda^{-(n+n_0)} \sum_{J'\in\mathcal{P}_{n+n_0}:\tau(J')=i} G_{n+n_0,J'}(\sigma)(T_{J'}^{-(n+n_0)}x_0) f(T_{J'}^{-(n+n_0)}x_0) \end{aligned}$$

would imply that for each $J \in \mathcal{P}_n$, we could find a point $T_{J'_i}^{-(n+n_0)}x_0 \in J$. Thus f should vanish on a dense set in X, consequently f could be identically zero. This is a contradiction.

If f is an eigenfunction of $L(\sigma)$ on $\mathcal{A}_b(D(X))$, $f_X = f|_X$ is clearly an eigenfunction of $L_X(\sigma)$ on $C(X)$ belonging to the same eigenvalue. If we note that $L_X(\sigma)$ defined in §5 is a positive operator, it is not hard to see that $\lambda(\sigma)$ is maximal in the above sense. □

The following lemma gives information on the eigenvalues of $L(s)$ with maximal modulus.

LEMMA 6.2. *Assume that the Markov system $\mathcal{T}_\Gamma$ satisfies Condition 3.1. Let $L_X(s) : C(X) \to C(X)$ be as in §5 with $\Re s > 1/2$. Then:*

(1) *if s is real, the maximal eigenvalue $\lambda(s)$ is simple;*
(2) *$L_X(s)$ has an eigenvalue λ with $|\lambda| = \lambda(\Re s)$ if and only if s is real; and*
(3) *in particular, there is no eigenvalue of modulus $|\lambda(s)|$ except for $\lambda(s)$ for any real $s > 1/2$.*

Proof. Suppose $s \in \mathbb{R}$. As in the proof of the previous lemma, we note that if f is an eigenfunction of $L(s)$ on $\mathcal{A}_b(D(X))$, $f_X = f|_X$ is an eigenfunction of $L_X(s)$ on $C(X)$ belonging to the same eigenvalue. Consider the eigenfunction $f(s)$ of $L(s)$ belonging to the maximal eigenvalue $\lambda(s)$ which is found in Lemma 6.1. Since $f_X(s) = f(s)|_X$ is positive valued everywhere, we can define $\tilde{L}_X(s) : C(X) \to C(X)$ by

$$\tilde{L}_X(s)f(x) = \lambda(s)^{-1} f_X(s)^{-1}(x) L_X(s)(f_X(s)f)(x).$$

Then λ is an eigenvalue of $L_X(s)$ if and only if $\lambda(s)^{-1}\lambda$ is an eigenvalue of $\tilde{L}_X(s)$. Since $\tilde{L}_X(s)1 = 1$ from the definition, $\tilde{L}_X(s)$ is a positive operator of norm 1. Thus in order to show assertion (1), it suffices to show that if $\tilde{L}_X(s)g = g$ holds for a real valued function $g \in C(X)$, then g turns out to be a constant function.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Suppose that g takes its maximum at $x_0 \in A(i)$. Then we have

$$\begin{aligned} g(x_0) &= \tilde{L}_X(s)^n g(x_0) \\ &= \lambda(s)^{-n} f_X(s)(x_0)^{-1} \sum_{J\in\mathcal{P}_n:\tau(J)=i} G_{n,J}(s)(T_J^{-n}x_0) f_X(s)(T_J^{-n}x_0) g(T_J^{-n}x_0) \end{aligned}$$

for each positive integer n. Hence $g(T_J^{-n}x_0)$ must coincide with $g(x_0)$, otherwise the right-hand side becomes smaller than the left-hand side. Using Condition 3.1 we can see that g takes value $g(x_0)$ on the dense subset of X as in the proof of Lemma 6.1. Now the proof of (1) is complete.

