An Extension of the Thermodynamic Formalism Approach to Selberg's Zeta Function for General Modular Groups

Cheng-Hung Chang* and Dieter H. Mayer**

Institut für Theoretische Physik
TU Clausthal
Arnold-Sommerfeld-Straße 6
38678 Clausthal-Zellerfeld

Abstract. In the framework of the thermodynamic formalism for dynamical systems [Rue78] Selberg's zeta function [Sel56] for the modular group $PSL(2,\mathbb{Z})$ can be expressed through the Fredholm determinant of the generalized Ruelle transfer operator for the dynamical system defined by the geodesic flow on the modular surface [May91b]. In the present paper we generalize this result to modular subgroups Γ of $PSL(2,\mathbb{Z})$ with finite index. The corresponding surfaces of constant negative curvature with finite hyperbolic volume are in general ramified covering surfaces of the modular surface for $PSL(2,\mathbb{Z})$. Selberg's zeta function for these modular subgroups can be expressed through the generalized transfer operator for $PSL(2,\mathbb{Z})$ belonging to the representation of $PSL(2,\mathbb{Z})$ induced by the trivial representation of the subgroup Γ. The decomposition of this induced representation into its irreducible components leads to a decomposition of the transfer operator for these modular groups in analogy to a well known factorization formula of Venkov and Zograf for Selberg's zeta function for modular subgroups [VZ83]. The transfer operator can be continued into the entire complex β plane and its Fredholm determinant defines a meromorphic function in this plane.

1 Introduction

The thermodynamic formalism in ergodic theory is concerned with the investigation of dynamical systems and their properties by employing ingredients and techniques from statistical mechanics of lattice spin systems [Rue78]. Besides other quantities like partition functions, the free energy, different entropies, the transfer operator is playing a very special role in this approach. This operator was originally devised to determine the partition functions of periodic configurations of the 1-dimensional Ising lattice-spin model [Isi25] of statistical mechanics. Based on the analogy between the dynamics of the shift operator on the configuration space of such spin models and the symbolic dynamics of discrete time dynamical systems the thermodynamic formalism for hyperbolic dynamical systems was introduced by D. Ruelle, Y. Sinai and R.

* e-mail: ch.chang@tu-clausthal.de
** e-mail: dieter.mayer@tu-clausthal.de

Bowen to study ergodic properties of such discrete time and, through the Poincaré map, also continuous time dynamical systems.

For very special systems like the geodesic flow on the modular surface $PSL(2,\mathbb{Z})\backslash\mathbb{H}$, this approach can be made quite explicit and leads to an interesting new approach to Selberg's zeta function for this group. This function, which counts in a certain way the length spectrum of the geodesic flow on this surface, is a special case of a dynamical zeta function introduced in the sixties by S. Smale and D. Ruelle. Contrary to the general case, for the geodesic flow on the modular surface the symbolic dynamics and hence the transfer operator can be written down explicitly so that many questions about this operator and hence Selberg's zeta function can be answered in great detail [May91b].

Selberg's classical approach to his zeta function through the trace formula shows that this function encodes through its zeros and poles interesting spectral properties of the Laplace-Beltrami operator on the modular surface respectively topological properties of the surface. In this sense this zeta function relates properties of a chaotic classical Hamiltonian system, namely the geodesic flow on a surface of constant negative curvature to properties of the corresponding quantum system, namely the free motion of a quantized particle on this surface. This answers at least for a very special case Einstein's by now famous question in [Ein17] which is at the basis of all modern developments in the field of 'quantum chaos' [Gut90].

Up to now properties of the distribution of the eigenvalue spectrum of quantum systems with chaotic classical limit and their relation to random matrix models respectively the topography of their eigenfunctions have been studied mostly numerically. Rigorous results have been obtained only for very special systems with nice arithmetic properties like the geodesic flow on arithmetic surfaces of constant negative curvature [Sar95]. Results for the eigenfunctions are even harder to get since the trace formulas are not of great help in this case [see the contribution of Steiner et al. to this volume]

The transfer operator approach to Selberg's zeta function where this function gets expressed as the Fredholm determinant of this operator could open a new possibility to attack such questions. In this approach spectral properties of an operator which is completely determined by properties of the underlying classical system like its length spectrum describe the analytic properties of the Selberg function and hence also the quantum spectrum of the quantized dynamical system. This surprising relation is achieved simply by analytic continuation of the classical transfer operator into the complex β plane. For a special value of this parameter namely $\beta = 1$ the transfer operator is just the well known Perron-Frobenius operator from ergodic theory.

Rather surprising is furthermore a result known up to now only for the group $PSL(2,\mathbb{Z})$: the eigenfunctions of the transfer operator with eigenvalue 1 at the β values corresponding to eigenvalues of the Laplace Beltrami operator $-\triangle$ can be explicitly related to the corresponding eigenfunctions of $-\triangle$, the so called Maass wave forms for $PSL(2,\mathbb{Z})$ [CM99], [CM98]. The explicit form of this transformation was derived by J. Lewis in [LZ97]. This is the first time where such a direct relation between quantum states and classical objects has been established for a chaotic system where the usual construction of quasimodes [Laz93] based on invariant Lagrangian manifolds cannot be applied. In the present case the quantum mechanical eigenfunctions are just related to eigenfunctions of the classical transfer operator. An example is the ground state of our quantum system on the modular surface which is related through the Lewis transformation to the invariant density of the Gauss map which on the other hand is closely related to the Poincare map of the classical flow. One should therefore expect that also the other eigenfunctions of the transfer operator related to the Maass wave forms should have an interesting interpretation for the classical system.

In this paper however we address another question namely how far is it possible to extend the transfer operator approach to Selberg's zeta function to more general Fuchsian groups than $PSL(2,\mathbb{Z})$. Hence we study subgroups Γ of $PSL(2,\mathbb{Z})$ with finite index. Since the geodesic flow on the corresponding surface $\Gamma\backslash\mathbb{H}$ is just a lift of the one for $PSL(2,\mathbb{Z})$ to a finite, in general ramified, covering surface of $PSL(2,\mathbb{Z})\backslash\mathbb{H}$, its symbolic dynamics and hence also its generalized transfer operator are closely related to the ones for $PSL(2,\mathbb{Z})$. Indeed, its transfer operator coincides with the generalized transfer operator for $PSL(2,\mathbb{Z})$ when one considers the geodesic flow with some representation of this group , corresponding in fact to a lift of the geodesic flow to some vector bundle [Fri86], [BO95] over the modular surface.

In detail the paper is organized as follows: in chapter 2 we recall the congruence subgroups of $PSL(2,\mathbb{Z})$ and introduce the special groups Γ_2, $\Gamma_0(2)$, $\Gamma^0(2)$, Γ_ϑ and $\Gamma(2)$ whose transfer operators we will discuss later in great detail. In chapter 3 we deduce the symbolic dynamics of the geodesic flow on the covering surfaces $\Gamma\backslash\mathbb{H}$ and give an explicit expression for the Poincaré map for these flows. In chapter 4 we discuss the transfer operators $\tilde{\mathcal{L}}_\beta$ for subgroups Γ, their analyticity properties in the 'temperature' β and the relation to transfer operators for $PSL(2,\mathbb{Z})$ with a representation, especially the one induced from a representation of Γ. We show that $\tilde{\mathcal{L}}_\beta$ is nuclear and that Selberg's zeta function for Γ can be simply expressed as the Fredholm determinant of $\tilde{\mathcal{L}}_\beta$. We derive a decomposition of $\tilde{\mathcal{L}}_\beta$ into a direct sum corresponding to the decomposition of the induced representation into its irreducible components and the resulting product formula of Venkov and Zograf for zeta functions for subgroups of $PSL(2,\mathbb{Z})$.

In a forthcoming second part to this paper we will discuss further spectral properties of the transfer operators for the special groups Γ_2, $\Gamma_0(2)$, $\Gamma^0(2)$, Γ_ϑ and $\Gamma(2)$ and how the theory of Lewis and Zagier [LZ97] of period functions for $PSL(2,\mathbb{Z})$ can be extended to these groups. We will also show how the transfer operator approach for these subgroups reflects the well known theory of Atkin-Lehner of old and new forms for cofinite Fuchsian groups [AL70].

2 The hyperbolic surfaces M_Γ

The Poincaré upper half-plane $\mathbb{H}$ is the half-plane

$$\mathbb{H} = \{x + iy \,|\, x,\, y \in \mathbb{R}, y > 0\} \tag{1}$$

equipped with the Poincaré metric (arc length)

$$d\,s^2 = y^{-2}\,(\,d\,x^2 + d\,y^2\,). \tag{2}$$

The surface $\mathbb{H}$ is a hyperbolic surface with constant negative Gaussian curvature -1 [Ter85].

The geodesics γ on $\mathbb{H}$ are the shortest paths through two points with respect to the Poincaré metric (2). They are either perpendicular half lines based on the real axis $\mathbb{R}$ or semi-circles based on two basepoints $\gamma_{-\infty}$ and $\gamma_{+\infty}$ in $\mathbb{R}$. A free particle on $\mathbb{H}$ moves along the geodesics with constant velocity. This motion is known to be hyperbolic; the geodesic flow on $\mathbb{H}$ is an Anosov flow but completely integrable. When introducing however the action of Fuchsian groups and their fundamental domains the induced geodesic flow on the corresponding quotient surface becomes highly chaotic and not predictable.

2.1 The modular groups

Consider first the group

$$PSL(2,\mathbb{R}) = SL(2,\mathbb{R})/\{\pm\, id\}\,,$$

where

$$SL(2,\mathbb{R}) := \{g = \begin{pmatrix} a\ b \\ c\ d \end{pmatrix} \mid a,b,c,d \in \mathbb{R},\ ad - bc = 1\},$$

and its action on $\mathbb{H}$ by Möbius-transformations

$$g\,z = \frac{az+b}{cz+d}, \tag{3}$$

for $z \in \mathbb{H}$. The group $PSL(2,\mathbb{R})$ is the isometry group of analytic actions on the surface $\mathbb{H}$. The discrete subgroups of $SL(2,\mathbb{R})$ are called Fuchsian groups. Examples are the full modular group

$$\Gamma(1) := PSL(2,\mathbb{Z}) = \{\begin{pmatrix} a\ b \\ c\ d \end{pmatrix} \mid a,\, b,\, c,\, d, \in \mathbb{Z},\, ad - bc = 1\}/\{\pm\, id\} \tag{4}$$

and its subgroups. The subgroups $\Gamma \subseteq \Gamma(1)$ with finite index $[\Gamma(1) : \Gamma]$ are called the modular groups.

The group $\Gamma(1)$ has two generators, e.g.

$$Q := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad T := \begin{pmatrix} 1\,1 \\ 0\,1 \end{pmatrix} \tag{5}$$

acting as $Qz = -\frac{1}{z}$ and $Tz = z + 1$ for $z \in \mathbb{H}$, which fulfill the following two relations

$$Q^2 = (QT)^3 = id. \tag{6}$$

The elements $\sigma \neq \pm id$ of $\Gamma(1)$ fall into three classes [Ter85]:

(i) σ is parabolic $\Leftrightarrow$ $|\mathrm{trace}\,\sigma| = 2$ and σ has Jordan normal form $\pm \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ with $a \neq 0$.

(ii) σ is elliptic $\Leftrightarrow$ $|\mathrm{trace}\,\sigma| < 2$ and σ has Jordan normal form $\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix}$ with $a \notin \mathbb{R}$ and $|a| = 1$.

(iii) σ is hyperbolic $\Leftrightarrow$ $|\mathrm{trace}\,\sigma| > 2$ and σ has Jordan normal form $\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix}$ with $a \in \mathbb{R}$ and $|a| > 1$.

An element $\sigma \in \Gamma$ is primitive if σ is not a power of another element in Γ.

The principal congruence subgroups $\Gamma(N)$ of level N, defined as

$$\Gamma(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) \,|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \; mod\, N, \; N \in \mathbb{N} \}, \tag{7}$$

are subgroups of the modular group $\Gamma(1)$. They are normal subgroups with index [Sch74]

$$\mu = [\Gamma(1) : \Gamma(N)] = \begin{cases} 6 & N = 2\,, \\ \frac{1}{2} N^3 \prod_{p\,|\,N} (1 - \frac{1}{p^2}) & N > 2\,, \end{cases}$$

where the devisor p of N is a prime number. The quotient groups $\Gamma(N)\backslash\Gamma(1)$ with $N = 2, 3, 4, 5$ are isomorphic to symmetric groups S_ν respectively alternating groups A_ν [KF66], [Sch74]:

$$\Gamma(2)\backslash\Gamma(1) \cong S_3, \; \Gamma(3)\backslash\Gamma(1) \cong A_4, \; \Gamma(4)\backslash\Gamma(1) \cong S_4 \text{ and } \Gamma(5)\backslash\Gamma(1) \cong A_5 \tag{8}$$

A subgroup Γ of $\Gamma(1)$ with the property $\Gamma(N) \subseteq \Gamma \subseteq \Gamma(1)$ for some $N \in \mathbb{N}$ is called a congruence subgroup. The following are examples of congruence subgroups:

$$\Gamma_0(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) |\; c = 0 \; mod\, N, \; N \in \mathbb{N} \}, \tag{9}$$

$$\Gamma^0(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) |\; b = 0 \; mod\, N, \; N \in \mathbb{N} \}. \tag{10}$$

These groups are not normal subgroups of $\Gamma(1)$ and have index μ in the full modular group $\Gamma(1)$ with [Sch74]

$$\mu = [\Gamma(1) : \Gamma_0(N)] = [\Gamma(1) : \Gamma^0(N)] = N \prod_{p|N} (1 + \frac{1}{p}). \tag{11}$$

Hence the groups $\Gamma_0(2)$ and $\Gamma^0(2)$ both have index three. Another congruence subgroup of index three is the theta group Γ_ϑ with

$$\Gamma_\vartheta := \Gamma(2) \cup \Gamma(2)Q \,. \tag{12}$$

The three groups $\Gamma_0(2)$, $\Gamma^0(2)$ and Γ_ϑ are conjugate to each other [Ran77]:

$$\Gamma^0(2) = G_1\,\Gamma_0(2)\,G_1^{-1} \;\text{ respectively }\; \Gamma_\vartheta = G_2\,\Gamma_0(2)\,G_2^{-1}\,,$$

with

$$G_1 = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \quad \text{and} \quad G_1^{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{respectively}$$
$$G_2 = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad G_2^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} .$$

Besides the principal congruence subgroups $\Gamma(N)$ the group $\Gamma(1)$ has other normal subgroups, e.g. [Sch74]

$$\Gamma_2 := \Gamma(2) \cup \Gamma(2)QT \cup \Gamma(2)(QT)^2 \tag{13}$$

with index two. Indeed all the subgroups of $\Gamma(1)$ with index $\mu \le 6$ are congruence subgroups. Non-congruence subgroups appear only for $\mu \ge 7$ [Ran77].

The following table collects some of the subgroups Γ of $\Gamma(1)$, their generators and representatives of the quotient sets $\Gamma\backslash\Gamma(1)$:

Group	Generators	Representatives of $\Gamma\backslash\Gamma(1)$
$\Gamma(1)$	Q, T	id
Γ_2	QT, QT^{-1}	id, Q
$\Gamma_0(2)$	T, $QT^{-2}Q$, $T^{-1}QT^{-2}Q$	id, Q, QT
$\Gamma^0(2)$	T^2, $QT^{-1}Q$, TQT	id, Q, T
Γ_ϑ	T^2, Q, $TQT^{-1}QT^{-1}$	id, T, TQ
$\Gamma(2)$	T^2, $QT^{-2}Q$	id, T, Q, TQ, QT^{-1}, TQT
$\Gamma^0(3)$	T^3, QTQ, TQT	id, T, T^{-1}, Q

Table 1. *Generators and representatives of different subgroups of* $\Gamma(1)$

2.2 The modular surfaces M_Γ

Let Γ be a subgroup of $\Gamma(1)$ with finite index $\mu = [\Gamma(1) : \Gamma] < \infty$. Two points z_1, $z_2 \in \mathbb{H}$ are Γ-equivalent if there exists an element $\sigma \in \Gamma$ with $z_1 = \sigma\, z_2$ as given in (3). The modular surface $M_\Gamma := \Gamma\backslash\mathbb{H}$ is the quotient surface defined by the Γ-equivalent points on $\mathbb{H}$.

A fundamental domain for Γ is a connected region $\mathcal{F}_\Gamma$ in $\mathbb{H}$ such that $\cup_{g\in\Gamma}\, g(\mathcal{F}_\Gamma) = \mathbb{H}$ with the property

$$g(\mathcal{F}_\Gamma^0) \cap g'(\mathcal{F}_\Gamma^0) = \emptyset\,, \;\; \forall\, g,\, g' \in \Gamma \;\text{ and }\; g \ne g'\,,$$

where $\mathcal{F}_\Gamma^0$ denotes the interior of $\mathcal{F}_\Gamma$. Conventionally one selects connected regions bounded by geodesics as fundamental domains. An example for such a domain for $\Gamma = \Gamma(1)$ is

$$\mathcal{F}_{\Gamma(1)} = \left\{ z \,\middle|\, |\Re z| \leq \frac{1}{2},\, |z| \geq 1,\, z \in \mathbb{H} \right\}. \tag{14}$$

For $\Gamma \subseteq \Gamma(1)$ one can then choose as a fundamental domain

$$\mathcal{F}_\Gamma = \cup_{\{g\} \in \Gamma \backslash \Gamma(1)} \, g(\mathcal{F}_{\Gamma(1)}) , \tag{15}$$

where the union runs over representatives of the μ different equivalence classes of the quotient $\Gamma \backslash \Gamma(1)$.

