BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 25, Number 1, July 1991

THE THERMODYNAMIC FORMALISM APPROACH TO SELBERG'S ZETA FUNCTION FOR PSL(2, Z)

DIETER H. MAYER

I. INTRODUCTION

Besides the classical approach to Selberg's zeta function for cofinite Fuchsian groups [S] through the trace formula [V] there has been developed recently another one based on the thermodynamic formalism [R2] applied to the dynamical zeta function of Smale and Ruelle [F] which for geodesic flows on surfaces of constant negative curvature (c.n.c.) is closely related to Selberg's function for the corresponding Fuchsian group [Sm, R1]. This latter approach however has been worked out up to now only for cocompact groups.

In this announcement we discuss the first example of a cofinite, noncocompact Fuchsian group where the aforementioned thermodynamic formalism approach works also, namely the modular group $\mathrm{PSL}(2, \mathbf{Z})$. The most remarkable fact with this group is that the whole formalism can be made rather explicit contrary to the general case where many of the constructions used are rather difficult to come by. The reason for this is a quite simple construction of symbolic dynamics for geodesic flows on surfaces of c.n.c. due to Bowen and Series [BS]. Instead of an usually only inductively defined Markov partition [F] their symbolic dynamics is based on a piecewise analytic Markov map of the limit set of the Fuchsian group, determined by the group generators. Through this symbolic dynamics the Smale-Ruelle function for the flow gets transformed into a generating function for partition functions for the B-S map to which the transfer operator method of statistical mechanics applies [M1, R1]. Since for cocompact groups the B-S maps are expanding [BS] their transfer operators can be chosen as nuclear operators [G], and the Selberg function finally gets

Received by the editors May 14, 1990 and, in revised form, November, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 58F20, 58F25; Secondary 11F72, 11M26.

expressed in terms of Fredholm determinants of these operators [Po2].

In trying to extend this approach to general cofinite groups one faces two problems: the B-S map is not expanding for a noncocompact group and hence its transfer operators are not nuclear [M3]. Furthermore, in this case this map has infinitely many branches leading to a more involved discussion of the analytic dependence of its transfer operators on possible parameters. For the group $PSL(2, \mathbf{Z})$ both these problems can be resolved. To achieve this we use the remarkable results of Series [Se], respectively Adler and Flatto [AF], showing that for the modular surface the nonexpanding B-S map can indeed be replaced by an expanding map, induced from the former one on a certain subset of the limit set. Quite surprisingly, this new map turns out to be the classical continued fraction map $T_G x = x^{-1} \bmod 1$ on the unit interval, whose importance for the modular surface was recognized already by Artin in [A]. This map still has infinitely many monotone branches so that the analytic properties of its transfer operators in exterior parameters are more involved. They have been worked out only recently in [M2].

The thermodynamic formalism then leads to a rather explicit representation of the Smale-Ruelle function and hence also of the Selberg function for $PSL(2, \mathbf{Z})$ in terms of Fredholm determinants of transfer operators of the map T_G. Finally, combining our results with classical ones for the Selberg function derived from the trace formula suggests also a seemingly new formulation of Riemann's hypothesis on his zeta function in terms of the transfer operators of T_G.

II. Transfer operators and Ruelle's zeta function for the Gauss map

The thermodynamic formalism for the Gauss map $T_G x = x^{-1} \bmod 1$ on the unit interval has been discussed recently in [M2]. In this formalism, a central role is played by the partition functions $Z_n(T_G, A)$ defined for $n \in \mathbf{N}$ through the n-periodic points $x \in \operatorname{Fix} T_G^n$ of T_G by the formula

$$Z_n(T_G, A) = \sum_{x \in \operatorname{Fix} T_G^n} \exp \sum_{k=0}^{n-1} A(T_G^k x),$$

where $A = A_s(x) = -s \log |T_G'(x)| = s \log x^2$ with s a complex parameter known in physics as "inverse temperature." Obviously $|\operatorname{Fix} T_G^n| = \infty$ for all n since the n-periodic points of T_G are just all irrationals x in the unit interval with n-periodic continued fraction expansion. Hence the partition functions $Z_n(T_G, A_s)$ are defined only for $\operatorname{Re} s > \frac{1}{2}$. The transfer matrix was invented in statistical mechanics to calculate partition functions similar to the ones introduced above [M1], by transforming the combinatorial problem into an algebraic one. Under the name transfer operator it became a standard tool in the ergodic theory of dynamical systems [R2]. To apply this technique here consider the following operators $L_s : A_\infty(D) \to A_\infty(D)$ acting on the B-space of functions holomorphic on the disc $D = \{z : |z-1| < \frac{3}{2}\}$ and continuous on $\overline{D}$:

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

The infinite sum reflects the infinite monotone branches of the map T_G with inverses $(n+x)^{-1}$, so that also L_s is well defined only for complex values of s with $\operatorname{Re} s > \frac{1}{2}$. Using essentially only Grothendieck's theory of nuclear operators [G] one shows [M2]

Proposition 1. *For* $\operatorname{Re} s > \frac{1}{2}$ *the operators* L_s *are nuclear of order zero and fulfill the trace formulas*

$$Z_n(T_G, A_s) = \operatorname{trace} L_s^n - \operatorname{trace}(-L_{s+1})^n.$$

Consider next for $k \in \mathbf{N}$ the functions

$$\zeta_k(z, s) = \exp \sum_{n=1}^{\infty} \frac{z^n}{n} Z_{nk}(T_G, A_s)$$

introduced originally for $k = 1$ in [R1] and well defined for $|z| < \exp -(kP(s))$ with $P(s) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(T_G, A_s)$ the topological pressure of A_s. We get, applying Proposition 1 and Grothendieck's Fredholm theory [G],

Corollary 1. *The functions* $\zeta_k(z, s)$ *can be expressed for* $\operatorname{Re} s > \frac{1}{2}$ *as* $\zeta_k(z, s) = \det(1 - z(-L_{s+1})^k)/\det(1 - zL_s^k)$ *and extend for such* s *as meromorphic functions into the entire* z*-plane. The functions* $\zeta_k(s) := \zeta_k(1, s)$ *are meromorphic in the half plane* $\operatorname{Re} s > \frac{1}{2}$ *for all* $k \in \mathbf{N}$.

That the functions $\zeta_k(s)$ of Corollary 1 extend as meromorphic functions even into the entire s-plane follows from the next result proved in [M2].

Theorem 1. *The map $s \to L_s$ extends as a meromorphic function into the entire s-plane with values nuclear operators of order zero in $A_\infty(D)$. It has simple poles for $s = s_k = (1-k)/2$, $k = 0, 1, \ldots$ with residues the rank 1 operators $N_k f(z) = \frac{1}{2}\frac{1}{k!}f^{(k)}(0)$ in $A_\infty(D)$. The Fredholm determinants* $\det(1 \pm L_s^n)$ *extend as meromorphic functions into the entire s-plane with (possibly removable) singularities at the above s_k values.*

III. Selberg's zeta function for PSL(2, Z)

The Selberg function $Z(s)$ for a general cofinite Fuchsian group Γ [S] has a simple interpretation in terms of the dynamical Smale-Ruelle zeta function $\zeta_{SR}(s)$ for the geodesic flow ϕ_t on the surface of c.n.c. defined by Γ [Sm, R1]:

$$(2) \qquad Z(s) = \prod_{k=0}^{\infty} \zeta_{SR}(s+k)^{-1} = \prod_{\gamma}\prod_{k=0}^{\infty}(1 - e^{-(s+k)l(\gamma)}),$$

where the product is over the length spectrum $L(\phi_t)$ of ϕ_t consisting of all periodic orbits γ with prime period $l(\gamma)$. The products are known to converge for $\mathrm{Re}\, s > 1$ [R1]. To discuss this function for the group $\Gamma = \mathrm{PSL}(2, \mathbf{Z})$ we use the results in [Se], respectively [AF], on the symbolic dynamics of the geodesic flow ϕ_t on the modular surface. It was shown there that ϕ_t can be described by a special flow built over a natural extension $\widetilde{T}_G$ of the Gauss map with $\widetilde{T}_G(x, y, \epsilon) = (T_G x, (y + [\frac{1}{x}])^{-1}, -\epsilon)$, $\epsilon = \pm 1$, (x, y) in the unit square, from which the length spectrum $L(\phi_t)$ can be derived as [Po1]

(3)
$$L(\phi_t) = \left\{ -\sum_{k=0}^{2r-1} \overline{A}_1(\widetilde{T}_G^{\,k}([w])) : \quad [w] \text{ periodic orbit of } \widetilde{T}_G \text{ of prime period } 2r,\ r \in \mathbf{N} \right\}$$

with $\overline{A}_1$ the special case $s = 1$ of the function $\overline{A}_s(w) = A_s(x)$ for $w = (x, y, \epsilon)$. By a standard chain of arguments [P] the function $\zeta_{SR}(s)$ can be written as

$$\zeta_{SR}(s) = \exp \sum_{n=1}^{\infty} \frac{z^n}{n} Z_n(\widetilde{T}_G, \overline{A}_s)$$

with $Z_n(\widetilde{T}_G, \overline{A}_s)$ the partition functions for $\widetilde{T}_G$ and $\overline{A}_s$. A simple calculation however shows that $Z_n(\widetilde{T}_G, \overline{A}_s) = 0$ for n odd and $Z_n(\widetilde{T}_G, \overline{A}_s) = 2Z_n(T_G, A_s)$ for n even. Hence the Smale-Ruelle zeta function $\zeta_{SR}(s)$ for ϕ_t is just the function $\zeta_2(s)$ as defined for T_G in Corollary 1. Combining formulas 2 and 3 then proves our main result:

Theorem 2. *The Selberg zeta function* $Z(s)$ *for the modular group* PSL(2, **Z**) *can be written as* $Z(s) = \det(1-L_s)\det(1+L_s)$ *with* L_s *the transfer operator of the Gauss map* T_G. $Z(s)$ *is meromorphic in the entire s-plane with (possibly removable) singularities at the points* $s_k = (1-k)/2$, $k = 0, 1, 2, \ldots$.

The nontrivial zeros of $Z(s)$ are hence given by those s-values for which L_s has $\lambda = 1$ or $\lambda = -1$ as an eigenvalue. The trace formula approach shows [V] that these zeros are either $\frac{1}{2}$ times the nontrivial zeros of Riemann's zeta function or they determine via the formula $s = \frac{1}{2} + ir$ the eigenvalues $\rho = \frac{1}{4} + r^2$ of the Laplacian $-\Delta$ on the modular surface. The value $s = 1$ is such a zero since $\det(1 - L_1) = 0$, corresponding to the eigenvalue $\rho = 0$ of $-\Delta$. One can then ask if it is generally true that all the eigenvalues of $-\Delta$ are determined by the factor $\det(1-L_s)$ whereas the nontrivial zeros of Riemann's zeta function are determined by $\det(1 + L_s)$. This would mean that the factorization in Theorem 2 corresponds to the one found for cocompact Fuchsian groups in [Sa]. In this case the Riemann Hypothesis would be equivalent to the operator L_s having eigenvalue $\lambda = -1$ only for s-values on the line $\operatorname{Re} s = \frac{1}{4}$.

REFERENCES

[AF] R. Adler and L. Flatto, *Cross section maps for the geodesic flow on the modular surface*, Contemporary Math., vol. 26, Amer. Math. Soc., Providence, RI, 1984, pp. 9–24.

[A] E. Artin, *Ein mechanisches System mit quasi-ergodischen Bahnen*, Collected Papers, Addison-Wesley, Reading, MA, 1965, pp. 499–504.

[BS] R. Bowen and C. Series, *Markov maps associated to Fuchsian groups*, Inst. Hautes Études Sci. Publ. Math. **50** (1979), 153–170.

[F] D. Fried, *The zeta functions of Ruelle and Selberg*. I, Ann. Sci. École Norm. Sup. (4) **19** (1986), 491–517.

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

[M1] D. Mayer, *The Ruelle-Araki transfer operator in classical statistical mechanics*, Lecture Notes in Phys., vol. 123, Springer-Verlag, Berlin and New York, 1980.

[M2] ——, *On the thermodynamic formalism for the Gauss map*, Comm. Math. Phys. **130** (1990), 311–333.

[M3] ——, *Continued fractions and related transformations*, Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces, (T. Bedford, M. Keane, and C. Series, ed.), Oxford Univ. Press, Oxford, 1991.

[P] W. Parry, *An analogue of the prime number theorem for closed orbits of shifts of finite type and their suspensions*, Israel J. Math. **45** (1983), 41–52.

[Po1] M. Pollicott, *Distribution of closed geodesics on the modular surface and quadratic irrationals*, Bull. Soc. Math. France **114** (1986), 431–446.

[Po2] ——, *Some applications of the thermodynamic formalism to manifolds of constant negative curvature*, Adv. Math. (to appear).

[R1] D. Ruelle, *Zeta functions and statistical mechanics*, Asterisque **40** (1976), 167–176.

[R2] ——, *Thermodynamic formalism*, Addison-Wesley, Reading, MA, 1978.

[Sa] P. Sarnak, *Determinants of Laplacians*, Comm. Math. Phys. **110** (1987), 113–120.

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

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

[Sm] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.

[V] A. B. Venkov, *Spectral theory of automorphic functions, the Selberg zeta function, and some problems of analytic number theory and mathematical physics*, Russian Math. Surveys (3) **34** (1979), 79–153.

MAX-PLANCK-INSTITUT FÜR MATHEMATIK, GOTTFRIED-CLAREN-STRASSE 26, D-5300 BONN, FEDERAL REPUBLIC OF GERMANY