Next we prove that if $e^{\sqrt{-1}\alpha}$, $\alpha \in [0, 2\pi)$, is an eigenvalue of $\tilde{L}_X(s)$, $s = \sigma + \sqrt{-1}\,t$ with $\sigma > 1/2$, then both α and t must be zero. Suppose $\tilde{L}_X(s)g = e^{\sqrt{-1}\alpha}g$ for some $g \in C(X)$. Then we have $|g| = |\tilde{L}_X(s)^n g| \le \tilde{L}_X(\sigma)^n |g|$ for any positive integer n. In the same way as in the proof of the simplicity of the maximal eigenvalue in (1), we can prove that $|g| \equiv |g|(x_0)$ for x_0 attaining the maximum of $|g|$. Thus we may assume that $|g| \equiv 1$. From the definition of $\tilde{L}_X(s)$, we again write

$$\begin{aligned} & e^{\sqrt{-1}n\alpha} g(x) \\ &= \tilde{L}_X(s)^n g(x) \\ &= \sum_{J\in\mathcal{P}_n:T_J^n J\ni x} f_X(\sigma)^{-1}(x)\lambda(\sigma)^{-n}|(T_J^n)'(T_J^{-n}x)|^{-(\sigma+\sqrt{-1}\,t)} f_X(\sigma)(T_J^{-n}x) g(T_J^{-n}x) \end{aligned}$$

for any $x \in X$ and for any $n \ge 1$. Since the weights $f_X(\sigma)^{-1}(x)\lambda(\sigma)^{-n}|(T_J^n)'(T_J^{-n}x)|^{-\sigma} f_X(\sigma)(T_J^{-n}x)$ of the summand are positive and their sum equals one, we can conclude that

$$g(T_J^{-n}x)|(T_J^n)'(T_J^{-n}x)|^{\sqrt{-1}\,t} = e^{\sqrt{-1}n\alpha} g(x)$$

holds for any x and for any $J \in \mathcal{P}_n$ with $x \in T_J^n J$. Therefore for any $\mathcal{T}_\Gamma$-periodic point x of period n, we obtain $(T_J^n)'(x)^{\sqrt{-1}\,t} = e^{\sqrt{-1}n\alpha}$ if $T_J^n x = x$. Now we have seen that if the next claim is verified, the proof of the other assertions is complete.

CLAIM 6.3. *Let $\mathcal{T}_\Gamma$ be as in Lemma 6.2. Let $t \in \mathbb{R}$ and $\alpha \in [0, 2\pi)$. Assume that for any $\mathcal{T}_\Gamma$-periodic point x, we have*

$$(T_J^n)'(x)^{\sqrt{-1}\,t} = e^{\sqrt{-1}n\alpha},$$

where n is the period of x and J is the element in $\mathcal{P}_n$ containing x. Then t and α turn out to be zero.

Proof. By virtue of Condition 3.1, we can find a positive integer k and $J_1, J_2 \in \mathcal{P}_k$ and $J \in \mathcal{P}$ such that $J_1, J_2 \subset J$, $T_{J_1}^k J_1 = T_{J_2}^k J_2 = A(\iota(J)) \supset J$. Consequently, we obtain a sequence $x, x_0, x_1, \ldots$ of distinct $\mathcal{T}_\Gamma$-periodic points with $T_{J_2}^k x = x$, $T_{J_1}^k T_{J_2}^{nk} x_n = x_n$, $n = 0, 1, \ldots$. Setting $\ell = \log(T_{J_2}^k)'(x)$ and $\ell_n = \log(T_{J_1}^k T_{J_2}^{nk})'(x_n)$ for $n = 0, 1, \ldots$, we have

$$e^{\sqrt{-1}\,(t\ell - k\alpha)} = 1 \quad \text{and} \quad e^{\sqrt{-1}\,(t\ell_n) - (n+1)k\alpha} = 1, \quad n = 0, 1, \ldots$$

by the assumption.