The modular surface M_Γ can be constructed by identifying boundary points of the fundamental domain $\mathcal{F}_\Gamma$ related by the generators of the group Γ [Ran77]. Topologically, the surface M_Γ is a sphere with finitely many handles and finitely many cusps located at ∞ and at the inequivalent rational points in $\mathbb{R}$. The group Γ is not cocompact but cofinite, i.e., its fundamental region is not compact but has a finite hyperbolic area.

The Poincaré upper half-plane $\mathbb{H}$ is the universal covering space of the quotient space M_Γ. The quotient space M_Γ is an in general ramified μ-fold covering space of the modular surface $M_{\Gamma(1)} = \Gamma(1) \backslash \mathbb{H}$.

3 The geodesic flow on M_Γ

3.1 Coding of the geodesics on M_Γ

We recall briefly the symbolic description of geodesics on the modular surfaces [Ser85]. For this, let us consider a special class of geodesics in $\mathbb{H}$, namely the set $\mathcal{A}$ of all oriented semi-circle running from $\gamma_{-\infty}$ to $\gamma_{+\infty}$ with $\gamma_{-\infty},\ \gamma_{+\infty} \in \mathbb{R}$ such that

$$\mathcal{A} := \{\, \gamma \mid 0 < |\gamma_{-\infty}| \leq 1 \leq |\gamma_{+\infty}|,\, \gamma_{-\infty}\gamma_{+\infty} < 0 \}. \tag{16}$$

The last condition $\gamma_{+\infty}\gamma_{-\infty} < 0$ makes sure that the geodesics in $\mathcal{A}$ cut the imaginary axis I. Every geodesic γ in $\mathbb{H}$ can be brought to $\mathcal{A}$ by means of an element in $g \in \Gamma(1)$ such that $g\gamma \in \mathcal{A}$. To describe the geodesics in $\mathcal{A}$ we consider the following two representations:

The first representation makes use of the continued fraction expansion. Since E. Artin [Art24] it is known that the geodesics on the modular surface $M_{\Gamma(1)}$ are closely related to the continued fraction transformation (Gauß transformation). Suppose $\gamma \in \mathcal{A}$ is a geodesic in $\mathbb{H}$ with the two basepoints $\gamma_{\pm\infty}$. Since $\gamma_{\pm\infty} \in \mathbb{R}$ these two points can be expressed by continued fractions:

$$\gamma_{-\infty} = -\varepsilon \left(\cfrac{1}{n_{-1} + \cfrac{1}{n_{-2} + \cdots}} \right) := -\varepsilon \, [n_{-1}, n_{-2}, n_{-3}, \cdots] \,, \tag{17}$$

$$\gamma_{+\infty} = \varepsilon \left(n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cdots}} \right) := \varepsilon \, [n_0, n_1, n_2, \cdots]^{-1} \,, \tag{18}$$

with $n_i \in \mathbb{N}$, where $\varepsilon = 1$ respectively $\varepsilon = -1$ denote the orientation $\gamma_{+\infty} > \gamma_{-\infty}$ respectively $\gamma_{+\infty} < \gamma_{-\infty}$. Hence every geodesic γ in $\mathcal{A}$ is determined by

$$\gamma := ([n_0, n_1, n_2, \cdots], [n_{-1}, n_{-2}, n_{-3}, \cdots], \varepsilon) \tag{19}$$
$$= (\varepsilon \, \gamma_{+\infty}^{-1}, -\varepsilon \, \gamma_{-\infty}, \varepsilon). \tag{20}$$

The second representation for the geodesics in $\mathcal{A}$ uses the so-called cutting sequences [Ser85]. To define this representation one needs the Farey tessellation $\mathbb{F}$ of the Poincaré upper half-plane $\mathbb{H}$. This tessellation $\mathbb{F}$ is defined by the hyperbolic triangle

$$\triangle_{\mathbb{F}} := \{ z \in \mathbb{H} \mid 0 \le \Re z \le 1, \ \frac{1}{2} \le |z - \frac{1}{2}| \} \tag{21}$$

and its translations under the group $\Gamma(1)$, hence $\mathbb{H} = \cup_{g \in \Gamma(1)} \, g \triangle_{\mathbb{F}}$. The boundary of the tessellation $\mathbb{F}$ consists of all boundaries of all the hyperbolic triangles $g \triangle_{\mathbb{F}}$ with $g \in \Gamma(1)$. They are either the half-lines $\Re z \in \mathbb{Z}$ with $z \in \mathbb{H}$ or semi-circles with rational basepoints p/q and p'/q' with p, q, p', $q' \in \mathbb{Z}$, q, $q' \neq 0$ and $pq' - qp' = \pm 1$ [Ser85].

Almost every oriented geodesic $\gamma \in \mathcal{A}$ intersects infinitely many hyperbolic triangles in $\mathbb{F}$ (up to those that run along the boundaries or run into the cusps of the hyperbolic triangles) and hence are cut by the boundaries of the triangles into infinitely many pieces of arcs. The three vertices of the triangle that the geodesic γ crosses are separated into the two sides of γ. Depending on which side (left respectively right) the single vertex of the triangle is located with respect to the oriented geodesic, the piece of the arc of γ in this triangle will be denoted by L respectively R. Therefore the geodesic γ intersecting the imaginary axis I in y, i.e., $y = \gamma \cap I$, can be described by the symbol sequence:

$$\gamma := \begin{cases} \cdots R^{n_{-3}} L^{n_{-2}} R^{n_{-1}} y L^{n_0} R^{n_1} L^{n_2} \cdots & \text{for } -1 \le \gamma_{-\infty} < 0, \ \ \varepsilon = 1 \,, \\ \cdots L^{n_{-3}} R^{n_{-2}} L^{n_{-1}} y R^{n_0} L^{n_1} R^{n_2} \cdots & \text{for } \ 0 < \gamma_{-\infty} \le 1, \ \ \varepsilon = -1 \,, \end{cases} \tag{22}$$

where $n_i \in \mathbb{N}$ denotes the number of consecutive arcs L respectively R.

Comparing the representations (20) and (22) for γ shows that the numbers n_i in the continued fraction expansions (17) respectively (18) coincide exactly with the numbers n_i of the consecutive arcs L respectively R in (22) [Ser85]. The sequence (22) stops in case the geodesic runs into a cusp of $\mathbb{H}$, namely the rational numbers $\mathbb{Q}$, or starts out in such a cusp. Its continued fraction expansion in (18) respectively (17) then is finite. In this case $\gamma_{-\infty}$ respectively $\gamma_{+\infty}$ is rational. The representation (22) of γ is called a cutting sequence [Ser85]. Later we will use it to determine the Poincaré return map for the geodesic flow in $M_{\Gamma(1)}$.

3.2 The geodesics on M_Γ and the geodesic flow

As mentioned earlier the Poincaré upper half-plane $\mathbb{H}$ is an infinite covering of the quotient space M_Γ for $\Gamma \subseteq \Gamma(1)$ and $[\Gamma(1) : \Gamma] < \infty$. Two different geodesics $\gamma^{(1)}$ and $\gamma^{(2)}$ in $\mathbb{H}$ are identified on the surface M_Γ if $\gamma^{(1)}$ and $\gamma^{(2)}$ are related by $\sigma\gamma^{(1)} = \gamma^{(2)}$ with some element $\sigma \in \Gamma$. Denote by

$$\pi : \mathbb{H} \to M_\Gamma \tag{23}$$

the projection of $\mathbb{H}$ onto M_Γ. Every geodesic γ in $\mathbb{H}$ is projected by π to a geodesic $\hat{\gamma}$ in M_Γ. Conversely a geodesic $\hat{\gamma}$ in M_Γ can be lifted to infinitely many γ in $\mathbb{H}$. Almost all geodesics in M_Γ cut the imaginary axis I, more precisely the projection of I in M_Γ, infinitely often.

A free particle on the surface $\mathbb{H}$ moves along a geodesic with constant velocity v, say $|v| = 1$. The physical phase space for this free particle is hence the unit tangent bundle $T_1\mathbb{H}$ of $\mathbb{H}$ of dimension three. The geodesic flow on $T_1\mathbb{H}$ is

$$\begin{aligned} \phi_t &: T_1\mathbb{H} \to T_1\mathbb{H} \\ &(\gamma_{\mathbb{H}}(0), \dot{\gamma_{\mathbb{H}}}(0)) \mapsto (\gamma_{\mathbb{H}}(t), \dot{\gamma_{\mathbb{H}}}(t)), \quad t \in \mathbb{R}, \end{aligned} \tag{24}$$

where $\gamma_{\mathbb{H}} : \mathbb{R} \to \mathbb{H}$ denotes the geodesic in $\mathbb{H}$ parameterized by the arc length respectively the time t through the initial position $\gamma_{\mathbb{H}}(0)$ with the initial tangent vector $\dot{\gamma_{\mathbb{H}}}(0)$. A (geodesic) orbit of the geodesic flow ϕ_t is a path in $T_1\mathbb{H}$ through the initial point $(\gamma_{\mathbb{H}}(0), \dot{\gamma_{\mathbb{H}}}(0)) \in T_1\mathbb{H}$.

Let $T_1 M_\Gamma$ be the unit tangent bundle of the configuration space of a free particle on M_Γ . Analogous to (24), the geodesic flow reads

$$\begin{aligned} \hat{\phi}_t &: T_1 M_\Gamma \to T_1 M_\Gamma \\ &(\hat{\gamma}_{M_\Gamma}(0), \dot{\hat{\gamma}}_{M_\Gamma}(0)) \mapsto (\hat{\gamma}_{M_\Gamma}(t), \dot{\hat{\gamma}}_{M_\Gamma}(t)) \end{aligned} \tag{25}$$

where $\hat{\gamma}_{M_\Gamma} : \mathbb{R} \to M_\Gamma$ denotes the geodesic in M_Γ through the point $\hat{\gamma}_{M_\Gamma}(0)$ with tangent vector $\dot{\hat{\gamma}}_{M_\Gamma}(0)$. The projection π in (23) relates the orbits of the geodesic flows on $\mathbb{H}$ and M_Γ in the obvious way:

$$\begin{aligned} \pi_* &: T_1\mathbb{H} \to T_1 M_\Gamma \\ &(\gamma_{\mathbb{H}}(t), \dot{\gamma_{\mathbb{H}}}(t)) \mapsto (\pi\gamma_{\mathbb{H}}(t), \pi\dot{\gamma_{\mathbb{H}}}(t)). \end{aligned} \tag{26}$$

Every periodic orbit of ϕ_t of period l in $T_1 M_\Gamma$ can be identified with a closed geodesic in M_Γ of length l, and vice versa [Pol91].

3.3 The Poincaré map for the geodesic flow on $M_{\Gamma(1)}$

To understand the construction of a Poincaré map for the geodesic flow on an arbitrary modular surface M_Γ we recall briefly the case of the full modular group $\Gamma(1)$. We use the approach by Series [Ser85].

To construct a Poincaré section for the geodesic flow $\hat{\phi}_t : T_1 M_{\Gamma(1)} \to T_1 M_{\Gamma(1)}$ for the modular surface for $\Gamma(1)$, consider first the half line $S = [i, i\infty) \in \mathbb{H}$ on the imaginary axis I. Suppose a point $y \in S$ is given. Obviously not all the geodesics through y change their type L or R in the cutting representation at the point y. Define $C(y)$ as the set of all unit vectors $v(y)$ at the point y such that the corresponding geodesic belongs to $\mathcal{A}$ and hence changes its type at y. One can then choose the Poincaré section X for the geodesic flow on $T_1 M_{\Gamma(1)}$ as [Ser85]

$$X := \bigcup_{y \in S,\, v(y) \in C(y)} \pi_*(y, v(y))\,. \tag{27}$$

To find the Poincaré return map for the Poincaré section X in (27) we consider a geodesic $\hat{\gamma}$ in $M_{\Gamma(1)}$ and lift $\hat{\gamma}$ to a geodesic γ_0 in $\mathcal{A}$. Suppose γ_0 has the cutting sequence representation

$$\gamma_0 = \cdots R^{n_{-3}} L^{n_{-2}} R^{n_{-1}} y_0 L^{n_0} R^{n_1} L^{n_2} \cdots\,,\ \ \varepsilon = +1 \tag{28}$$

with $y_0 = \gamma_0 \cap S$ and has basepoints

$$\gamma_{-\infty} = -[n_{-1}, n_{-2}, n_{-3}, \cdots] \quad \text{and} \quad \gamma_{+\infty} = [n_0, n_1, n_2, \cdots]^{-1}\,.$$

The corresponding orbit in $T_1 M_{\Gamma(1)}$ of the geodesic $\hat{\gamma}$ in $M_{\Gamma(1)}$ has then the crossing point

$$\pi_*(y_0, v(y_0)) \in X$$

with the Poincaré section X where $v(y_0)$ denotes the unit tangent vector of γ_0 at y_0.

To find the next crossing point of this orbit with X, we follow the oriented geodesic γ_0. The next change of the type along γ_0 takes place on the axis $\Re z = n_0 := n$ which can be identified with the imaginary axis by means of the elements $QT^{-n} \in \Gamma(1)$ or $T^{-n} \in \Gamma(1)$, where Q and T are defined in (5). The geodesic $T^{-n}\gamma_0$ however doesn't belong to $\mathcal{A}$, because its basepoints $|\gamma_{+\infty}| \leq 1$ and $|\gamma_{-\infty}| \geq 1$ violate the conditions imposed on $\mathcal{A}$. Therefore, the point $\pi_*(y_1', v(y_1'))$ with $y_1' = T^{-n}\gamma_0 \cap S$ cannot be the next crossing point in X. In contrast, the element QT^{-n} brings the geodesic $\gamma_0 \in \mathcal{A}$ in (28) into the set $\mathcal{A}$ again. The action of Q however leads to a change of the orientation of the geodesic γ_0, namely the basepoints $\gamma'_{\pm\infty}$ of $QT^{-n}\gamma_0$ have $\gamma'_{-\infty} > \gamma'_{+\infty}$

instead of $\gamma_{-\infty} < \gamma_{+\infty}$. The geodesics $\gamma_1 = QT^{-n}\gamma_0$ and γ_0 are equivalent under $\Gamma(1)$, i.e., γ_1 and γ_0 define the same geodesic in $M_{\Gamma(1)}$. The basepoints of the equivalent geodesic $\gamma = \gamma_1 = QT^{-n}\gamma_0$ are

$$\gamma'_{-\infty} = [n_0, n_{-1}, n_{-2}, n_{-3}, \cdots] \quad \text{and} \quad \gamma'_{+\infty} = -[n_1, n_2, n_3 \cdots]^{-1} .$$

The corresponding cutting sequence representation on the other hand is

$$\gamma_1 = QT^{-n_0}\gamma_0 = \cdots R^{n_{-3}} L^{n_{-2}} R^{n_{-1}} L^{n_0} y_1 R^{n_1} L^{n_2} \cdots , \quad \varepsilon = -1 ,$$

where $y_1 = \gamma_1 \cap S$ denotes the point of intersection of γ_1 with S. Let $v(y_1)$ be the tangent vector of γ_1 at y_1. Then the first return point of γ_0 to X is $\pi_*(y_1, v(y_1))$. The Poincaré return map $P : X \to X$ transforms the initial crossing point $\pi_*(y_0, v(y_0)) \in X$ into the next crossing point $\pi_*(y_1, v(y_1)) \in X$. Obviously the above arguments are valid also for the geodesics in $\mathcal{A}$ with $\varepsilon = -1$.