Suppose that $t \ne 0$. Then we have

$$\ell - k\frac{\alpha}{t} \in \frac{2\pi}{t}\mathbb{Z} \quad \text{and} \quad \ell_n - (n+1)k\frac{\alpha}{t} \in \frac{2\pi}{t}\mathbb{Z}, \quad n = 0, 1, \ldots. \tag{6.2}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

It will be more convenient to handle the elements in $PSL(2,\mathbb{R})$. By using an appropriate conformal transform, we can find non-commutative hyperbolic elements

$$G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$$

of $SL(2,\mathbb{R})$ with $\lambda > 1$ so that GH^n are all hyperbolic for $n = 0, 1, \ldots$ and

$$\begin{aligned} \ell &= 2\cosh^{-1}\frac{\operatorname{tr} H}{2} = 2\cosh^{-1}\frac{\lambda+\lambda^{-1}}{2} \\ \ell_n &= 2\cosh^{-1}\frac{\operatorname{tr} GH^n}{2} = 2\cosh^{-1}\frac{a\lambda^n + d\lambda^{-n}}{2} \end{aligned} \tag{6.3}$$

hold. Consider the sequence

$$a_n = \frac{\cosh(\ell_{n+1}/2)}{\cosh(\ell_n/2)} = \frac{a\lambda^{n+1} + d\lambda^{-(n+1)}}{a\lambda^n + d\lambda^{-n}}, \quad n = 0, 1, 2, \ldots.$$

Since GH^n are all hyperbolic, $|\operatorname{tr} GH^n| > 2$ for all $n \geq 0$. Therefore we have $a \neq 0$, consequently $a_n \to \lambda = e^{\ell/2}$ and $\ell_n \to +\infty$ as $n \to \infty$ by (6.3). This implies that $\ell_{n+1} - \ell_n \to \ell$ as $n \to \infty$. Thus $(\ell_{n+1} - \ell_n) - (\ell_n - \ell_{n-1}) \to 0$ as $n \to \infty$. On the other hand, from (6.2) we see that

$$(\ell_{n+1} - \ell_n) - (\ell_n - \ell_{n-1}) \in \beta\mathbb{Z}, \quad n = 1, 2, \ldots \tag{6.4}$$

for $\beta = 2\pi/t$. Hence we see that

$$\ell_{n+1} = \ell_n + \ell \tag{6.5}$$

holds for all n sufficiently large.

Next we look at the equations in (6.3). First we obtain the relation

$$2\cosh\frac{\ell}{2}\cosh\frac{\ell_n}{2} = \cosh\frac{\ell_{n+1}}{2} + \cosh\frac{\ell_{n-1}}{2}, \quad n = 1, 2, \ldots. \tag{6.6}$$

Secondly, solving (6.3) with respect to a and d, we obtain, for $n = 1, 2, \ldots$,

$$\begin{cases} a = 2(\lambda - \lambda^{-1})^{-1}\left(\lambda^{-(n-1)}\cosh\dfrac{\ell_n}{2} - \lambda^{-n}\cosh\dfrac{\ell_{n-1}}{2}\right) \\ b = 2(\lambda - \lambda^{-1})^{-1}\left(\lambda^{n}\cosh\dfrac{\ell_{n-1}}{2} - \lambda^{n-1}\cosh\dfrac{\ell_n}{2}\right). \end{cases} \tag{6.7}$$

Combining (6.6) and the validity of (6.5) for n large, we can see that (6.5) holds for any $n \geq 0$. Applying the relation $\ell_1 = \ell_0 + \ell$ to (6.7), we have $a = e^{\ell_0/2}$ and $d = e^{-\ell_0/2}$. Hence G has the form $\left(\begin{smallmatrix} a & b \\ 0 & a^{-1} \end{smallmatrix}\right)$ $\left(\text{respectively } \left(\begin{smallmatrix} a & 0 \\ c & a^{-1} \end{smallmatrix}\right)\right)$. Since $GH \neq HG$, we have $b \neq 0$ (respectively $c \neq 0$). Hence we have

$$GH^{-n}G^{-1}H^n = \begin{pmatrix} 1 & ab(1-\lambda^{-2n}) \\ 0 & 1 \end{pmatrix}$$

$$\left(\text{respectively } GH^nG^{-1}H^{-n} = \begin{pmatrix} 1 & 0 \\ a^{-1}c(1-\lambda^{-2n}) & 1 \end{pmatrix}\right).$$

This contradicts the discreteness of Γ. Hence we conclude that t must be zero.

Now we notice that the proof of the assertion that $\alpha = 0$ becomes very easy. In fact, if it were not zero, we would have (6.4) with $\beta = 1$. Thus we reach a contradiction in the same way as above. □

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Now we can state the main theorem of this section.

THEOREM 6.4. *Assume that the Markov system $\mathcal{T}_\Gamma$ satisfies Condition 3.1. Consider the $\mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$-valued analytic function $L(s)$ in $\{s \in \mathbb{C} : \Re s > 1/2\}$. Then the spectral radius of $L(s)$ with $s \notin \mathbb{R}$ is less than $\lambda(\Re s)$ and there is a neighborhood $\mathcal{U} \subset \mathbb{C}$ of the interval $(1/2, \infty) \subset \mathbb{R}$ where $L(s)$ has the spectral decomposition*

$$L(s)^n = \lambda(s)^n E(s) + R(s)^n, \quad n = 1, 2, \ldots \tag{6.8}$$

so that the following hold.

(a) *$s \mapsto E(s)$ and $s \mapsto R(s)$ are $\mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$-valued analytic functions with*

$$E(s)^2 = E(s), \quad E(s)R(s) = R(s)E(s) = 0.$$

In particular, the spectral radius of $R(s)$ is less than $|\lambda(s)|$. Moreover, $E(s)$ has the form

$$E(s)f = F(s)(f)h(s),$$

where $h(s) = E(s)1$ is an $\mathcal{A}_b(D(X))$-valued analytic function and $F(s)$ is an $\mathcal{A}_b(D(X))'$-valued analytic function.

(b) *$\lambda(s)$ is an analytic function in $\mathcal{U}$ which coincides with the eigenvalue of $L(s)$ of maximal modulus. In particular, $\lambda(s)$ is strictly decreasing on the interval $(1/2, \infty)$.*

(c) *For real s, $h(s)|_X$ is positive valued and $F(s) : \mathcal{A}_b(D(X)) \to \mathbb{C}$ is given by*

$$F(s)f = \int_X f|_X(x)\nu(s)(dx) \tag{6.9}$$

for a probability measure $\nu(s)$ on X. In particular, $\nu(1)$ is the normalized Lebesgue measure on X and consequently $\lambda(1) = 1$.

Proof. The first assertion is a direct consequence of assertion (2) in Lemma 6.2. For real $s > 1/2$, let $\lambda(s)$ be the maximal eigenvalue of $L(s)$. By virtue of assertion (3) of Lemma 6.2, the modulus of the eigenvalues other than $\lambda(s)$ is less than $\lambda(s)$. Combining this with assertion (1) of Lemma 6.2, we have the spectral decomposition (6.8) for each real $s > 1/2$. Since $s \mapsto L(s)$ is an $\mathcal{L}(\mathcal{A}_b(D(X)) \to \mathcal{A}_b(D(X)))$-valued analytic function in $\{s \in \mathbb{C} : \Re s > 1/2\}$, the general perturbation theory of an analytic family of linear operators in [**10**] implies that there is a neighborhood $\mathcal{U}$ of the interval $(1/2, \infty)$ in $\mathbb{C}$ such that (a) and the first assertion of (b) are satisfied by setting $h(s) = E(s)1$.

Next we fix a real s with $s > 1/2$. Now we write the decomposition (6.8) as

$$\lambda(s)^{-n}L(s)^n f = F(s)(f)h(s) + \lambda(s)^{-n}R(s)^n. \tag{6.10}$$

The positivity of $h(s)$ in (c) is shown as follows. Substituting 1 for f, we can see that $\lambda(s)^{-n}L(s)^n 1$ is positive valued on X and converges uniformly to $h(s)$ on $D(X)$ as $n \to \infty$. Hence $h(s)|_X$ is positive valued on X.