Combining the cases $\varepsilon = 1$ and $\varepsilon = -1$, suppose $\gamma \in \mathcal{A}$ is described as

$$\gamma = (\varepsilon\gamma_{+\infty}^{-1}, -\varepsilon\gamma_{-\infty}, \varepsilon) = ([n_0, n_1, n_2, \cdots], [n_{-1}, n_{-2}, n_{-3}, \cdots], \varepsilon), \quad \varepsilon = \pm 1 . \tag{29}$$

Then the Poincaré return map $P : X \to X$ transforms $\pi_*(y, v(y)) \in X$ into $\pi_*(y', v(y')) \in X$ with $y = \gamma \cap S$ and $y' = QT^{-n\varepsilon}\gamma \cap S$ where $n = n_0$ is the integer part of $\varepsilon\gamma_{+\infty}$. However it is known [Ser85] that the map

$$\rho : X \to \mathcal{A}$$

is bijective up to the two geodesics with basepoints $\gamma_{\pm\infty} = \pm 1$ respectively $\gamma_{\pm\infty} = \mp 1$ which are mapped into each other under Q and hence correspond to a single point in X [Ser85]. The map ρ allows to describe the points of X through the geodesics in $\mathcal{A}$ and hence by their basepoints $\gamma_{-\infty}$ and $\gamma_{+\infty}$. As in (29) the geodesics γ in $\mathcal{A}$ can be described as

$$\gamma = (\varepsilon\gamma_{+\infty}^{-1}, -\varepsilon\gamma_{-\infty}, \varepsilon)$$

with $\varepsilon = \pm 1$ and $\varepsilon\gamma_{+\infty}^{-1}$, $-\varepsilon\gamma_{-\infty}$ two real numbers in $[0, 1]$. In this coordinate system the Poincaré return map P has the following explicit form:

$$\begin{aligned} P : [0,1] \times [0,1] \times \mathbb{Z}_2 &\to [0,1] \times [0,1] \times \mathbb{Z}_2 \\ (x_1, x_2, \varepsilon) &\mapsto (-T^n Q x_1, -QT^n x_2, -\varepsilon) \end{aligned} \tag{30}$$

where for $x := (x_1, x_2, \varepsilon)$ the number $n = n(x) = [\frac{1}{x_1}]$ denotes the integer part of $\frac{1}{x_1}$ and the change of the orientation $\varepsilon \to -\varepsilon$ reflects the action of Q.

Since the first entry in the map (30) is nothing but the Gauß transformation

$$T_G : [0,1] \to [0,1] \quad \text{with} \quad T_G : z \mapsto \frac{1}{z} - [\frac{1}{z}], \tag{31}$$

we can rewrite (30) as

$$P(x_1, x_2, \varepsilon) = (\frac{1}{x_1} - [\frac{1}{x_1}], \frac{1}{[\frac{1}{x_1}] + x_2}, -\varepsilon)$$
$$= (T_G x_1, \frac{1}{[\frac{1}{x_1}] + x_2}, -\varepsilon). \qquad (32)$$

Obviously the x_1-direction is the expanding and the x_2-direction the contracting direction of the map P. The ergodic properties of the map P are determined by the behavior in the expanding directions of P [Rue78], [Bow75] which are (x_1, ε) with $\varepsilon = \pm 1$.

Apparently the geodesics in $\mathbb{H}$ are not closed. In contrast, due to the identifications through the group Γ there exist closed geodesics $\hat{\gamma}$ in M_Γ. Obviously this $\hat{\gamma}$ must define a periodic point of the Poincaré map P. An easy calculation then shows that for such a closed geodesic one has

$$x_1 = [\overline{n_0, n_1, \cdots, n_{m-2}, n_{m-1}}],\ x_2 = [\overline{n_{m-1}, n_{m-2}, \cdots, n_1, n_0}] \text{ and } \varepsilon = \pm 1,$$

where the bar denotes the periodic repetition of the sequence. For $\Gamma = \Gamma(1)$, suppose $\hat{\gamma}$ is a periodic geodesic in $M_{\Gamma(1)}$. Then the lifted geodesic γ in $\mathbb{H}$ belonging to $\mathcal{A}$ can be expressed in the coordinates as used in (19):

$$\gamma = ([\overline{n_0, n_1, \cdots, n_{m-2}, n_{m-1}}], [\overline{n_{m-1}, n_{m-2}, \cdots, n_1, n_0}], \varepsilon).$$

Every hyperbolic element $\sigma \in \Gamma$ determines a closed geodesic in M_Γ by fixing the two basepoints $\gamma_{-\infty}$ and $\gamma_{+\infty}$ on $\mathbb{R}$, namely $\sigma\gamma_{\pm\infty} = \gamma_{\pm\infty}$. The length $l(\gamma)$ of γ is known to be related to the trace of σ [Hej76]:

$$2 \cosh \frac{l(\gamma)}{2} = (e^{l(\gamma)})^{1/2} + (e^{l(\gamma)})^{-1/2} = |\mathrm{trace}(\sigma)|. \qquad (33)$$

3.4 The Poincaré map for the geodesic flow on M_Γ

Consider now any subgroup $\Gamma \subseteq \Gamma(1)$ with finite index $\mu = [\Gamma(1) : \Gamma] < \infty$. The surface $M_\Gamma = \Gamma\backslash\mathbb{H}$ is then a finite, in general ramified, covering surface of $M_{\Gamma(1)}$. Denote by $\pi_\Gamma : M_\Gamma \to M_{\Gamma(1)}$ and $\pi : \mathbb{H} \to M_\Gamma$ the corresponding projection maps. To generalize the construction of the Poincaré map to arbitrary subgroups Γ of $\Gamma(1)$ consider the set $\mathcal{A}_\Gamma = \cup_{i=1}^{\mu} g_i \mathcal{A}$ where $\mathcal{A}$ was defined in (16) and the g_i are representatives of the μ different equivalence classes $\{g_i\}$ in $\Gamma\backslash\Gamma(1)$. As a natural Poincaré section for the geodesic flow on M_Γ one can choose

$$X_\Gamma = \pi_* \Big(\bigcup_{i=1}^{\mu} g_i \bigcup_{y \in S,\, v(y) \in C(y)} (y, v(y)) \Big). \qquad (34)$$

which obviously is identical to $\pi_{\Gamma,*}{}^{-1}(X)$, where X is the Poincaré section in $M_{\Gamma(1)}$ and hence the lift of X to the covering sphere bundle $T_1 M_\Gamma$ of

$T_1 M_{\Gamma(1)}$. There is again a bijective map of $\mathcal{A}_\Gamma$ onto X_Γ, apart from possibly a finite set of geodesics in $\mathcal{A}$: for $\gamma \in \mathcal{A}_\Gamma$ there exists a $\{g_i\} \in \Gamma\backslash\Gamma(1)$ such that $g_i^{-1}\gamma \in \mathcal{A}$. To this $g_i^{-1}\gamma$ there corresponds exactly one point $(y, v(y))$ with $y \in S$ and $v(y) \in C(y)$ and hence there corresponds to γ exactly one point $g_i(y, v(y))$ with $y \in S$ and $v(y) \in C(y)$ which on the other hand defines exactly one point in X_Γ. If Q belongs to Γ then there exist again geodesics in $\mathcal{A}_\Gamma$ which correspond to the same point in the Poincaré section namely the geodesics γ with basepoints $\gamma_{\pm\infty} = \pm 1$ or ∓ 1 and perhaps some or all of their images $g_i\gamma$.

To determine the explicit form of the Poincaré map $P_\Gamma : X_\Gamma \to X_\Gamma$ we proceed as follows: consider an orbit intersecting X_Γ and the geodesic γ in $\mathcal{A}_\Gamma$ corresponding to this point in X_Γ. Then there exists a $\{g_i\} \in \Gamma\backslash\Gamma(1)$ with $g_i^{-1}\gamma \in \mathcal{A}$. ¿From our discussion of the group $\Gamma(1)$ we know that the geodesic $QT^{-n\varepsilon}g_i^{-1}\gamma \in \mathcal{A}$ describes the action of the Poincaré map for the group $\Gamma(1)$. Obviously, the geodesics γ and $QT^{-n\varepsilon}g_i^{-1}\gamma$ do not describe in general the same orbit for the geodesic flow on M_Γ. But there exists a unique $\{g_j\} \in \Gamma\backslash\Gamma(1)$ such that $g_j QT^{-n\varepsilon}g_i^{-1} \in \Gamma$: consider namely $g := hg_iT^{n\varepsilon}Q$ with $h \in \Gamma$ arbitrary, then $gQT^{-n\varepsilon}g_i^{-1} = h \in \Gamma$ and g defines a unique g_j with $\{g\} = \{g_j\} \in \Gamma\backslash\Gamma(1)$. Hence $g_jQT^{-n\varepsilon}g_i^{-1}\gamma \in g_j\mathcal{A} \in \mathcal{A}_\Gamma$ defines an unique point on the Poincaré surface X_Γ. That this point is the first point in X_Γ to which the orbit returns follows immediately from our discussion for $\Gamma(1)$. We can introduce now on X_Γ coordinates in analogy to the section X, namely if $\gamma' = g_i\gamma$ with $\{g_i\} \in \Gamma\backslash\Gamma(1)$ and $\gamma \in \mathcal{A}$ described as in (29) as $\gamma = (\varepsilon\gamma_{+\infty}^{-1}, -\varepsilon\gamma_{-\infty}, \varepsilon)$, then we can describe γ' as

$$\gamma' = (\varepsilon\gamma_{+\infty}^{-1}, -\varepsilon\gamma_{-\infty}, \varepsilon, \{g_i\}) .$$

In these coordinates the Poincaré map $P_\Gamma : X_\Gamma \to X_\Gamma$ then can be written as

$$\begin{aligned} P_\Gamma : [0,1] \times [0,1] \times \mathbb{Z}_2 \times \Gamma\backslash\Gamma(1) &\to [0,1] \times [0,1] \times \mathbb{Z}_2 \times \Gamma\backslash\Gamma(1) \\ (x_1, x_2, \varepsilon, \{g\}) &\mapsto (T_G x_1, \frac{1}{[\frac{1}{x_1}] + x_2}, -\varepsilon, \{gT^{n\varepsilon}Q\}) , \end{aligned} \tag{35}$$

where $n = n(x) = [\frac{1}{x_1}]$ is the integer part of $\frac{1}{x_1}$ for $x := (x_1, x_2, \varepsilon, \{g\})$. It is just an extension of the Poincaré map for $\Gamma(1)$ by the set $\Gamma\backslash\Gamma(1)$.

Obviously we get back from P_Γ the Poincaré map P in the case $\Gamma = \Gamma(1)$ since then $\Gamma\backslash\Gamma(1)$ consists just of one element. The obvious interpretation of the Poincaré map P_Γ in (35) is as the lift of P to the different sheets of T_1M_Γ as a μ-fold covering of the space $T_1M_{\Gamma(1)}$ characterized by the μ classes $\{g_i\}$, $i = 1, 2, \cdots, \mu$ of $\Gamma\backslash\Gamma(1)$. For the construction of the generalized Ruelle transfer operator for the geodesic flow on M_Γ we need the action of P_Γ along the expanding directions. Obviously they are the x_1 directions in the different sheets described by $\varepsilon = \pm 1$ and $\{g\} \in \Gamma\backslash\Gamma(1)$. Hence we get

$$P_\Gamma|_{ex} \; : \; [0,1] \times \mathbb{Z}_2 \times \Gamma\backslash\Gamma(1) \to [0,1] \times \mathbb{Z}_2 \times \Gamma\backslash\Gamma(1)$$
$$P_\Gamma|_{ex}(z,\varepsilon,\{g\}) = (T_G\, z, -\varepsilon, \{gT^{n\varepsilon}Q\}) = (\frac{1}{z} - [\frac{1}{z}], -\varepsilon, \{gT^{n\varepsilon}Q\})\,, \quad (36)$$

with $n = n(x) = [\frac{1}{z}]$ and $x = (z, \varepsilon, \{g\})$. This allows us now to construct the transfer operator for the geodesic flow on M_Γ.

4 The transfer operator for the geodesic flow on M_Γ

4.1 The transfer operator for modular subgroups $\Gamma \subseteq \Gamma(1)$

The generalized Ruelle transfer operator for an expanding map $\tau : M \to M$ with inverse temperature β and weight function $A : M \to \mathbb{R}$ is defined in [Rue78], [May91a] as

$$\mathcal{L}_{\tau,\beta} f(x) = \sum_{y \in \tau^{-1}x} \exp\left(-\beta\, A(y)\right) f(y)\,,$$

where the sum runs over all preimages of x under the map τ and where f is some function on M, which has to be chosen appropriately. The weight function in the case of the Poincaré map $P_\Gamma|_{ex}$ which describes the ergodic properties of the geodesic flow is the function $A_{P_\Gamma|_{ex}}(x) = log|T_G'(z)|$ with $x = (z,\varepsilon,\{g\})$, quite similar to the one for $\Gamma(1)$ [May91a]. Furthermore, the preimages $P_\Gamma|_{ex}^{-1} x$ of the point $x = (z,\varepsilon,\{g\})$ are just

$$\left\{ (\frac{1}{z+n}, -\varepsilon, \{gQT^{n\varepsilon}\}) \mid n \in \mathbb{N} \right\}, \quad (37)$$

because

$$\begin{aligned} P_\Gamma|_{ex}(\frac{1}{z+n}, -\varepsilon, \{gQT^{n\varepsilon}\}) &= (z, \varepsilon, \{gQT^{n\varepsilon}\, T^{-n(x')\varepsilon} Q\}\,) \\ &= (z,\varepsilon,\{g\})\,, \end{aligned} \quad (38)$$

where we used $n(x') = n((\frac{1}{z+n}, -\varepsilon, \{gQT^{n\varepsilon}\})) = [z+n] = n$. The transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma,\chi}$ for a subgroup $\Gamma \subseteq \Gamma(1)$ with finite index $\mu = [\Gamma(1) : \Gamma] < \infty$ and representation $\chi : \Gamma \to end\, V$ can be defined as follows

$$\tilde{\mathcal{L}}_\beta^{\Gamma,\chi} \underline{f}(z,\varepsilon,\{g\}) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \chi(gQT^{n\varepsilon}g'^{-1}) \underline{f}(\frac{1}{z+n}, -\varepsilon, \{gQT^{n\varepsilon}\})\,, \quad (39)$$

where g' is the unique element among $g_1, \cdots, g_\mu$, such that $\{g'\} = \{gQT^{n\varepsilon}\}$, that means $\Gamma g' = \Gamma g Q T^{n\varepsilon}$. Therefore $gQT^{n\varepsilon}g'^{-1} \in \Gamma$ and $\chi(gQT^{n\varepsilon}g'^{-1})$ is well defined as acting on the function $\underline{f}$ if $\underline{f}$ takes values in the space V. The exact properties of the functions in the variable z will be discussed later. Before doing this we will give a simple interpretation of the above transfer operator in terms of a generalized transfer operator for the group $\Gamma(1)$ with a certain representation .

4.2 Transfer operators for $\Gamma(1)$ with representation U^χ

If $\psi : \Gamma(1) \to end W$ is a finite dimensional representation on the vector space W then the transfer operator for the geodesic flow on the modular surface for $\Gamma(1)$ with representation ψ has the following form

$$\tilde{\mathcal{L}}_\beta^{\Gamma,\psi} \underline{f}(z,\varepsilon) = \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta}\, \psi(QT^{n\varepsilon})\, \underline{f}(\frac{1}{z+n}, -\varepsilon)\,, \tag{40}$$

where the function $\underline{f}$ takes values in the space W. This operator generalizes the transfer operator used in [CM98] to describe the Selberg zeta function for the geodesic flow on the modular surface for $\Gamma(1)$. Indeed, this operator will be shown to describe the generalized Selberg zeta function for the group $\Gamma(1)$ with representation ψ [VZ83]. To understand the form of the transfer operator in (39) for Γ we briefly recall the definition of the representation of a group G induced by a representation χ of a subgroup $G_1 \subseteq G$ with $[G : G_1] = m < \infty$.