Next we note that for any $x \in X$ we have

$$\begin{aligned} \lambda(s)^{-n}L(s)^n f(x) &\le \lambda(s)^{-n}L(s)\left(\frac{h(s)}{\inf_{y\in X} h(s)(y)}\right)(x)\|f|_X\|_\infty \\ &\le \frac{\sup_{y\in X} h(s)(y)}{\inf_{y\in X} h(s)(y)}\|f|_X\|_\infty. \end{aligned} \tag{6.11}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Combining the explicit form (5.3) of $L(s)$ with the estimate (6.11), we have

$$\begin{aligned}\lambda(s)^{-n}\lambda(s+t)^n &\leq \lambda(s)^{-n}\frac{(L(s+t)^n h(s+t))(x)}{h(s+t)(x)}\\ &\leq \frac{\sup_{y\in X} h(s+t)(y)}{\inf_{y\in X} h(s+t)(y)} \sup_{J\in\mathcal{P}_n}\sup_{x\in J} |(T_J^n)'z|^{-t}\lambda(s)^{-n}L(s)^n 1(x)\\ &\leq \sup_{J\in\mathcal{P}_n}\sup_{x\in J} |(T_J^n)'z|^{-t}\frac{\sup_{y\in X} h(s+t)(y)}{\inf_{y\in X} h(s+t)(y)}\frac{\sup_{y\in X} h(s)(y)}{\inf_{y\in X} h(s)(y)}\end{aligned}$$

for any $t > 0$ and for any $x \in X$. By virtue of the Lasota–Yorke condition, $\sup_{J\in\mathcal{P}_n}\sup_{x\in J}|(T_J^n)'z|^{-t}$ goes to zero exponentially fast as n goes to infinity. Therefore, we see that $\lambda(s)$ is strictly decreasing in $(1/2,\infty)$. The proof of (b) is now complete.

We consider equation (6.10) at any $x \in X$: we have

$$|F(s)(f)| \leq \frac{\sup_{y\in X} h(s)(y)}{(\inf_{y\in X} h(s)(y))^2}\|f|_X\|_\infty + \frac{\lambda(s)^{-n}|R(s)^n f(x)|}{\inf_{y\in X} h(s)(y)}.$$

Letting n tend to infinity we obtain

$$|F(s)(f)| \leq \frac{\sup_{y\in X} h(s)(y)}{(\inf_{y\in X} h(s)(y))^2}\|f|_X\|_\infty$$

by virtue of (6.11). Thus we can extend $F(s)$ uniquely to a bounded linear functional on $C(X)$. From (6.10), the positivity and $F(1) = 1$ are easily proved. Hence we can see the representation (6.9).

In the case when $s = 1$, it is not hard to see from the change of variable formula that $\int_X L(1)f(x)m(dx) = \int_X f(x)m(dx)$ holds for any $f \in C(X)$, where m is the Lebesgue measure on X. Combining this with (6.10), it follows that $\lambda(1) = 1$ and $\nu(1)$ is the normalized Lebesgue measure on X. The proof of (c) is complete. □

7. *The Selberg zeta function*

Now we apply the ergodic theory of transfer operators developed in the preceding sections to the Selberg zeta function and give proofs of Theorems A–C given in the Introduction.

First we prepare the following.

LEMMA 7.1. (Cf. [**6**, **16**]) *Let D be a bounded domain in $\mathbb{C}$ and let $\varphi : D \to D$ be a univalent function with $\sup_{z\in D}|\varphi'(z)| < 1$. Consider an operator $L : \mathcal{A}_b(D) \to \mathcal{A}_b(D)$ defined by $Lf(z) = F(\varphi(z))f(\varphi(z))$ using an analytic function F on D with $F(a) \neq 0$, where a is a unique fixed point of φ. Then the point spectrum of L coincides with the set*

$$\Lambda(L) = \{F(a)\varphi'(a)^k : k = 0, 1, \ldots\}$$

and any element in $\Lambda(L)$ is simple.