Let $\chi : G_1 \to V$ be a representation of the subgroup G_1 of G. Then the representation U^χ of G induced from the representation χ of G_1 can be defined as follows [VZ83]: for $\underline{v} = (v_i)_{i=1}^m \in V^m$ one defines for $g \in G$

$$(U^\chi(g)\underline{v})_i = \chi(g_i\, g\, g_j^{-1})\, v_j \tag{41}$$

where the $g_i \in G$, $i = 1, \cdots, m$ are fixed by the condition that $G = \bigcup_{i=1}^m G_1 g_i$ with $g_1 = id$ the unit element in G, and g_j in (41) is uniquely determined by the condition $g_i g g_j^{-1} \in G_1$ that means $G_1 g_i g = G_1 g_j$. Coming back now to our transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma,\chi}$ in (39) and identifying the functions $\underline{f}(z,\varepsilon,\{g_i\}) \in V$, $i = 1, \cdots, \mu$ with the elements v_i in (41) and writing $(\underline{\underline{f}}(z,\varepsilon))_{\{g_i\}} = \underline{f}(z,\varepsilon,\{g_i\})$ we find

$$(\tilde{\mathcal{L}}_\beta^{\Gamma,\chi} \underline{\underline{f}}(z,\varepsilon))_{\{g\}} = \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta} \chi(gQT^{n\varepsilon}g'^{-1}) \underline{f}(\frac{1}{z+n}, -\varepsilon, \{g'\}) \tag{42}$$

and hence

$$\tilde{\mathcal{L}}_\beta^{\Gamma,\chi} \underline{\underline{f}}(z,\varepsilon) = \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta} U^\chi(QT^{n\varepsilon}) \underline{\underline{f}}(\frac{1}{z+n}, -\varepsilon) \tag{43}$$

Hence we have shown

Theorem 1. *For $\Gamma \subseteq \Gamma(1)$ a subgroup of $\Gamma(1)$ with finite index and $\chi : \Gamma \to end V$ a finite dimensional representation of Γ the transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma,\chi}$ of the geodesic flow on $\Gamma\backslash\mathbb{H}$ with representation χ of Γ is identical up to isomorphy to the transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma(1),U^\chi}$ of the geodesic flow on $\Gamma(1)\backslash\mathbb{H}$ with representation U^χ of $\Gamma(1)$ induced by the representation χ of Γ.*

For the special case when χ is the trivial 1-dimensional representation of Γ the transfer operator (43) can be reduced to

$$\tilde{\mathcal{L}}_\beta^{\Gamma,\chi}\underline{f}(z,\varepsilon)=\sum_{n=1}^{\infty}\left(\frac{1}{z+n}\right)^{2\beta}U^\chi(QT^{n\varepsilon})\underline{f}\left(\frac{1}{z+n},-\varepsilon\right) \tag{44}$$

where the representation U^χ is defined as $(U^\chi(g)\underline{v})_i=v_j$ where j is determined again by the condition $g_i g g_j^{-1}\in\Gamma$. Employing the function $\omega:\Gamma(1)\to\mathbb{R}$ with

$$\omega(g)=\begin{cases}1 & g\in\Gamma\,,\\ 0 & g\notin\Gamma\,,\end{cases}$$

one can also write

$$(U^\chi(g)\underline{v})_i=\sum_{j=1}^{\mu}\omega(g_i g g_j^{-1})\,v_j\,.$$

Consequently the linear operator U^χ in this case has the following matrix representation $\chi^\Gamma(g)$:

$$\chi^\Gamma(g)=\begin{pmatrix}\omega(g_1gg_1^{-1}) & \omega(g_1gg_2^{-1}) & \cdots & \omega(g_1gg_\mu^{-1})\\ \omega(g_2gg_1^{-1}) & \omega(g_2gg_2^{-1}) & \cdots & \omega(g_2gg_\mu^{-1})\\ \cdots & \cdots & \cdots & \cdots\\ \omega(g_\mu gg_1^{-1}) & \omega(g_\mu gg_2^{-1}) & \cdots & \omega(g_\mu gg_\mu^{-1})\end{pmatrix}. \tag{45}$$

The matrix χ^Γ is a μ-dimensional permutation matrix. In every row respectively column there is only one entry different from zero and its value is 1. For fixed μ, there exist only $\mu!$ different matrices of this type. Thus, for every $g\in G$ we can always find integers $r_1, r_2\in\mathbb{N}$ with $r_1<r_2$ such that $\chi^\Gamma(g^{r_1})=\chi^\Gamma(g^{r_2})$ and consequently $\chi^\Gamma(g^{r_2-r_1})=1$. That is, there always exists $r\in\mathbb{N}$ such that $0<r<\infty$ and $\chi^\Gamma(g^r)=1$.

As an example let us consider the group $\Gamma_0(2)$. According to (11) it is a subgroup of $\Gamma(1)$ with index 3 and has furthermore the following coset decomposition

$$\Gamma(1)=\Gamma_0(2)g_1\cup\Gamma_0(2)g_2\cup\Gamma_0(2)g_3\,,$$

where $g_1=id$, $g_2=Q$ and $g_3=QT$ are representatives of the three different equivalence classes in $\Gamma_0(2)\backslash\Gamma(1)$. The representations $\chi^{\Gamma_0(2)}$ of Q and T read

$$\chi^{\Gamma_0(2)}(Q)=\begin{pmatrix}\omega(id\,Q\,(id)^{-1}) & \omega(id\,Q\,(Q)^{-1}) & \omega(id\,Q(QT)^{-1})\\ \omega(QQ(id)^{-1}) & \omega(QQ(Q)^{-1}) & \omega(QQ(QT)^{-1})\\ \omega(QT\,Q(id)^{-1}) & \omega(QT\,Q(Q)^{-1}) & \omega(QT\,Q(QT)^{-1})\end{pmatrix}=\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$$

respectively

$$\chi^{\Gamma_0(2)}(T)=\begin{pmatrix}\omega(id\,T(id)^{-1}) & \omega(id\,T(Q)^{-1}) & \omega(id\,T(QT)^{-1})\\ \omega(Q\,T(id)^{-1}) & \omega(Q\,T(Q)^{-1}) & \omega(Q\,T(QT)^{-1})\\ \omega(QT\,T(id)^{-1}) & \omega(QT\,T(Q)^{-1}) & \omega(QT\,T(QT)^{-1})\end{pmatrix}=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}.$$

The induced representations χ^{Γ} of Q and T for the groups Γ_2, $\Gamma_0(2)$, $\Gamma^0(2)$, Γ_ϑ and $\Gamma^0(3)$ are summarized in the following table:

group Γ	$\chi^{\Gamma}(Q)$	$\chi^{\Gamma}(T)$	representatives of $\Gamma\backslash\Gamma(1)$
$\Gamma(1)$	1	1	id
Γ_2	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$g_1 = id$ $g_2 = Q$
$\Gamma_0(2)$	$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$	$g_1 = id$ $g_2 = Q$ $g_3 = QT$
$\Gamma^0(2)$	$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$	$\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$	$g_1 = id$ $g_2 = Q$ $g_3 = T$
Γ_ϑ	$\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$	$\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$	$g_1 = id$ $g_2 = T$ $g_3 = TQ$
$\Gamma^0(3)$	$\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}$	$\begin{pmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1\end{pmatrix}$	$g_1 = id$ $g_2 = T^{-1}$ $g_3 = T$ $g_4 = Q$

Table 2. *The induced representations χ^{Γ} of Q and T for different subgroups Γ of $\Gamma(1)$*

In the following we restrict our discussion to the trivial 1-dimensional representation χ of Γ, but our results can be extended without problem to the more general case of non-trivial χ's. The corresponding transfer operator with the representation (45) reads

$$\tilde{\mathcal{L}}_\beta^{\Gamma,\chi}\underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty}\left(\frac{1}{z+n}\right)^{2\beta}\chi^{\Gamma}(QT^{n\varepsilon})\underline{f}\left(\frac{1}{z+n},-\varepsilon\right). \tag{46}$$

4.3 Decomposition of the transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma}$

In general, the induced representation U^χ will be reducible, that is U^χ can be decomposed into its irreducible components

$$\chi^{\Gamma} = \oplus_i \chi_i^{\Gamma}. \tag{47}$$

In the special case of a normal subgroup $\Gamma \subseteq \Gamma(1)$ the quotient set $\Gamma\backslash\Gamma(1)$ is itself a group and the induced representation of the group $\Gamma\backslash\Gamma(1)$ is isomorphic to the right regular representation of the group $\Gamma\backslash\Gamma(1)$. Since $U^\chi(h) = 1$ for $h \in \Gamma$ the representation U^χ defines also a representation of $\Gamma\backslash\Gamma(1)$ by $U^\chi(\Gamma g) = U^\chi(g)$. Then one finds [Vin89]

$$\chi^{\Gamma} = \bigoplus_{\chi_i^{\Gamma}\in\chi^*(\Gamma\backslash\Gamma(1))} n_i\chi_i^{\Gamma}, \tag{48}$$

where $\chi^*(\Gamma\backslash\Gamma(1))$ denotes the set of all inequivalent unitary irreducible representations of $\Gamma\backslash\Gamma(1)$ and n_i the dimension of χ_i^Γ.

According to (48) the transfer operator (46) can then be decomposed as

$$\begin{aligned}
&\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}\,\underline{f}(z,\varepsilon) \\
&= \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta} \Big(\bigoplus_{\chi_i^\Gamma\in\chi^*(\Gamma\backslash\Gamma(1))} n_i\,\chi_i^\Gamma(QT^{n\varepsilon})\Big)\,\underline{f}_i(\frac{1}{z+n},-\varepsilon) \qquad (49)\\
&= \Big(\bigoplus_{\chi_i^\Gamma\in\chi^*(\Gamma\backslash\Gamma(1))} n_i\,\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi_i^\Gamma}\Big)\underline{f}_i(z,\varepsilon) \qquad (50)
\end{aligned}$$

with

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi_i^\Gamma}\,\underline{f}_i(z,\varepsilon) = \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta}\chi_i^\Gamma(QT^{n\varepsilon})\,\underline{f}_i(\frac{1}{z+n},-\varepsilon),$$

where the function $\underline{f}_i$ takes values in the representation space of χ_i^Γ and $\underline{f} = \oplus_i \underline{f}_i$.

The simplest example is the transfer operator for $\Gamma(1)$ with the trivial representation $\chi^\Gamma = 1$:

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),1} f(z,\varepsilon) = \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta}\, f(\frac{1}{z+n},-\varepsilon)\,. \qquad (51)$$

A non-trivial example is the transfer operator for Γ_2. The induced representations χ^{Γ_2} of the generators Q and T in table 2 read in this case

$$\chi^{\Gamma_2}(Q) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \chi^{\Gamma_2}(T) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,. \qquad (52)$$

Applying the matrices $M = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ respectively $M^{-1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ one gets

$$M\chi^{\Gamma_2}(Q)M^{-1} = M\chi^{\Gamma_2}(T)M^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,.$$

That is, under the above basis transformation the 2-dimensional representation χ^{Γ_2} in (52) can be decomposed into two 1-dimensional unitary irreducible representations $\chi^{\Gamma_2} = \chi_1^{\Gamma_2} \oplus \chi_{-1}^{\Gamma_2}$ with

$$\chi_1^{\Gamma_2} = 1 \quad \text{and} \quad \chi_{-1}^{\Gamma_2}(Q) = \chi_{-1}^{\Gamma_2}(T) = -1\,. \qquad (53)$$

Therefore the transfer operator for Γ_2 can be decomposed as

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^{\Gamma_2}}\,\underline{f}(z,\varepsilon) = \sum_{n=1}^\infty (\frac{1}{z+n})^{2\beta} \begin{pmatrix} \chi_1^{\Gamma_2}(QT^{n\varepsilon}) & 0 \\ 0 & \chi_{-1}^{\Gamma_2}(QT^{n\varepsilon}) \end{pmatrix} \underline{f}(\frac{1}{z+n},-\varepsilon)\,. \qquad (54)$$

Another non-trivial example of a normal subgroup of $\Gamma(1)$ is the principal congruence subgroup $\Gamma(2)$. The quotient group $\Gamma(2)\backslash\Gamma(1)$ has index 6 and is isomorphic to the permutation group S_3 in (8). The induced representation $\chi^{\Gamma(2)}$ of $\Gamma(2)$ can be decomposed as $\chi^{\Gamma(2)} = \chi_1^{\Gamma(2)} \oplus \chi_{-1}^{\Gamma(2)} \oplus \chi_2^{\Gamma(2)} \oplus \chi_2^{\Gamma(2)}$ with the following three unitary irreducible representations $\chi_i^{\Gamma(2)}$, $i = 1, -1, 2$ [Vin89]: $\chi_1^{\Gamma(2)}$ is the trivial 1-dimensional representation, i.e., $\chi_1^{\Gamma(2)}(Q) = \chi_1^{\Gamma(2)}(T) = 1$. $\chi_{-1}^{\Gamma(2)}$ is the non-trivial 1-dimensional representation with

$$\chi_{-1}^{\Gamma(2)}(Q) = \chi_{-1}^{\Gamma(2)}(T) = -1. \tag{55}$$

The representation $\chi_2^{\Gamma(2)}$ is 2-dimensional and is usually realized by the matrices

$$\hat{\chi}_2(Q) = \begin{pmatrix} -\cos\frac{2\pi}{3} & \sin\frac{2\pi}{3} \\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3} \end{pmatrix} \quad \text{and} \quad \hat{\chi}_2(T) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

An alternative choice more convenient for our later discussion is the following:

$$\tilde{\chi}_2(Q) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix} \quad \text{and} \quad \tilde{\chi}_2(T) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{56}$$

which is related to the former by $\tilde{\chi}_2 = M\,\hat{\chi}_2\,M^{-1}$ with

$$M = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{6} \\ -\frac{1}{2} & \frac{\sqrt{3}}{6} \end{pmatrix} \quad \text{and} \quad M^{-1} = \begin{pmatrix} 1 & -1 \\ \sqrt{3} & \sqrt{3} \end{pmatrix}.$$

In the following we choose $\chi_2^{\Gamma(2)} = \tilde{\chi}_2$. The transfer operator for $\Gamma(2)$ can then be decomposed as

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^{\Gamma(2)}} \underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \tag{57}$$
$$\times \begin{pmatrix} \chi_1^{\Gamma(2)}(QT^{n\varepsilon}) & & & 0 \\ & \chi_{-1}^{\Gamma(2)}(QT^{n\varepsilon}) & & \\ & & \chi_2^{\Gamma(2)}(QT^{n\varepsilon}) & \\ 0 & & & \chi_2^{\Gamma(2)}(QT^{n\varepsilon}) \end{pmatrix} \underline{f}(\frac{1}{z+1}, -\varepsilon).$$

The representation U^χ for non-normal subgroups Γ is in general also reducible. As examples we consider the groups $\Gamma_0(2)$, $\Gamma^0(2)$ and Γ_ϑ. The induced representations χ^Γ in table 2 for the groups $\Gamma \in \{\Gamma_0(2), \Gamma^0(2), \Gamma_\theta\}$ can be decomposed into the following irreducible components:

$$M_\Gamma \chi^\Gamma(Q) M_\Gamma^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & -1 \end{pmatrix} = \begin{pmatrix} \chi_1^\Gamma(Q) & 0 \\ 0 & \chi_2^\Gamma(Q) \end{pmatrix},$$

$$M_\Gamma \chi^\Gamma(T) M_\Gamma^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \chi_1^\Gamma(T) & 0 \\ 0 & \chi_2^\Gamma(T) \end{pmatrix}, \tag{58}$$

with

$$\begin{aligned} &\chi_1^\Gamma(Q)=1,\ \ \chi_1^\Gamma(T)=1 \quad \text{and} \\ &\chi_2^\Gamma(Q)=\begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix},\ \ \chi_2^\Gamma(T)=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned} \tag{59}$$

The matrices M_Γ and M_Γ^{-1} for the different groups Γ have the form

$$\begin{aligned} M_{\Gamma_0(2)} &= \begin{pmatrix} -\frac{1}{9} & -\frac{1}{9} & -\frac{1}{9} \\ -1 & -1 & 2 \\ -1 & 2 & -1 \end{pmatrix}, \quad M_{\Gamma_0(2)}^{-1} = \begin{pmatrix} -3 & -\frac{1}{3} & -\frac{1}{3} \\ -3 & 0 & \frac{1}{3} \\ -3 & \frac{1}{3} & 0 \end{pmatrix}, \\ M_{\Gamma^0(2)} &= \begin{pmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ 1 & 1 & -2 \\ -2 & 1 & 1 \end{pmatrix}, \quad M_{\Gamma^0(2)}^{-1} = \begin{pmatrix} 3 & 0 & -\frac{1}{3} \\ 3 & \frac{1}{3} & \frac{1}{3} \\ 3 & -\frac{1}{3} & 0 \end{pmatrix}, \\ M_{\Gamma_\vartheta(2)} &= \begin{pmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ -2 & 1 & 1 \\ 1 & -2 & 1 \end{pmatrix}, \quad M_{\Gamma_\vartheta(2)}^{-1} = \begin{pmatrix} 3 & -\frac{1}{3} & 0 \\ 3 & 0 & -\frac{1}{3} \\ 3 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}. \end{aligned}$$

According to (58) the transfer operators $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$ for $\Gamma \in \{\Gamma_0(2),\, \Gamma^0(2),\, \Gamma_\theta\}$ can then be decomposed as follows:

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma} \underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \begin{pmatrix} \chi_1^\Gamma(QT^{n\varepsilon}) & 0 \\ 0 & \chi_2^\Gamma(QT^{n\varepsilon}) \end{pmatrix} \underline{f}(\frac{1}{z+n}, -\varepsilon). \tag{60}$$

Comparing (55) with (53) and (59) with (56) one concludes

$$\chi_{-1}^{\Gamma_2} = \chi_{-1}^{\Gamma(2)} \text{ and } \chi_2^{\Gamma(2)} = \chi_2^\Gamma \ \text{ for } \ \Gamma \in \{\Gamma_0(2),\, \Gamma^0(2),\, \Gamma_\theta\}.$$

Using the notations

$$\chi_1(Q) = 1, \qquad \chi_1(T) = 1, \tag{61}$$

$$\chi_{-1}(Q) = -1, \qquad \chi_{-1}(T) = -1, \tag{62}$$

$$\chi_2(Q) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \chi_2(T) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{63}$$

one hence finds for the decompositions of the different induced representations χ^Γ of $\Gamma(1)$:

$$\begin{aligned} \chi^{\Gamma(1)} &= \chi_1\,, \\ \chi^{\Gamma_2} &= \chi_1 \oplus \chi_{-1}\,, \\ \chi^{\Gamma} &= \chi_1 \oplus \chi_2\,, \qquad \Gamma \in \{\Gamma_0(2), \Gamma^0(2), \Gamma_\vartheta\}, \\ \chi^{\Gamma(2)} &= \chi_1 \oplus \chi_{-1} \oplus \chi_2 \oplus \chi_2 \end{aligned} \tag{64}$$

and for the corresponding decompositions of the transfer operators $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$:

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),1} f(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} f(\frac{1}{z+n}, -\varepsilon)\,,$$

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^{\Gamma_2}} \underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \begin{pmatrix} \chi_1(QT^{n\varepsilon}) & 0 \\ 0 & \chi_{-1}(QT^{n\varepsilon}) \end{pmatrix} \underline{f}(\frac{1}{z+n}, -\varepsilon)\,,$$

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^{\Gamma}} \underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \begin{pmatrix} \chi_1(QT^{n\varepsilon}) & 0 \\ 0 & \chi_{2}(QT^{n\varepsilon}) \end{pmatrix} \underline{f}(\frac{1}{z+n}, -\varepsilon)\,,$$

$$\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^{\Gamma(2)}} \underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta}$$
$$\times \begin{pmatrix} \chi_1(QT^{n\varepsilon}) & & & 0 \\ & \chi_{-1}(QT^{n\varepsilon}) & & \\ & & \chi_2(QT^{n\varepsilon}) & \\ 0 & & & \chi_2(QT^{n\varepsilon}) \end{pmatrix} \underline{f}(\frac{1}{z+1}, -\varepsilon)\,.$$

As we shall see later, these decompositions of the transfer operators are closely related to the Venkov-Zograf factorization of the Selberg zeta function for the different subgroups $\Gamma \subseteq \Gamma(1)$. Before doing this we will first discuss spectral properties of the transfer operators.

4.4 Spectral properties of the transfer operator

In the following we write the transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$ in (46) for an arbitrary subgroup $\Gamma \subseteq \Gamma(1)$ with finite index simply as $\tilde{\mathcal{L}}$:

$$\tilde{\mathcal{L}}\,\underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \chi^\Gamma(QT^{n\varepsilon}) \underline{f}(\frac{1}{z+n}, -\varepsilon)\,. \tag{65}$$

Combining the functions $\underline{f}(z,\varepsilon)$ for $\varepsilon = \pm 1$ into the vector function

$$\underline{\underline{f}}(z) := \begin{pmatrix} \underline{f}(z,+1) \\ \underline{f}(z,-1) \end{pmatrix} = \begin{pmatrix} \underline{f}_+(z) \\ \underline{f}_-(z) \end{pmatrix}$$

the operator $\tilde{\mathcal{L}}$ can be written as

$$\tilde{\mathcal{L}}\,\underline{\underline{f}}(z) = \begin{pmatrix} 0 & \mathcal{L}_- \\ \mathcal{L}_+ & 0 \end{pmatrix} \begin{pmatrix} \underline{f}_+(z) \\ \underline{f}_-(z) \end{pmatrix} \tag{66}$$

with

$$(\mathcal{L}_+\,\underline{f}_+)(z) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \chi^\Gamma(QT^{n}) \underline{f}_-(\frac{1}{z+n})\,, \tag{67}$$

$$(\mathcal{L}_-\,\underline{f}_-)(z) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \chi^\Gamma(QT^{-n}) \underline{f}_+(\frac{1}{z+n})\,, \tag{68}$$

with μ-dimensional vectors $\underline{f}_{\pm}$. Inserting (67) and (68) into (66) gives

$$\tilde{\mathcal{L}}\,\underline{\underline{f}}(z) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \begin{pmatrix} 0 & \chi^{\Gamma}(QT^{-n}) \\ \chi^{\Gamma}(QT^{n}) & 0 \end{pmatrix} \underline{\underline{f}}(\frac{1}{z+n}) \tag{69}$$

with a 2μ-dimensional vector $\underline{\underline{f}}$.

Due to this structure trace $\tilde{\mathcal{L}}^n$ vanishes for odd n. The dynamical reason for this is that the orbits of the geodesic flow have at consecutive points of intersections with the Poincaré section different orientations described by $\varepsilon \to -\varepsilon$. Hence the Poincaré map P_Γ cannot have any fixed point. Only at the next intersection the orientation is the same and there can exist fixed points P_Γ^2 and so on.

An operator of the form (66) has very special spectral properties: Suppose $\tilde{\lambda}$ is an eigenvalue of the transfer operator $\tilde{\mathcal{L}}$ with eigenfunction $\underline{\underline{f}}(z) = \begin{pmatrix} \underline{f}_+(z) \\ \underline{f}_-(z) \end{pmatrix}$, i.e.,

$$\begin{pmatrix} 0 & \mathcal{L}_- \\ \mathcal{L}_+ & 0 \end{pmatrix} \begin{pmatrix} \underline{f}_+(z) \\ \underline{f}_-(z) \end{pmatrix} = \tilde{\lambda} \begin{pmatrix} \underline{f}_+(z) \\ \underline{f}_-(z) \end{pmatrix} . \tag{70}$$

Then $-\tilde{\lambda}$ is also an eigenvalue of $\tilde{\mathcal{L}}$ with eigenfunction $\begin{pmatrix} \underline{f}_+(z) \\ -\underline{f}_-(z) \end{pmatrix}$ as one can verify easily. Furthermore, equation (70) is equivalent to the equations

$$\mathcal{L}_+ \underline{f}_+(z) = \tilde{\lambda} \underline{f}_-(z) \quad \text{and} \quad \mathcal{L}_- \underline{f}_-(z) = \tilde{\lambda} \underline{f}_+(z) .$$

They imply immediately

$$\mathcal{L}_- \mathcal{L}_+ \underline{f}_+(z) = \tilde{\lambda} \mathcal{L}_- \underline{f}_-(z) = \tilde{\lambda}^2 \underline{f}_+(z)$$

and

$$\mathcal{L}_+ \mathcal{L}_- \underline{f}_-(z) = \tilde{\lambda} \mathcal{L}_+ \underline{f}_+(z) = \tilde{\lambda}^2 \underline{f}_-(z) .$$

That is, both the operators $\mathcal{L}_-\mathcal{L}_+$ respectively $\mathcal{L}_+\mathcal{L}_-$ have eigenvalue $\tilde{\lambda}^2$ with corresponding eigenfunction $\underline{f}_+(z)$ respectively $\underline{f}_-(z)$. Indeed the two operators $\mathcal{L}_-\mathcal{L}_+$ and $\mathcal{L}_+\mathcal{L}_-$ have the same eigenvalues: if $\underline{g}(z)$ is an eigenfunction with eigenvalue λ of $\mathcal{L}_-\mathcal{L}_+$ respectively $\mathcal{L}_+\mathcal{L}_-$, then $\mathcal{L}_+\underline{g}(z)$ respectively $\mathcal{L}_-\underline{g}(z)$ is an eigenfunction of $\mathcal{L}_+\mathcal{L}_-$ respectively $\mathcal{L}_-\mathcal{L}_+$ with the same eigenvalue, as long as the two operators $\mathcal{L}_+$ and $\mathcal{L}_-$ have trivial kernel. For the Fredholm determinant of $\tilde{\mathcal{L}}$ one finds

$$\begin{aligned} \det(1-\tilde{\mathcal{L}}) &= \det(1 - \begin{pmatrix} 0 & \mathcal{L}_- \\ \mathcal{L}_+ & 0 \end{pmatrix}) \\ &= \det(1-\mathcal{L}_+\mathcal{L}_-) = \det(1-\mathcal{L}_-\mathcal{L}_+) , \end{aligned} \tag{71}$$

under the assumption that the operator $\tilde{\mathcal{L}}$ is for instance nuclear in some Banach space in the sense of Grothendieck, which will be shown later.

The only difference in the definition of the two operators $\mathcal{L}_+$ in (67) and $\mathcal{L}_-$ in (68) is the sign of the power of T in the representation χ^Γ. Hence $\mathcal{L}_+$ and $\mathcal{L}_-$ are identical, iff $\chi^\Gamma(QT^n) = \chi^\Gamma(QT^{-n})$ for all $n \in \mathbb{N}$, i.e., iff $\chi^\Gamma(T^2) = 1$ holds. For example this is true for the groups $\Gamma(1)$, $\Gamma(2)$, $\Gamma_0(2)$ and $\Gamma^0(2)$. In this case we set $\mathcal{L}_+ = \mathcal{L}_- := \mathcal{L}$ and relation (71) then says

$$\det(1-\tilde{\mathcal{L}}) = \det(1-\mathcal{L}^2) = \det(1+\mathcal{L})\,\det(1-\mathcal{L})\,. \tag{72}$$

To understand in this case the spectrum of $\tilde{\mathcal{L}} = \begin{pmatrix} 0 & \mathcal{L} \\ \mathcal{L} & 0 \end{pmatrix}$, it's obviously enough to study the spectrum of the operator

$$\mathcal{L}\,\underline{f}(z) = \sum_{n=1}^{\infty}(\frac{1}{z+n})^{2\beta}\chi^\Gamma(QT^n)\,\underline{f}(\frac{1}{z+n}) \tag{73}$$

where, compared to (65), the dependence on the parameter ε has disappeared. Assuming again nuclearity of $\tilde{\mathcal{L}}$ its spectrum consists of all numbers $+\lambda$ respectively $-\lambda$ with λ an eigenvalue of $\mathcal{L}$. The corresponding eigenfunctions of $\tilde{\mathcal{L}}$ are $\begin{pmatrix} \underline{f} \\ \underline{f} \end{pmatrix}$ respectively $\begin{pmatrix} \underline{f} \\ -\underline{f} \end{pmatrix}$ with $\mathcal{L}^2\underline{f} = \lambda^2\underline{f}$. For $\Gamma = \Gamma(1)$ and χ^Γ the trivial representation (73) is nothing but the transfer operator for the Gauß transformation [May91a], [CM99]

$$\mathcal{L}_\beta f(z) = \sum_{n=1}^{\infty}(\frac{1}{z+n})^{2\beta}\,f(\frac{1}{z+n})\,.$$

4.5 The analytic continuation of the transfer operator

We have still to find an appropriate Banach space of functions on which the transfer operators we have discussed up to now in a formal way are well defined nuclear operators. Consider vector valued functions $\underline{\underline{f}}$ given by

$$\underline{\underline{f}}(z) = \bigoplus_{\varepsilon=\pm 1,\, i=1\cdots\mu} f_{\varepsilon\, i}(z) = \begin{pmatrix} f_{+1}(z) \\ \vdots \\ f_{+\mu}(z) \\ f_{-1}(z) \\ \vdots \\ f_{-\mu}(z) \end{pmatrix} \tag{74}$$

where $\mu = [\Gamma(1) : \Gamma]$ is the dimension of the induced representation χ^Γ. For general subgroups $\Gamma \subseteq \Gamma(1)$ we choose in analogy to the case $\Gamma(1)$

the functions $f_{\varepsilon,\,i}$ to belong to the Banach space $B(D)$ of all holomorphic functions on the disk [May91a]

$$D := \{z | z \in \mathbb{C},\, |z-1| < \frac{3}{2}\} \tag{75}$$

which are continuous on the closure $\bar{D}$. Consider then the transfer operator $\tilde{\mathcal{L}}_\beta$ as acting on the space $\oplus_{i=1\cdots 2\mu} B(D)$:

$$\tilde{\mathcal{L}}_\beta : \bigoplus_{i=1\cdots 2\mu} B(D) \to \bigoplus_{i=1\cdots 2\mu} B(D)\,. \tag{76}$$

Due to the contraction property [May91a] of the transformation $\psi_n(z) = \frac{1}{z+n}$ in $\tilde{\mathcal{L}}_\beta$ in (65) for all $n \in \mathbb{N}$ the transfer operator $\tilde{\mathcal{L}}_\beta$ is well defined on this space for $\Re\beta > \frac{1}{2}$ and even holomorphic in this half plane. We will show next, that the operator $\tilde{\mathcal{L}}_\beta$ can be analytically continued to a meromorphic family of operators in the entire complex β plane as follows:

In a first step we select a number $\kappa \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. According to our discussion in paragraph 4.3 there exists a smallest number r with $0 < r < \infty$ and $\chi^\Gamma(T^r) = 1$. Using this property and writing $n = r(n'-1)+m$, $1 \le m \le r$ and $n' \ge 1$ the transfer operator (65) can be written as

$$\begin{aligned}
&\tilde{\mathcal{L}}_\beta \underline{f}(z,\varepsilon) = \sum_{n=1}^{\infty} (\frac{1}{z+n})^{2\beta} \chi^\Gamma(QT^{n\varepsilon})\, \underline{f}(\frac{1}{z+n}, -\varepsilon) \\
&= \sum_{n=1}^{\infty}\sum_{m=1}^{r} \Big(\frac{1}{z+m+r(n-1)}\Big)^{2\beta} \chi^\Gamma(QT^{(m+r(n-1))\varepsilon})\, \underline{f}\Big(\frac{1}{z+m+r(n-1)}, -\varepsilon\Big) \\
&= \Big\{ \sum_{n=1}^{\infty}\sum_{m=1}^{r} \Big(\frac{1}{z+m+r(n-1)}\Big)^{2\beta} \chi^\Gamma(QT^{m\varepsilon}) \\
&\quad \times \Big[\underline{f}\Big(\frac{1}{z+m+r(n-1)}, -\varepsilon\Big) - \sum_{l=0}^{\kappa} \frac{\underline{f}^{(l)}(0,-\varepsilon)}{l\,!} \Big(\frac{1}{z+m+r(n-1)}\Big)^{l}\Big]\Big\} \\
&\quad + \sum_{l=0}^{\kappa}\sum_{m=1}^{r} \chi^\Gamma(QT^{m\varepsilon}) \frac{\underline{f}^{(l)}(0,-\varepsilon)}{l\,!} (\frac{1}{r})^{2\beta+l} \sum_{n=1}^{\infty} \Big(\frac{1}{n+(\frac{z+m}{r}-1)}\Big)^{2\beta+l}.
\end{aligned} \tag{77}$$

The last sum can be expressed by the Hurwitz zeta function $\zeta(s, z+1) := \sum_{n=1}^{\infty} (\frac{1}{n+z})^s$ as follows:

$$\sum_{n=1}^{\infty} \Big(\frac{1}{n+(\frac{z+m}{r}-1)}\Big)^{2\beta+l} = \zeta(2\beta+l, \frac{z+m}{r})\,. \tag{78}$$

Due to well known analyticity properties of this function and the property [MOS66]

$$\lim_{s\to 1} \Big[\zeta(s,z) - \frac{1}{s-1}\Big] = -\psi(z) = -\frac{\Gamma'(z)}{\Gamma(z)}\,, \tag{79}$$

where $\Gamma(z)$ is the gamma function and $\psi(z)$ is the psi function, the function (78) is meromorphic in the entire complex β-plane with poles of first order at those β values where $2\beta + l = 1$ with $l \in \mathbb{N}_0$. Substituting (78) into (77) shows that the operator $\tilde{\mathcal{L}}_\beta$ can be written as

$$\tilde{\mathcal{L}}_\beta = \tilde{\mathcal{A}}_\beta^{(\kappa)} + \tilde{\mathcal{L}}_\beta^\kappa \,, \tag{80}$$

with

$$\tilde{\mathcal{A}}_\beta^{(\kappa)} \underline{f}(z,\varepsilon) := \sum_{l=0}^{\kappa} (\frac{1}{r})^{2\beta+l} \sum_{m=1}^{r} \chi^\Gamma(QT^{m\varepsilon}) \frac{\underline{f}^{(l)}(0,-\varepsilon)}{l\,!} \zeta(2\beta+l, \frac{z+m}{r}) \tag{81}$$

and

$$\begin{aligned} \tilde{\mathcal{L}}_\beta^\kappa \underline{f}(z,\varepsilon) := \sum_{n=1}^{\infty} \sum_{m=1}^{r} \chi^\Gamma(QT^{m\varepsilon}) \Big(\frac{1}{z+m+r(n-1)}\Big)^{2\beta} \\ \times \Bigg[\underline{f}\Big(\frac{1}{z+m+r(n-1)}, -\varepsilon\Big) - \sum_{l=0}^{\kappa} \frac{\underline{f}^{(l)}(0,-\varepsilon)}{l\,!} \Big(\frac{1}{z+m+r(n-1)}\Big)^{l}\Bigg] . \end{aligned} \tag{82}$$

The operator $\tilde{\mathcal{L}}_\beta^\kappa$ is obviously holomorphic in the region $\Re\beta > -\frac{\kappa}{2}$ and the operator $\tilde{\mathcal{A}}_\beta^{(\kappa)}$ is meromorphic in $\mathbb{C}$ with possible poles only at $\beta = \beta_l = \frac{1-l}{2}$, $l = 0, 1, \ldots, \kappa$. Consequently, the operator $\tilde{\mathcal{L}}_\beta$ is meromorphic in the region $\Re\beta > -\frac{\kappa}{2}$ with possible poles at the points $\beta = \beta_l$. Since $\kappa \in \mathbb{N}_0$ was arbitrary, $\tilde{\mathcal{L}}_\beta$ is meromorphic in the entire β-plane.