Proof. Clearly L is a compact operator and its spectrum consists of eigenvalues only. Let λ be an eigenvalue of L and let f be a corresponding eigenfunction. Consider the smallest non-negative integer k satisfying $f^{(k)}(a) \neq 0$. Differentiating both sides of $Lf = \lambda f$ k times at $z = a$, we have $F(a)\varphi'(a)^k f^{(k)}(a) = \lambda f^{(k)}(a)$. Consequently, we have $\lambda = F(a)\varphi'(a)^k$. Hence any eigenvalue of L is an element of $\Lambda(L)$.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Next we show that any element in $\Lambda(L)$ must be an eigenvalue of L. Combining the Banach inverse mapping theorem and the fact that the spectrum of L consists of eigenvalues only, we can see that a number λ is an eigenvalue of L if and only if $L-\lambda$ is not surjective. Consider $L - F(a)\varphi'(a)^k$ for a non-negative integer k. Then we can see that there is no function $f \in \mathcal{A}_b(D)$ satisfying $(L - F(a)\varphi'(a)^k)f(z) = (z-a)^k$. Contrary to the assertion, suppose that we can find such a function f. Let n be the smallest non-negative integer with $f^{(n)}(a) \neq 0$. Then differentiating both sides of $(L - F(a)\varphi'(a)^k)f(z) = (z-a)^k$ at a, we have $F(a)(\varphi'(a)^n - \varphi'(a)^k) = k!\delta_{kn}$, where δ_{kn} is the Kronecker delta. Since $F(a) \neq 0$ and $0 < |\varphi'(a)| < 1$, this is impossible.

Finally, we show that $F(a)\varphi'(a)^k$ is simple. Suppose that g is a non-trivial function satisfying $(L - F(a)\varphi'(a)^k)^\ell g = 0$. That is,

$$\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j F(a)^{\ell-j} \varphi'(a)^{(\ell-j)k} F(\varphi^j(z)) F(\varphi^{j-1}(z)) \cdots F(\varphi(z)) g(\varphi^j(z)) \tag{7.1}$$

holds for any $z \in D$. Again, let n be the smallest non-negative integer with $g^{(n)}(a) \neq 0$. Differentiating both sides of (7.1) n times at $z = a$, we have

$$\sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j F(a)^\ell \varphi'(a)^{(\ell-j)k+nj} g^{(n)}(a) = 0.$$

Consequently we have $(\varphi'(a)^k - \varphi'(a)^n)^\ell = 0$. Therefore we must have $n = k$. Differentiating both sides (7.1) repeatedly at $z = a$, it is easy to see that there are constants $c_1, c_2, \ldots$ independent of g such that $g^{(k+j)}(a) = c_j g^{(k)}(a)$ holds for each $j = 1, 2, \ldots$. From this fact we can conclude that if $f \in \mathcal{A}_b(D)$ satisfies $(L - F(a)\varphi'(a)^k)^\ell f = 0$, we have $f = (f^{(k)}(a)/g^{(k)}(a))g$. This completes the proof of Lemma 7.1. □

Now we can prove the following theorem.

THEOREM 7.2. *Let $\mathcal{T}_\Gamma = (\mathcal{R}, \mathcal{P}, T)$ be the Markov system associated with the cofinite Fuchsian group Γ. For s with $\Re s > 1$, the Fredholm determinant* $\mathrm{Det}(I - L(s))$ *of the transfer operator with respect to $\mathcal{T}_\Gamma$ is represented by an absolutely convergent series as*

$$\mathrm{Det}(I - L(s)) = \exp\left(- \sum_{\tau \in PO(\mathcal{T}_\Gamma)} \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \frac{1}{n} \exp(-(s+k)n\ell(\Phi(\tau))) \right).$$

Proof. By Corollary 5.2 and Theorem 6.4, $L(s)$ is a nuclear operator of order zero with spectral radius less than one if $\Re s > 1$. Therefore the series $\sum_{n=1}^{\infty} (1/n) \operatorname{Tr} L(s)^n$ is absolutely convergent and is an analytic function in the domain $\{s \in \mathbb{C} : \Re s > 1\}$. By a formula in [**7**] and [**8**], we can write

$$\mathrm{Det}(I - L(s)) = \exp\left(- \sum_{n=1}^{\infty} \operatorname{Tr} L(s)^n \right). \tag{7.2}$$