In the limit $\beta \to \beta_\kappa = \frac{1-\kappa}{2}$ the term $l = \kappa$ in the sum of $\tilde{\mathcal{A}}_\beta^{(\kappa)}$ in (81) is the only one which can become singular: due to (79) we have

$$\zeta(2\beta+\kappa, z) = \frac{1}{2}\frac{1}{\beta-\beta_\kappa} + O(1) \qquad \text{for } \beta \to \beta_\kappa \,. \tag{83}$$

Hence the operator $\tilde{\mathcal{A}}_\beta^{(\kappa)}$ for $\beta \to \beta_\kappa$ behaves like

$$\tilde{\mathcal{A}}_\beta^{(\kappa)} \underline{f}(z,\varepsilon) = \Big(\sum_{m=1}^{r} \chi^\Gamma(QT^{m\varepsilon})\Big) \Big[\underline{a}_\kappa \frac{\underline{f}^{(\kappa)}(0,-\varepsilon)}{\beta-\beta_\kappa} + O(1)\Big] , \tag{84}$$

with

$$\underline{a}_\kappa := \frac{1}{2r\ \kappa\,!} . \tag{85}$$

Whether the operator $\tilde{\mathcal{A}}_\beta^{(\kappa)}$ has in $\beta = \beta_\kappa$ really a singularity, depends on the representation χ^Γ. For $r = 2$ and $\chi^\Gamma(Q) = \chi^\Gamma(T) = -1$ for example the operator $\mathcal{A}_\beta^{(\kappa)}$ is regular and therefore the operator $\tilde{\mathcal{L}}_\beta$ is holomorphic in the entire β-plane. The transfer operator for the group Γ_2 in the irreducible representation $\chi_2^{\Gamma_2}$ of the induced representation of χ^{Γ_2} belongs to this case.

4.6 Nuclearity of the transfer operator for $\Gamma \subseteq \Gamma(1)$

It is known that the transfer operator for the group $\Gamma(1)$ in the Banach space $B(D)$ is a nuclear operator [May91a] in the sense of Grothendieck [Gro55]. Here we will show that also the transfer operator for an arbitrary subgroup $\Gamma \subseteq \Gamma(1)$ with $[\Gamma(1):\Gamma] < \infty$ is a nuclear operator in the Banach space $\oplus_{n=1}^{2\mu} B(D)$.

For this consider the analytically continued transfer operator $\tilde{\mathcal{L}}_\beta$ as defined in (80). Obviously the operator $\tilde{\mathcal{A}}_\beta^{(\kappa)}$ for $\beta \neq \beta_l$, $l = 1, \cdots, \kappa$ is a nuclear operator of order zero, because $\tilde{\mathcal{A}}_\beta^{(\kappa)}$ has finite rank $\kappa + 1$. To show the operator $\tilde{\mathcal{L}}_\beta^\kappa$ to be nuclear, we write $\tilde{\mathcal{L}}_\beta^\kappa$ analogous to $\tilde{\mathcal{L}}_\beta$ in (69) as follows:

$$\tilde{\mathcal{L}}_\beta^\kappa \underline{\underline{f}}(z) := \sum_{n=1}^{\infty} \sum_{m=1}^{r} \begin{pmatrix} 0 & \chi^\Gamma(QT^{-m}) \\ \chi^\Gamma(QT^m) & 0 \end{pmatrix} \Big(\frac{1}{z+m+r(n-1)}\Big)^{2\beta}$$
$$\times \left[\underline{\underline{f}}\Big(\frac{1}{z+m+r(n-1)}\Big) - \sum_{l=0}^{\kappa} \frac{\underline{\underline{f}}^{(l)}(0)}{l\,!} \Big(\frac{1}{z+m+r(n-1)}\Big)^l \right] . \tag{86}$$

Suppose $\underline{\underline{f}} \in \mathbb{C}^{2\mu} \otimes B(D)$ with

$$\underline{\underline{f}} = \sum_{i=1}^{2\mu} \underline{e}_i \otimes f_i \,, \tag{87}$$

where $\{\underline{e}_1, \underline{e}_2, \cdots, \underline{e}_{2\mu}\}$ is a basis in $\mathbb{C}^{2\mu}$ and $f_i \in B(D)$ for $i = 1, 2, \cdots, 2\mu$. Then the operator $\tilde{\mathcal{L}}_\beta^\kappa$ in (82) can be written as

$$\tilde{\mathcal{L}}_\beta^\kappa \underline{f}(z)$$
$$= \sum_{i=1}^{2\mu} \sum_{m=1}^{r} \begin{pmatrix} 0 & \chi^\Gamma(QT^{-m}) \\ \chi^\Gamma(QT^m) & 0 \end{pmatrix} \underline{e}_i \otimes \sum_{n=1}^{\infty} \Big(\frac{1}{z+m+r(n-1)}\Big)^{2\beta}$$
$$\times \left[f_i\Big(\frac{1}{z+m+r(n-1)}\Big) - \sum_{l=0}^{\kappa} \frac{f_i^{(l)}(0)}{l\,!} \Big(\frac{1}{z+m+r(n-1)}\Big)^l \right] . \tag{88}$$

Using the notations $\chi_m : \mathbb{C}^{2\mu} \to \mathbb{C}^{2\mu}$

$$\chi_m := \begin{pmatrix} 0 & \chi^\Gamma(QT^{-m}) \\ \chi^\Gamma(QT^m) & 0 \end{pmatrix} ,$$

and

$$\mathcal{L}_{\beta,m} f(z) := \sum_{n=1}^{\infty} \Big(\frac{1}{z+m+r(n-1)}\Big)^{2\beta}$$
$$\times \left[f\Big(\frac{1}{z+m+r(n-1)}\Big) - \sum_{l=0}^{\kappa} \frac{f^{(l)}(0)}{l\,!} \Big(\frac{1}{z+m+r(n-1)}\Big)^l \right] \tag{89}$$

the operator (88) reads as

$$\tilde{\mathcal{L}}_\beta^\kappa \underline{f}(z) = \sum_{i=1}^{2\mu} \sum_{m=1}^{r} \chi_m \, \underline{e}_i \otimes \mathcal{L}_{\beta,\,m} f_i(z) \tag{90}$$

and hence we find

$$\tilde{\mathcal{L}}_\beta^\kappa = \sum_{m=1}^{r} \chi_m \otimes \mathcal{L}_{\beta,\,m} \, . \tag{91}$$

Since the operator $\mathcal{L}_{\beta,\,m}$ for $\Re\beta > -\frac{\kappa}{2}$ is a nuclear operator of order zero [May91a], the operator $\tilde{\mathcal{L}}_\beta^\kappa$ is also a nuclear operator of order zero for all $\Re\beta > -\frac{\kappa}{2}$, $\beta \neq \beta_\kappa$ and consequently of trace class. The same holds then for the transfer operator $\tilde{\mathcal{L}}_\beta$.

5 The Thermodynamic formalism approach to Selberg's zeta function for $\Gamma \subseteq \Gamma(1)$

Having established the trace class property for $\tilde{\mathcal{L}}_\beta$ we can now apply this operator to dynamical zeta functions and especially the Selberg zeta function for $\Gamma \subseteq \Gamma(1)$.

5.1 Generalized Selberg zeta functions for $PSL(2,\mathbb{Z})$

The dynamical zeta function $\zeta_{RS}(\beta;\chi)$ of Ruelle and Smale for the group $\Gamma(1)$ with some representation $\chi : \Gamma(1) \to GL(V)$ is defined as [Rue94]

$$\zeta_{RS}(\beta;\chi) = \prod_\gamma \left[\det\left(1 - \chi(\sigma_\gamma)\, e^{-\beta\, l(\gamma)}\right)\right]^{-1}, \tag{92}$$

where the product runs over all primitive periodic orbits[1] γ of the geodesic flow (25) on the modular surface, $l(\gamma)$ denotes the period of γ and σ_γ the hyperbolic element in the group $\Gamma(1)$ which fixes the geodesic γ in $\mathbb{H}$, i.e., $\sigma\gamma\gamma = \gamma$. Obviously σ_γ is determined only up to conjugation with some element in $\Gamma(1)$.

We briefly recall the description of the periodic orbits for the geodesic flow for $\Gamma(1)$ since for general subgroups $\Gamma \subseteq \Gamma(1)$ we will heavily rely on this. Consider the Poincaré map $P : X \to X$ in (32) given as

$$Px = P(x_1,\, x_2,\, \varepsilon) = (T_G x_1, \frac{1}{[\frac{1}{x_1}] + x_2}, -\varepsilon\,)\,, \tag{93}$$

[1] Notice that geodesics on $M_{\Gamma(1)}$ and geodesic orbits on $T_1 M_{\Gamma(1)}$ are not the same. But we use γ to denote both of them, since they are closely related to each other. $l(\gamma)$ stands for the period of the periodic geodesic orbit respectively the length of the geodesic.

where $x = \gamma \cap X$ denotes the point of intersection of γ with the Poincaré section X. The orbit γ is closed, iff there exists some $m \geq 1$ such that $P^m x = x$. This implies

$$T_G^m x_1 = x_1 \quad \text{and} \quad (-1)^m \varepsilon = \varepsilon \tag{94}$$

and therefore

$$x_1 = [\overline{n_0, n_1, \cdots, n_{m-2}, n_{m-1}}] \quad \text{with even } m\,.$$

A simple calculation shows that x_2 in (93) must then be of the form:

$$x_2 = [\overline{n_{m-1}, n_{m-2}, \cdots, n_1, n_0}].$$

That means up to conjugation a periodic orbit γ exactly corresponds to a geodesic with basepoints

$$\gamma_{-\infty} = -\varepsilon\,[\overline{n_{m-1}, n_{m-2}, \cdots, n_1, n_0}], \quad \gamma_{+\infty} = \varepsilon\,[\overline{n_0, n_1, \cdots, n_{m-2}, n_{m-1}}]^{-1}$$

in the representation (20). Since

$$QT^{n_{m-1}\varepsilon}QT^{-n_{m-2}\varepsilon}\cdots QT^{n_1\varepsilon}QT^{-n_0\varepsilon}\gamma_{\pm\infty} = \gamma_{\pm\infty},$$

this closed geodesic γ is fixed by the hyperbolic element

$$\sigma_\gamma = QT^{n_{m-1}\varepsilon}QT^{-n_{m-2}\varepsilon}\cdots QT^{n_1\varepsilon}QT^{-n_0\varepsilon} \quad \text{with even } m\,. \tag{95}$$

The period of such an orbit respectively the length of the corresponding closed geodesic can be expressed in terms of the recurrence time function (roof function) $r(x)$ [Pol86] for the geodesic flow

$$l(\gamma) = \sum_{k=0}^{m-1} r(P^k x)$$

where $x \in \gamma \cap X$ and $r(x)$ denotes the recurrence time respectively the length of the geodesic orbit γ from x to Px. For the geodesic flow for $\Gamma(1)$, this function $r(x)$ is equal to $r(x) = r((x_1, x_2, \varepsilon)) = log|T_G'(x_1)| = -log\, x_1^2$ [May91a]. Since the hyperbolic length is invariant under $PSL(2,\mathbb{R})$, this function is also the recurrence time function for the geodesic flow of any subgroup, $\Gamma \subseteq \Gamma(1)$ with respect to the Poincaré section X_Γ.

To apply this to the generalized Ruelle-Smale function in (92) we use a result of Ruelle [Rue94]:

Lemma 2. *Let $\tau : M \to M$ be a discrete time dynamical system and $\Phi : M \to GL(\mu, \mathbb{C})$ a matrix-valued function. Then the following identity holds*

$$\zeta_R(z, \Phi) = \prod_\gamma \Big[\det\Big(1 - z^{n(\gamma)} \prod_{k=0}^{n(\gamma)-1} \Phi(\tau^k \xi_\gamma)\Big)\Big]^{-1} \tag{96}$$

$$= \exp \sum_{n=1}^{\infty} \frac{z^m}{m} Z_m(\Phi) \tag{97}$$

where the first product runs over all primitive periodic orbit $\gamma \in M$ *with* $\xi_\gamma \in \gamma$ *and* $n(\gamma) \in \mathbb{N}$ *is the period of* γ*, i.e.,* $\tau^{n(\gamma)}\xi_\gamma = \xi_\gamma$*, and* $Z_m(\Phi)$ *denotes the partition function*

$$Z_m(\Phi) = \sum_{\xi \in Fix\, \tau^m} \operatorname{trace}\Big(\prod_{k=0}^{m-1} \Phi(\tau^k \xi) \Big) \tag{98}$$

where the sum runs over all fixed point of τ^m*.*

The dynamical zeta function (92) is a special case of $\zeta_R(z,\Phi)$ in (96) with $z = 1$, $\tau = P$ and

$$\Phi(x) = \chi\,(QT^{N(x)})\, \exp(-\beta\, r(x)\,)\,, \tag{99}$$

where $N(x) = -\varepsilon\, n(x)$ with $n(x) = [\frac{1}{x_1}]$ for $x = (x_1, x_2, \varepsilon) \in X$: inserting (99) into definition (96) namely gives

$$\begin{aligned}
&\zeta_{RS}(\beta;\chi) = \zeta_R(1,\Phi) \\
&= \prod_\gamma \Big[\det\, \big[\, 1 - \prod_{k=0}^{m-1} \chi(QT^{N(\,P^k\,x)})\, \exp\,\big(-\beta\, r(P^k x) \big) \,\big] \Big]^{-1}
\end{aligned}$$

with $P^m x = x$ and therefore

$$\begin{aligned}
\zeta_{RS}(\beta;\chi) &= \prod_\gamma \Big[\det \Big(1 - \chi(\sigma_\gamma)\, \exp\,\big(-\beta \sum_{k=0}^{m-1} r(P^{\,k} x)\big)\Big)\Big]^{-1} \\
&= \prod_\gamma \Big[\det \Big(1 - \chi(\sigma_\gamma)\, e^{-\beta\, l(\gamma)} \Big)\Big]^{-1}.
\end{aligned} \tag{100}$$

Due to (97) the dynamical zeta function (92) hence can be expressed as

$$\zeta_{RS}(\beta;\chi) = \sum_{m=1}^{\infty} \frac{Z_m(\beta;P;\chi)}{m}\,, \tag{101}$$

with Z_m the partition function in (98) and Φ defined in (99):

$$\begin{aligned}
Z_m(\beta;P;\chi) &= \sum_{x \in\, Fix\, P^m} \operatorname{trace}\Big[\prod_{k=0}^{m-1} \Phi(\,P^k\, x\,) \Big] \\
&= \sum_{x \in\, Fix\, P^m} \operatorname{trace}\Big[\prod_{k=0}^{m-1} [\,\chi\,(QT^{N\,(P^k x)}\,)\, \exp\,(-\beta\, r\,(\,P^k\,x\,)\,)\,] \Big] \\
&= \sum_{x \in\, Fix\, P^m} \operatorname{trace} \chi\,(\sigma_\gamma)\;\; \exp\,\Big(-\beta \sum_{k=0}^{m-1} r\,(\,P^k\,x\,) \Big)\,.
\end{aligned} \tag{102}$$

Z_m obviously vanishes for odd integers m, because according to (94) P^m doesn't have fixed points for odd m.

The generalized Selberg zeta function for Γ with representation χ is defined as [Ven90]

$$Z_S(\beta;\Gamma(1);\chi)=\prod_{\gamma}\prod_{k=0}^{\infty}\det\left(1-\chi(\sigma_\gamma)\,e^{-(\beta+k)\,l(\gamma)}\right). \tag{103}$$

Comparing with expression (92) shows that Selberg's zeta function can be expressed as

$$Z_S(\beta;\Gamma(1);\chi)=\prod_{k=0}^{\infty}\zeta_{RS}(\beta+k;\chi)^{-1} \tag{104}$$

in terms of the Ruelle-Smale dynamical zeta function with representation χ.

5.2 Dynamical zeta functions and Fredholm determinants of transfer operators

As in statistical mechanics the transfer operator in the thermodynamic formalism of dynamical systems is used to determine partition functions as in (98) [May80]. In the case of the geodesic flow on modular surfaces this can be done as follows:

Using the notation $Uz=\frac{1}{z}$ the transfer operator $\tilde{\mathcal{L}}_\beta$ in (69) can be expressed as

$$\begin{aligned}\tilde{\mathcal{L}}_\beta\underline{f}(z)&=\sum_{n_1=1}^{\infty}\begin{pmatrix}0&\chi^\Gamma(QT^{-n_1})\\ \chi^\Gamma(QT^{n_1})&0\end{pmatrix}\left(\frac{1}{n_1+z}\right)^{2\beta}\underline{f}\Big(\frac{1}{n_1+z}\Big)\\ &=\sum_{n_1=1}^{\infty}\begin{pmatrix}0&\chi^\Gamma(QT^{-n_1})\\ \chi^\Gamma(QT^{n_1})&0\end{pmatrix}(UT^{n_1}z)^{2\beta}\underline{f}(UT^{n_1}z)\,,\end{aligned}$$

where for short $\underline{\underline{f}}(z)$ has been replaced by $\underline{f}(z)$. The m-th iterate of the operator $\tilde{\mathcal{L}}_\beta$ is then equal to

$$\begin{aligned}&\tilde{\mathcal{L}}_\beta^m\underline{f}(z)=\underbrace{(\tilde{\mathcal{L}}_\beta\circ\tilde{\mathcal{L}}_\beta\circ\cdots\circ\tilde{\mathcal{L}}_\beta)}_{m-times}\underline{f}(z)\\ &=\sum_{n_1,n_2,\dots,n_m=1}^{\infty}\chi_{n_1,n_2,\dots,n_m}\\ &\quad\times\Big[(UT^{n_m}z)(UT^{n_{m-1}}UT^{n_m}z)\cdots(UT^{n_1}UT^{n_2}\cdots UT^{n_{m-1}}UT^{n_m}z)\Big]^{2\beta}\\ &\quad\times\underline{f}(UT^{n_1}UT^{n_2}\cdots UT^{n_{m-1}}UT^{n_m}z)\,,\end{aligned} \tag{105}$$

with

$$\chi_{n_1,n_2,\dots,n_m} = \begin{cases} \begin{pmatrix} \chi^\Gamma(\sigma_{n_1,n_2,\dots,n_m,-1}) & 0 \\ 0 & \chi^\Gamma(\sigma_{n_1,n_2,\dots,n_m,+1}) \end{pmatrix} & \text{for } m \text{ even} \\ \\ \begin{pmatrix} 0 & \chi^\Gamma(\sigma_{n_1,n_2,\dots,n_m,+1}) \\ \chi^\Gamma(\sigma_{n_1,n_2,\dots,n_m,-1}) & 0 \end{pmatrix} & \text{for } m \text{ odd} \end{cases} \tag{106}$$

and

$$\sigma_{n_1,n_2,\dots,n_m,\varepsilon} = QT^{(-1)^m n_m \varepsilon} QT^{(-1)^{m-1} n_{m-1}\varepsilon} \cdots QT^{n_2\varepsilon} QT^{-n_1\varepsilon}\,, \quad \varepsilon = \pm 1\,. \tag{107}$$

Since $\tilde{\mathcal{L}}_\beta$ is nuclear, also $\tilde{\mathcal{L}}_\beta^m$ is nuclear for all $m \in \mathbb{N}$. With the abbreviations

$$\psi_{n_1,n_2,\dots,n_m}(z) = UT^{n_1}UT^{n_2}\cdots UT^{n_{m-1}}UT^{n_m}z\,, \tag{108}$$

$$\varphi_{n_1,n_2,\dots,n_m}(z) = \Big[\prod_{k=1}^{m}(UT^{n_k}UT^{n_{k+1}}\cdots UT^{n_{m-1}}UT^{n_m}z)\Big]^{2\beta}\,,$$

$$\Phi_{n_1,n_2,\dots,n_m}(z) = \chi_{n_1,n_2,\dots,n_m}\ \varphi_{n_1,n_2,\dots,n_m}(z)\,, \tag{109}$$

(105) can be rewritten as

$$\tilde{\mathcal{L}}_\beta^m \underline{f}(z) = \sum_{n_1,n_2,\dots,n_m=1}^{\infty} \Phi_{n_1,n_2,\dots,n_m}(z)\,\underline{f}(\psi_{n_1,n_2,\dots,n_m}(z))\,. \tag{110}$$

In analogy we find

$$(-\tilde{\mathcal{L}}_{\beta+1})^m \underline{f}(z) = \sum_{n_1,n_2,\dots,n_m=1}^{\infty} \Phi_{n_1,n_2,\dots,n_m}(z)\,\psi'_{n_1,n_2,\dots,n_m}(z)\,\underline{f}(\psi_{n_1,n_2,\dots,n_m}(z))\,. \tag{111}$$

To calculate the trace of the operators $\tilde{\mathcal{L}}_\beta^m$ and $(-\tilde{\mathcal{L}}_{\beta+1})^m$, we follow the arguments in [May91a] :

The operator (110) has the stature

$$L\underline{f}(z) = \sum_i L_i \underline{f}(z)\,, \tag{112}$$

where i stands for the multi-index $(n_1, n_2, \dots, n_m)$ and L_i is a composition operator [May91a] of the form

$$L_i \underline{f}(z) = \Phi_i(z)\,\underline{f}(\psi_i(z))\,, \tag{113}$$

with Φ_i a matrix-valued function. To determine trace L_i consider the eigenfunction equation

$$L_i \underline{f}(z) = \Phi_i(z) \underline{f}(\psi_i(z)) = \lambda \underline{f}(z)\,. \tag{114}$$

Since ψ_i in (108) is holomorphic in the disk D in (75) and maps the disk D strictly inside itself, ψ_i has exactly one fixed point z_i^* in D [Rue76]. At the point $z = z_i^*$ equation (114) reads

$$\Phi_i(z_i^*) \underline{f}(z_i^*) = \lambda \underline{f}(z_i^*)\,,$$

i.e., if $\underline{f}(z_i^*) \neq \underline{0}$, then λ must be an eigenvalue of $\Phi(z_i^*)$ with corresponding eigenvector $\underline{f}(z_i^*)$. If $\underline{f}(z_i^*) = \underline{0}$, then we differentiate equation (114) with respect to z and get

$$D\Phi_i(z) \underline{f}(\psi_i(z)) + \Phi_i(z) \underline{f}'(\psi_i(z)) \psi_i'(z) = \lambda \underline{f}'(z)\,. \tag{115}$$

Since $\underline{f}(z_i^*) = \underline{0}$, setting $z = z_i^*$ in (115) one gets the equation

$$\psi_i'(z_i^*) \Phi_i(z_i^*) \underline{f}'(z_i^*) = \lambda \underline{f}'(z_i^*) \tag{116}$$

and therefore $\lambda = \rho \psi_i'(z_i^*)$, with ρ an eigenvalue of the matrix $\Phi_i(z_i^*)$ with eigenvector $\underline{f}'(z_i^*) \neq \underline{0}$. If also $\underline{f}'(z_i^*) = \underline{0}$, one differentiates equation (115) again. Repeating the argument shows that the eigenvalues of L_i must belong to the set $\{\rho \psi_i'(z_i^*)^n\}$, $n \in \mathbb{N}_0$, where ρ runs over all eigenvalues of the matrix $\Phi_i(z_i^*)$.

Conversely, to show that every one of the numbers $\lambda_n := \rho \psi_i'(z_i^*)^n$ really belongs to the spectrum of L_i, one must show $(L_i - \rho \psi_i'(z_i^*)^n)$ is not invertible, that means there is no solution $\underline{f}(z)$ for the equation

$$(L_i - \rho \psi_i'(z_i^*)^n) \underline{f}(z) = \underline{g}(z) \tag{117}$$

respectively

$$\Phi_i(z) \underline{f}(\psi_i(z)) - \rho \psi_i'(z_i^*)^n \underline{f}(z) = \underline{g}(z) \tag{118}$$

for a certain function $\underline{g}(z)$ in $\oplus_{i=1}^{2\mu} B(D)$. Let us choose a function $\underline{g}(z)$ with the properties

$$\underline{g}^{(k)}(z_i^*) = 0 \quad \text{for } 0 \leq k \leq n-1 \quad \text{and} \quad \underline{g}^{(n)}(z_i^*) \neq \underline{0}\,. \tag{119}$$

Then for $n = 0$ it follows from (118) that

$$(\Phi_i(z_i^*) - \rho) \underline{f}(z_i^*) = \underline{g}(z_i^*) \neq \underline{0}\,. \tag{120}$$

One then chooses $\underline{g}(z_i^*)$ such that this matrix equation doesn't have a solution $\underline{f}(z_i^*)$. This is possible since ρ is an eigenvalue of $\Phi_i(z_i^*)$ and thus $(\Phi_i(z_i^*) - \rho)$ is not invertible. Therefore $\lambda_0 = \rho$ is in the spectrum of L_i.

For $n \geq 1$ we have to consider two cases. First, suppose $\rho\psi_i'(z_i^*)^k$ are not eigenvalues of $\Phi_i(z_i^*)$ for all $k \in \mathbb{N}$. Equation (118) at $z = z_i^*$ then reads:

$$(\Phi_i(z_i^*) - \rho\psi_i'(z_i^*)^n)\underline{f}(z_i^*) = \underline{g}(z_i^*) = \underline{0}. \tag{121}$$

This implies immediately $\underline{f}(z_i^*) \equiv \underline{0}$, because $\rho\psi_i'(z_i^*)^n$ is not an eigenvalue of $\Phi_i(z_i^*)$. Differentiating equation (118) once and setting $z = z_i^*$ gives

$$D\Phi_i(z_i^*)\underline{f}(z_i^*) + \Phi_i(z_i^*)\psi_i'(z_i^*)\underline{f}'(z_i^*) - \rho\psi_i'(z_i^*)^n\underline{f}'(z_i^*) = \underline{g}'(z_i^*) = \underline{0}\,.$$

Since $\underline{f}(z_i^*) = \underline{0}$ we get

$$\psi_i'(z_i^*)\,(\,\Phi_i(z_i^*) - \rho\psi_i'(z_i^*)^{n-1}\,)\,\underline{f}'(z_i^*) = g'(z_i^*) = \underline{0}$$

and therefore $\underline{f}'(z_i^*) \equiv \underline{0}$, because $\rho\psi_i'(z_i^*)^{n-1}$ is not an eigenvalue of $\Phi_i(z_i^*)$. Repeating this argument n-times one finds

$$\psi_i'(z_i^*)^n(\Phi_i(z_i^*) - \rho)\underline{f}^{(n)}(z_i^*) = \underline{g}^{(n)}(z_i^*) \neq \underline{0}\,. \tag{122}$$

Choose now the vector $\underline{g}^{(n)}(z_i^*)$ such that equation (122) doesn't have a solution $\underline{f}^{(n)}(z_i^*)$. This is possible, because ρ is an eigenvalue of $\Phi_i(z_i^*)$. This shows that for a function $\underline{g}(z)$ obeying the conditions (119) equation (117) does not have a solution and therefore all $\rho\psi_i'(z_i^*)^n$ with $n \in \mathbb{N}_0$ are eigenvalues of L_i. The fact that the eigenvalues $\rho\psi_i'(z_i^*)^n$ for a given ρ are indeed simple follows from arguments similar to the ones given in [May91a]. The trace of L_i is therefore the sum of the geometrical series $\rho\psi_i'(z_i^*)^n$ summed over all eigenvalues ρ of $\Phi_i(z_i^*)$, i.e.,

$$\operatorname{trace} L_i = \frac{\operatorname{trace} \Phi_i(z_i^*)}{1 - \psi_i'(z_i^*)}\,. \tag{123}$$

In the second case suppose some of the numbers $\rho\psi_i'(z_i^*)^k$ for $k \in \mathbb{N}$ are eigenvalues of $\Phi_i(z_i^*)$. Then we consider a new transformation $\psi_{i,\delta} : D \to D$ slightly deformed in a δ-neighborhood of the transformation ψ_i. Obviously $\psi_{i,\delta}$ will have new fixed points $z_{i,\delta}^*$, slightly different from the z_i^*. We choose $\psi_{i,\delta}$ such that all $\rho\psi_{i,\delta}'(z_{i,\delta}^*)^k$ for $k \in \mathbb{N}_0$ don't belong to the spectrum of $\Phi_i(z_i^*)$. This is possible, because $\Phi_i(z_i^*)$ in (109) is a finite-dimensional permutation matrix and has only finitely many non-vanishing eigenvalues and $\rho\psi_{i,\delta}'(z_{i,\delta}^*)^n$ converges to zero for large n. Repeating the arguments of the first case shows

$$\operatorname{trace} L_{i,\delta} = \frac{\operatorname{trace} \Phi_i(z_{i,\delta}^*)}{1 - \psi_{i,\delta}'(z_{i,\delta}^*)}$$

Since the trace is continuous in ψ_i, taking the limit $\delta \to 0$ one finds

$$\lim_{\delta\to 0} \operatorname{trace} L_{i,\delta} = \operatorname{trace} L_i \;=\; \frac{\operatorname{trace} \Phi_i(z_i^*)}{1 - \psi_i'(z_i^*)}$$

as in (123).

Summarizing, we have therefore shown for the operator L in (112):

$$\operatorname{trace} L = \sum_i \operatorname{trace} L_i = \sum_i \frac{\operatorname{trace} \Phi_i(z_i^*)}{1 - \psi_i'(z_i^*)} . \tag{124}$$

With the definition

$$L_i^{(s)} f(z) = \Phi_i(z) \, (\psi_i'(z))^{\,s} f(\psi_i(z)), \quad s = 0, 1$$

and $L^{(s)} = \sum_i L_i^{(s)}$ one finally gets the relation

$$\begin{aligned} \operatorname{trace} L^{(0)} - \operatorname{trace} L^{(1)} &= \sum_i \left(\operatorname{trace} L_i^{(0)} - \operatorname{trace} L_i^{(1)} \right) \\ &= \sum_i \operatorname{trace} \Phi_i(z_i^*) \, . \end{aligned} \tag{125}$$

Applying this relation to the operators (110) and (111) yields

$$\begin{aligned} &\operatorname{trace} \tilde{\mathcal{L}}_\beta^m - \operatorname{trace}(-\tilde{\mathcal{L}}_{\beta+1})^m \\ &= \sum_{n_1, n_2, \dots, n_m = 1}^{\infty} \operatorname{trace} (\chi_{n_1, n_2, \dots, n_m}) \, \varphi_{n_1, n_2, \dots, n_m}(z^*_{n_1, n_2, \dots, n_m}) \end{aligned} \tag{126}$$

with $z^*_{n_1, n_2, \dots, n_m}$ the fixed point of $\psi_{n_1, n_2, \dots, n_m}(z)$ in (108), i.e.,

$$UT^{n_1} UT^{n_2} \cdots UT^{n_{m-1}} UT^{n_m} z^*_{n_1, n_2, \dots, n_m} = z^*_{n_1, n_2, \dots, n_m}$$

and consequently

$$z^*_{n_1, n_2, \dots, n_m} = [\overline{n_1, n_2, \cdots, n_{m-1}, n_m}] \, . \tag{127}$$

This however is also a fixed point of T_G^m in (94). For this fixed point one has

$$\begin{aligned} \varphi_{n_1, n_2, \dots, n_m}(z^*_{n_1, n_2, \dots, n_m}) &= \Big(\prod_{k=1}^{m} [\overline{n_k, n_{k+1}, \cdots, n_m, n_1, \cdots, n_{k-1}}] \Big)^{2\beta} \\ &= \Big(\prod_{k=0}^{m-1} T_G^k \, z^*_{n_1, n_2, \dots, n_m} \Big)^{2\beta} . \end{aligned} \tag{128}$$

Furthermore, it follows from (106) that

$$\operatorname{trace}(\chi_{n_1, n_2, \dots, n_m}) = \begin{cases} 0 & m \text{ odd}, \\ \sum_{\varepsilon = \pm 1} \operatorname{trace} \chi^{\Gamma} \, (\sigma_{n_1, n_2, \dots, n_m, \varepsilon}) & m \text{ even}. \end{cases} \tag{129}$$

Inserting (128) and (129) in (126) one gets for odd m

$$\operatorname{trace} \tilde{\mathcal{L}}_\beta^m - \operatorname{trace}(-\tilde{\mathcal{L}}_{\beta+1})^m = 0$$

and for even m

$$\operatorname{trace}\tilde{\mathcal{L}}_\beta^m - \operatorname{trace}(-\tilde{\mathcal{L}}_{\beta+1})^m = \sum_{z\in Fix\, T_G^m} \sum_{\varepsilon=\pm1} \operatorname{trace}\chi^\Gamma(\sigma_{z,\varepsilon}) \Big(\prod_{k=0}^{m-1} T_G^k\, z \Big)^{2\beta}, \tag{130}$$

where the fixed points $z^*_{n_1,n_2,\dots,n_m}$ of $\psi_{n_1,n_2,\dots,n_m}(z)$ are replaced by the fixed points z of T_G^m and the element $\sigma_{n_1,n_2,\dots,n_m,\varepsilon}$, which depends on z, is abbreviated as $\sigma_{z,\varepsilon}$. The summation on the right-hand side of (130) gives nothing but the partition function $Z_m(\beta;P;\chi)$ in (102) for $\chi=\chi^\Gamma$. Notice that σ_γ in formula (102) coincides with $\sigma_{n_1,n_2,\dots,n_m,\varepsilon}$ in (107) and can be expressed in the matrix form (106). Hence we find

$$\operatorname{trace}\tilde{\mathcal{L}}_\beta^m - \operatorname{trace}(-\tilde{\mathcal{L}}_{\beta+1})^m = Z_m(\beta;P;\chi^\Gamma)\,, \tag{131}$$

which is valid not only for even but also for m odd, because for m odd both sides of (131) vanish.