One can evaluate $\operatorname{Tr} L(s)^n$ as follows. Recalling formulae (5.1)–(5.3), we write

$$\operatorname{Tr} L(s)^n = \sum_{J \in \mathcal{P}_n : T_J^n J \supset J} \operatorname{Tr} L(n, s, J), \tag{7.3}$$

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

where $L(n, s, J) : \mathcal{A}_b(D(\iota(J))) \to \mathcal{A}_b(D(\iota(J)))$ is the operator defined by

$$L(n, s, J)f(z) = G_{n,J}(s)(T_J^{-n}z)f(T_J^{-n}z).$$

We note that the absolute convergence of the series (7.3) for $s, \Re s > 1$, follows from (3) of Lemma 3.4 and the explicit form (5.3) of $L(s)$. By virtue of Lemma 7.1, the spectrum set of $L(n, s, J)$ consists of only simple eigenvalues of the form $(T_J^n)'(a_J)^{-(s+k)}, k = 0, 1, \ldots$, where a_J is a unique fixed point of T_J^n in J. Combining this and Lemma 4.2, the series in formula (7.2) is written as

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n}\operatorname{Tr}L(s)^n &= \sum_{n=1}^{\infty}\sum_{J\in\mathcal{P}_n:T_J^nJ\supset J}\sum_{k=0}^{\infty}((T_J^n)'(a_J))^{-(s+k)}\\ &= \sum_{\tau\in PO(\mathcal{T}_\Gamma)}\sum_{k=0}^{\infty}\sum_{n=1}^{\infty}\frac{1}{n}\exp(-(s+k)n\ell(\Phi(\tau))).\end{aligned}$$

Here we have used the absolute convergence of series in (7.2) and (7.3) to change the order of summations. Thus we obtain the theorem. □

The set of hyperbolic conjugacy classes is divided into the subsets $HC_0(\Gamma) = \Phi(PO(\mathcal{T}_\Gamma)_0)$ and $HC_1(\Gamma) = HC(\Gamma) \setminus HC_0(\Gamma)$, where $\Phi : PO(\mathcal{T}_\Gamma) \to HC(\Gamma)$ is the map defined in Lemma 4.2. Note that $HC_1(\Gamma)$ is a finite set. Put

$$\Xi(s) = \prod_{k=0}^{\infty}\prod_{c\in HC_1(\Gamma)}(1-\exp(-(s+k)\ell(c)))\,.$$

Then $\Xi(s)$ is meromorphic in $\mathbb{C}$ and, in particular, it is analytic in $\{s : \Re s > 0\}$ without zeros. By virtue of Lemma 4.2, Corollary 5.2, Theorem 6.4 and Theorem 7.2, we obtain the following theorem which implies Theorem A and Theorem B given in the Introduction.

THEOREM 7.3. *For s with $\Re s > 1$,*

$$Z(s)\Xi(s)^2 = \operatorname{Det}(I - L(s))$$

holds. $Z(s)$ is an analytic function in the domain $\{s \in \mathbb{C} : \Re s > 1\}$ without zeros having a meromorphic extension to the entire s-plane. In particular, there is a domain containing $\{s \in \mathbb{C} : \Re s \geq 1\}$ where $s = 1$ is a unique zero of the analytic function $Z(s)$ and it is simple. Moreover, the candidates of poles are located on the axes $\Re s = -k/2, k = -1, 0, 1, 2, \ldots$.

The following theorem and Ikehara's Tauberian theorem implies Theorem C (see [**14**, Chapter 6 and Appendix I]).