Finally we can express the dynamical zeta function in (101) in terms of the transfer operator $\tilde{\mathcal{L}}_\beta$ as follows:

$$\begin{aligned}\zeta_{RS}(\beta;\chi^\Gamma) &= \exp\sum_{m=1}^\infty \frac{Z_m(\beta;P;\chi^\Gamma)}{m}\\ &= \exp\sum_{m=1}^\infty \frac{1}{m}(\operatorname{trace}\tilde{\mathcal{L}}_\beta^m - \operatorname{trace}(-\tilde{\mathcal{L}}_{\beta+1})^m)\\ &= \frac{\exp\sum_{m=1}^\infty \frac{1}{m}\operatorname{trace}\tilde{\mathcal{L}}_\beta^m}{\exp\sum_{m=1}^\infty \frac{1}{m}\operatorname{trace}(-\tilde{\mathcal{L}}_{\beta+1})^m}\,.\end{aligned}$$

Since $\operatorname{trace}\tilde{\mathcal{L}}_\beta^m = 0$ for m odd, this yields immediately

$$\zeta_{RS}(\beta;\chi^\Gamma) = \frac{\exp\sum_{m=1}^\infty \frac{1}{m}\operatorname{trace}\tilde{\mathcal{L}}_\beta^m}{\exp\sum_{m=1}^\infty \frac{1}{m}\operatorname{trace}\tilde{\mathcal{L}}_{\beta+1}^m}\,. \tag{132}$$

Applying the identity [Gro56]

$$\begin{aligned}\det(1-z\mathcal{L}) &= \exp(\operatorname{trace} log\,(1-z\mathcal{L})) = \exp\Big(-\operatorname{trace}\sum_{m=1}^\infty\frac{(z\mathcal{L})^m}{m}\Big)\\ &= \frac{1}{\exp\sum_{m=1}^\infty \frac{1}{m}\operatorname{trace}(z\mathcal{L})^m}\end{aligned}$$

for nuclear operators $\mathcal{L}$ of order zero, we obtain from (132) the formula

$$\zeta_{RS}(\beta;\chi^\Gamma) = \frac{\det(1-\tilde{\mathcal{L}}_{\beta+1})}{\det(1-\tilde{\mathcal{L}}_\beta)}\,. \tag{133}$$

5.3 The Selberg zeta function for subgroups $\Gamma \subseteq \Gamma(1)$

Equality (133) gives a connection between the transfer operator $\tilde{\mathcal{L}}_\beta$ and the dynamical zeta function $\zeta_{RS}(\beta;\chi^\Gamma)$. Since on the other hand the dynamical zeta function is related to Selberg's zeta function (104) through [Ven90], [Rue94], [Cha99]

$$Z_S(\beta;\Gamma(1);\chi^\Gamma) = \prod_{k=0}^{\infty} \zeta_{RS}(\beta+k;\chi^\Gamma)^{-1}\,, \tag{134}$$

Selberg's zeta gets expressed in terms of the transfer operator by inserting (133) into (134):

Theorem 3. *Let Γ be a subgroup of $\Gamma(1)$ of finite index and χ^Γ be the representation of $\Gamma(1)$ induced from the trivial representation of the subgroup Γ. Furthermore, let $\tilde{\mathcal{L}}_\beta^\Gamma$ respectively $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$ be the transfer operators for the geodesic flows on $\Gamma\backslash\mathbb{H}$ respectively $\Gamma(1)\backslash\mathbb{H}$ with representation χ^Γ. Then the Selberg zeta functions $Z_S(\beta;\Gamma)$ respectively $Z_S(\beta;\Gamma(1);\chi^\Gamma)$ are related to the Fredholm determinants of these transfer operators by*

$$Z_S(\beta;\Gamma(1);\chi^\Gamma) = \det(1-\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}) = det(1-\tilde{\mathcal{L}}_\beta^\Gamma) = Z_S(\beta;\Gamma)\,. \tag{135}$$

In the framework of the traditional approach to Selberg's zeta function by means of Selberg's trace formula the following relation between Selberg's zeta functions for subgroups of different Fuchsian groups with representations is well known [VZ83]:

Theorem 4. *Let G be an arbitrary Fuchsian group of the first kind and G_1 be a subgroup of G of finite index. Let χ be an arbitrary finite-dimensional unitary representation of G_1 and U^χ the representation of G induced from χ. Then one has*

$$Z_S(\beta;G_1;\chi) = Z_S(\beta;G;U^\chi)\,. \tag{136}$$

Our Theorem 2 is just a dynamical proof of this result for the special case $G = \Gamma(1)$, $G_1 = \Gamma$ and χ the trivial representation.

Remark: Obviously, as soon as one can establish the transfer operator approach for a general Fuchsian group, Theorem 3 follows exactly along the line of arguments we have given in the case $G =$
$Gamma(1)$.

5.4 Factorization of the Selberg zeta functions

As mentioned in paragraph 4.3 the decomposition of the induced representation χ^Γ in its irreducible components implies also a decomposition of the transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$. Combined with (135) this implies a factorization of the Fredholm determinant $\det(1-\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma})$ and hence the factorization of the Selberg zeta function $Z_S(\beta;\Gamma(1);\chi^\Gamma)$.

Let χ^Γ be the representation of $\Gamma(1)$ induced from the trivial representation of the subgroup Γ of finite index which decomposes as $\chi^\Gamma = \oplus_i \chi_i^\Gamma$. Due to (47) the Fredholm determinant of $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$ can then be factorized as

$$\det(1-\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}) = \prod_i \det(1-\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi_i^\Gamma})\,.$$

This implies immediately a factorization of the Selberg zeta function as

$$Z_S(\beta;\Gamma(1);\chi^\Gamma) = \prod_i Z_S(\beta;\Gamma(1);\chi_i^\Gamma)\,.$$

If Γ is a normal subgroup of $\Gamma(1)$, the decomposition of the induced representation χ^Γ in (48) implies

$$\det(1-\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}) = \prod_{\chi_i^\Gamma\in\chi^*(\Gamma\backslash\Gamma(1))} \det(1-\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi_i^\Gamma})^{dim\,\chi_i^\Gamma}\,, \tag{137}$$

where $\chi^*(\Gamma\backslash\Gamma(1))$ denotes the set of all inequivalent irreducible unitary representations of the group $\Gamma\backslash\Gamma(1)$. This leads to the factorization of the Selberg zeta function as

$$Z_S(\beta;\Gamma(1);\chi^\Gamma) = \prod_{\chi_i^\Gamma\in\chi^*(\Gamma\backslash\Gamma(1))} Z(\beta;\Gamma(1);\chi_i^\Gamma)^{dim\,\chi_i^\Gamma}\,, \tag{138}$$

which is just a special case of the result of Venkov and Zograf in [VZ83]:

Proposition 5. *Let G be an arbitrary Fuchsian group of the first kind and G_1 be a normal subgroup of finite index in G. Let χ be an arbitrary finite dimensional unitary representation of G_1 and U^χ be the representation of G induced from χ. Then one has*

$$Z_S(\beta;G;U^\chi) = \prod_{\chi_i^\Gamma\in\chi^*(G_1\backslash G)} Z_S(\beta;G;\chi_i^\Gamma)^{dim\,\chi_i^\Gamma}\,.$$

If the subgroup Γ of $\Gamma(1)$ is not normal, the corresponding transfer operator $\tilde{\mathcal{L}}_\beta^{\Gamma(1),\chi^\Gamma}$ is in general also reducible. It follows from (64) that for example for the subgroups $\Gamma_0(2)$, $\Gamma^0(2)$ and Γ_ϑ the induced representation χ^Γ always contains the trivial representation $\chi^{\Gamma(1)}$. We will see in the second part of our paper that this decomposition of the transfer operators is closely related to the theory of new and old automorphic forms for subgroups of $Gamma(1)$. That the transfer operator $\mathcal{L}_\beta^{\Gamma(1)}$ is contained in the transfer operator for any subgroup just reflects the fact that any automorphic form for $\Gamma(1)$ is also an automorphic form for any of its subgroups.

According to the decomposition of the induced representations in (64), Selberg's zeta function $Z_S(\beta,\Gamma,1)$ for these groups can be factorized as followed:

$$\begin{aligned}
Z_S(\beta,\Gamma(1),1) &= \det(1-\tilde{\mathcal{L}}_\beta^{\chi_1})\,,\\
Z_S(\beta,\Gamma_2,1) &= \det(1-\tilde{\mathcal{L}}_\beta^{\chi_1})\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_{-1}})\,,\\
Z_S(\beta,\Gamma,1) &= \det(1-\tilde{\mathcal{L}}_\beta^{\chi_1})\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_2})\,,\quad \Gamma\in\{\Gamma_0(2),\,\Gamma^0(2),\,\Gamma_\vartheta\}\,,\\
Z_S(\beta,\Gamma(2),1) &= \det(1-\tilde{\mathcal{L}}_\beta^{\chi_1})\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_{-1}})\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_2})\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_2})\,,
\end{aligned} \tag{139}$$

where $\tilde{\mathcal{L}}_\beta^{\chi_i}$ denotes the transfer operator for $\Gamma(1)$ with representation χ_i, $i=1,-1,2$ as given in (61), (62) and (63). Relation (72) implies the further factorization

$$\det(1-\tilde{\mathcal{L}}_\beta^{\chi_i}) = \det(1-\mathcal{L}_\beta^{\chi_i})\,\det(1+\mathcal{L}_\beta^{\chi_i})\,.$$

Obviously, the relations in (139) imply also the following identities:

$$\begin{aligned}
Z_S(\beta,\Gamma(1),1) &= \det(1-\tilde{\mathcal{L}}_\beta^{\chi_1})\,,\\
Z_S(\beta,\Gamma_2,1) &= Z_S(\beta,\Gamma(1),1)\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_{-1}})\,,\\
Z_S(\beta,\Gamma_0(2),1) &= Z_S(\beta,\Gamma(1),1)\,\det(1-\tilde{\mathcal{L}}_\beta^{\chi_2})\,.
\end{aligned}$$

6 Conclusion

In this paper we have discussed the transfer operator for the dynamical system of a free particle moving on the modular surfaces $M_\Gamma=\Gamma\backslash\mathbb{H}$ belonging to the subgroups $\Gamma\subseteq\Gamma(1)$ of finite index $\mu=[\Gamma(1):\Gamma]$, which are μ-fold covering surfaces of the modular surface $M_{\Gamma(1)}=\Gamma(1)\backslash\mathbb{H}$. Topologically these surfaces are spheres with a finite number of handles with cusps at infinity. Starting from the well understood case of the modular surface $\Gamma(1)\backslash\mathbb{H}$ we construct Poincaré surfaces X_Γ and the Poincaré return maps P_Γ from which the transfer operators $\tilde{\mathcal{L}}_\beta^\Gamma$ for Γ can be determined in a standard way. We show that the operators $\tilde{\mathcal{L}}_\beta^\Gamma$ are meromorphic in $\beta\in\mathbb{C}$, determine their spectral properties as nuclear operators and show how they can be decomposed into their basic components. By using trace formulas similar to the Atiyah-Bott formula the dynamical zeta functions of Ruelle and Smale for the geodesic flows on the surfaces of constant negative curvature given by Γ can be expressed in the thermodynamic formalism through Fredholm determinants of the transfer operator. The aforementioned decomposition of the transfer operator leads to a factorization of the Selberg zeta function for Γ and gives a new dynamical interpretation of the Venkov-Zograf factorization proved originally within the classical approach by the Selberg trace formula. Finally we discuss some simple groups where the approach developed in this paper can be performed explicitly.

It would be interesting to compare the transfer operators we have constructed for subgroups Γ of $PSL(2,\mathbb{Z})$ with the operators Morita defined in his recent work on general cofinite Fuchsian groups in [Mor97]. In a second paper we will discuss the spectral properties of the transfer operators in more detail especially their eigenfunctions to the eigenvalue $\lambda = 1$ and their relation to automorphic forms of the groups Γ.

References

[AL70] A. O. L. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[Art24] E. Artin. Ein mechanisches System mit quasi-ergodischen Bahnen. *Abh. Math. Sem. d. Hamburgischen Universität*, 3:170–175, 1924.

[BO95] U. Bunke and M. Olbrich. *Selberg Zeta and Theta Functions.* Akademie Verlag, 1995.

[Bow75] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Springer-Verlag, 1975. L.N.in Mathematics 470.

[Cha99] C.-H. Chang. *Die Transferoperator-Methode für Quantenchaos auf den Modulflächen $\Gamma\backslash\mathbb{H}$.* Papierflieger, Clausthal-Zellerfeld, January 1999. ISBN 3-89720-253-0.

[CM98] C.-H. Chang and D. Mayer. The period function of the nonholomorphic Eisenstein series for $PSL(2,\mathbb{Z})$. *Math. Phys. Elec. J.*, 4(6), 1998.

[CM99] C.-H. Chang and D. Mayer. The transfer operator approach to Selberg's zeta function and modular and Maass wave forms for $PSL(2,\mathbb{Z})$. In D. Hejhal and M. Gutzwiller et al, editors, *IMA Volumes 109 'Emerging applications of number theory'*, pages 72–142. Springer-Verlag, 1999.

[Ein17] A. Einstein. Zum Quantensatz von Sommerfeld und Epstein. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 19:82–92, 1917.

[Fri86] D. Fried. The zeta functions of Ruelle and Selberg i. *Ann. scient. éc. norm. sup.* 4^e *Série*, 19:491–517, 1986.

[Gro55] A. Grothendieck. Produits tensoriels topologiques et espaces nucléaires. *Mem. Am. Math. Soc.*, 16, 1955.

[Gro56] A. Grothendieck. La theorie de Fredholm. *Bull. Soc. Math. France*, 84:319–384, 1956.

[Gut90] M. Gutzwiller. *Chaos in classical and quantum mechanics.* Springer-Verlag, 1990.

[Hej76] D. Hejhal. The Selberg trace formula and the Riemann zeta function. *Duke Math. J.*, 43(3):441–482, Sep. 1976.

[Isi25] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31:253–258, 1925.

[KF66] F. Klein and R. Fricke. *Vorlesungen über die Theorie der elliptischen Modulfunktionen. Ausgearbeitet und vervollständigt von Robert Fricke. 2 Vol.* Teubner, Leipzig, 1966.

[Laz93] V. Lazutkin. *KAM Theory and Semiclassical Approximations to Eigenfunctions.* Springer-Verlag, 1993.

[LZ97] J. Lewis and D. Zagier. Period functions and the Selberg zeta function for the modular group. In *The Mathematical Beauty of Physics*, Adv. Series in Math. Physics 24, pages 83–97. World Scientific, Singapore, 1997.

[May80] D. Mayer. *The Ruelle-Araki transfer operator in classical statistical mechanics.* L. N. in Phys. 123. Springer-Verlag, Berlin, 1980.

[May91a] D. Mayer. Continued fractions and related transformations. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, chapter 7, pages 175–222. Oxford Univ. Press, Oxford, 1991.

[May91b] D. Mayer. The thermodynamic formalism approach to Selberg's zeta function for $PSL(2,\mathbb{Z})$. *Bull. Am. Math. Soc.*, 25:55–60, 1991.

[Mor97] T. Morita. Markov systems and transfer operators associated with cofinite Fuchsian groups. *Ergodic Theory Dyn. Syst.*, 17(5):1147–1181, 1997.

[MOS66] W. Magnus, F. Oberhettinger, and R. Soni. *Formulas and Theorems for the Special Functions of Mathematical Physics.* Springer-Verlag, 1966.

[Pol86] M. Pollicott. Distribution of closed geodesics on the modular surface and quadratic irrationals. *Bull. Soc. math. France*, 114:431–446, 1986.

[Pol91] M. Pollicott. Closed geodesics and zeta function. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, chapter 7, pages 153–173. Oxford Univ. Press, Oxford, 1991.

[Ran77] R. Rankin. *Modular Forms and Functions.* Cambridge University Press, 1977.

[Rue76] D. Ruelle. Zeta-functions for expanding maps and Anosov flows. *Invent. Math.*, 34:231–242, 1976.

[Rue78] D. Ruelle. *Thermodynamic formalism.* Addison Wesley, Reading MA, 1978.

[Rue94] D. Ruelle. *Dynamical Zeta Functions for Piecewise Monotone Maps of the Interval.* AMS, 1994.

[Sar95] P. Sarnak. Arithmetic chaos. In *Israel Math. Conf. Proc.*, volume 8, pages 183–236, 1995.

[Sch74] B. Schoeneberg. *Elliptic Modular Functions.* Springer-Verlag, 1974.

[Sel56] A. Selberg. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Indian Math. Soc.*, 20:47–87, 1956.

[Ser85] C. Series. The modular surface and continued fractions. *J. London Math. Soc.*, (2), 31:69–80, 1985.

[Ter85] A. Terras. *Harmonic Analysis on Symmetric Spaces and Applications I.* Springer-Verlag, 1985.

[Ven90] A.B. Venkov. *Spectral Theory of Automorphic Functions and Its Applications.* Kluwer Academic Publishers, 1990.

[Vin89] E. Vinberg. *Linear Representation of Groups.* Birkhäuser, Basel, 1989.

[VZ83] A. B. Venkov and P. G. Zograf. On analogues of the Artin factorization formulas in the spectral theory of automorphic functions connected with induced representations of Fuchsian groups. *Math. USSR Izvestiya*, 21(3):435–443, 1983.