THEOREM 7.4. *Put*

$$\zeta(s) = \prod_{c\in HC(\Gamma)}(1-\exp(-s\ell(c)))^{-1}.$$

Then ζ is absolutely convergent in the domain $\{s \in \mathbb{C} : \Re s > 1\}$ and represents an analytic function without zeros. It has a meromorphic extension to the entire s-plane. In particular, on the axis $\Re s = 1$, $\zeta(s)$ has no zeros, $s = 1$ is the unique pole on the axis and it is simple.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press

Proof. For $s, \Re s > 1$, we have

$$\zeta(s) = \frac{Z(s+1)}{Z(s)} = \prod_{c \in HC_1(\Gamma)} (1 - e^{-s\ell(c)})^{-2} \frac{\mathrm{Det}(I - L(s+1))}{\mathrm{Det}(I - L(s))}$$

from Theorem 7.2. Therefore poles of ζ on the axis $\Re s = 1$ are zeros of $\mathrm{Det}(I - L(s))$ of the same order. By virtue of the Fredholm theory in [**8**] and Theorem 6.4, we see that $s = 1$ is the unique zero of $\mathrm{Det}(I - L(s))$ and it is simple. The other assertions are obvious from Theorem 7.2. □

Acknowledgement. The author would like to express his thanks to Professor Koyama for informing him of the existence of [**1**].

REFERENCES

[1] E. B. Bogomolney and M. Carioli. Quantum maps from transfer operators. *Physica* **D67** (1993), 88–112.
[2] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics 470).* Springer, Berlin, 1975.
[3] R. Bowen and C. Series. Markov maps associated with Fuchsian groups. *Publ. IHES* **50** (1979), 401–418.
[4] N. Dunford and J. T. Schwartz. *Linear Operators Part I*. Wiley, New York, 1988.
[5] I. Efrat. Dynamics of the continued fraction map and the spectral theory of $SL(2, \mathbb{Z})$. *Invent. Math.* **114** (1993), 207–218.
[6] C. Faivre. Perron–Frobenius operator for certain transformations of the interval. *Preprint*.
[7] A. Grothendieck. Produits tenssoriels topologiques et espaces necleaires. *Memoir. Amer. Math. Soc.* **16** (1955).
[8] A. Grothendieck. La Théorie de Fredholm. *Bull. Soc. Math. France* **84** (1956), 319–384.
[9] D. Hejhal. *The Selberg Trace formula for $PSL(2, \mathbb{R})$ I. (Lecture Notes in Mathematics 548).* Springer, Berlin, Heidelberg, New York, 1976. *The Selberg trace formula for $PSL(2, \mathbb{R})$ II. (Lecture Notes in Mathematics 1001).* Springer, Berlin, Heidelberg, New York, 1983.
[10] T. Kato. *Perturbation Theory for Linear Operators.* Springer, Berlin, Heidelberg, New York, 1976.
[11] D. Mayer. On the thermodynamic formalism for the Gauss map. *Commun. Math. Phys.* **130** (1990), 311–333.
[12] D. Mayer. The thermodynamic formalism approach to Selberg's zeta function for $PSL(2, \mathbb{Z})$. *Bull. Amer. Math. Soc.* **25** (1991), 55–60.
[13] T. Morita. One-dimensional Markov maps and Selberg zeta functions associated to cofinite Fuchsian groups. *Geometry and Analysis in Dynamical System*. Ed. H. Ito. World Scientific, Singapore, 1994, pp. 161–173.
[14] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187–188** (1990).
[15] A. Pietsch. *Eigenvalues and s-numbers*. Cambridge University Press, London, New York, 1987.
[16] M. Pollicott. Some applications of thermodynamic formalism to manifolds with constant negative curvature. *Adv. in Math.* **85** (1991), 161–192.
[17] D. Ruelle. Zeta functions for expanding maps and Anosov flows. *Invent. Math.* **34** (1976), 231–242.
[18] D. Ruelle. *Thermodynamic Formalism*. Addison-Wesley, Reading MA, 1978.
[19] A. Selberg. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian space with applications to Dirichlet series. *J. Indian Math. Soc.* **20** (1956), 47–87.
[20] C. Series. Symbolic dynamics for geodesic flows. *Acta Math.* **146** (1981), 103–128.

https://doi.org/10.1017/S014338579708632X Published online by Cambridge University Press