

SIMPLE CLOSED GEODESICS ON $H^+/\Gamma(3)$ ARISE FROM THE MARKOV SPECTRUM

BY J. LEHNER AND M. SHEINGORN

1. Let

$$H^+ = \{z = x + iy: y > 0\}$$

be the complex upper half-plane, and let

$$\Gamma(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{n}; a, b, c, d \in \mathbf{Z}, ad - bc = 1 \right\}$$

be the principal congruence subgroup of level n in the modular group $SL(2, \mathbf{Z}) = \Gamma(1)$. In this note we are concerned with $\Gamma(3)$. Let S be the Riemann surface $H^+/\Gamma(3)$ and let $\pi: H^+ \rightarrow S$ be the projection map. S is a sphere with four punctures.

A hyperbolic element γ is a Möbius transformation of H^+ that has two real fixed points; its axis A_γ is the circle with center on \mathbf{R} connecting the fixed points. Write ξ_γ, ξ'_γ for the fixed points of γ . If $\gamma \in \Gamma(3)$ is hyperbolic, A_γ projects to a closed geodesic on S ; conversely, every closed geodesic on S arises in this way. A *simple* closed geodesic is one that does not intersect itself.

The Markov Spectrum will be described in detail in §2. Here we note the definition of the Markov function $M(\theta)$. For real irrational θ set

$$(1.1) \quad M(\theta) = \sup\{c > 0: |\theta - p/q| < 1/cq^2 \text{ for infinitely many reduced fractions } p/q\}.$$

In the range $M(\theta) < 3$, M assumes only a denumerably infinite set of values $M_\nu \uparrow 3$. The numbers M_ν constitute the Markov Spectrum, which we denote by MS.

The connection between simple closed geodesics on S and MS is established in the following way. For $\beta \in \Gamma(3)$ write $A_\gamma \wedge \beta A_\gamma$ to mean $A_\gamma \cap \beta A_\gamma \neq \emptyset$, A_γ , i.e., the intersection is a single point in H^+ . The following criterion is easy to prove:

$$(1.2) \quad \pi(A_\gamma) \text{ is nonsimple if and only if } A_\gamma \wedge \beta A_\gamma \text{ for some } \beta \in \Gamma(3) - \langle \gamma \rangle.$$

But in this statement we know nothing about β except that it is not elliptic ($\Gamma(3)$ contains no elliptic elements).

THEOREM 1. *If $\pi(A_\gamma)$ is nonsimple, there is a parabolic element P in $\Gamma(3)$ such that $A_\gamma \wedge PA_\gamma$.*

Theorem 1 leads directly to the main result:

Received by the editors September 26, 1983 and, in revised form, May 1, 1984.
 1980 *Mathematics Subject Classification.* Primary 10D05.

©1984 American Mathematical Society
 0273-0979/84 \$1.00 + \$.25 per page

THEOREM 2. *Let $\gamma \in \Gamma(3)$ be hyperbolic. Then $\pi(A_\gamma)$ is simple if and only if $M(\xi_\gamma) = M(\xi'_\gamma) < 3$.*

Since Zagier [3] has recently given an asymptotic formula describing this portion of MS, we can deduce from Theorem 2:

COROLLARY 1. *Let $N_S(T)$ be the number of simple closed geodesics on S of hyperbolic length $\leq T$. Then $T \ll N_S(T) \ll T^2$.*

The implied constants are effective. This inequality contrasts with results on $N(T)$, the number of closed geodesics of length $\leq T$, first obtained by H. Huber [1].

Returning to Theorem 1, we make use in the proof of the following known facts:

(1.3) A simple loop L contained in $\pi(A_\gamma)$ cannot bound a disk, i.e., each component of $S - L$ must contain at least one puncture.

(1.4) If L bounds a disk with exactly one puncture, then L determines a conjugacy class of parabolic elements of $\Gamma(3)$.

We remark that $\pi(A_\gamma)$ has a finite number of self-intersections, since it is a real-analytic curve.

Now it can be shown that $\pi(A_\gamma)$, assumed nonsimple, contains a simple loop L surrounding a single puncture p . We are indebted to A. F. Beardon for a proof of this fact that is shorter and simpler than the original one.

There is a lift of $\pi(A_\gamma)$ lying on A_γ and starting from a point ζ , i.e., the lift is an interval $(\zeta, \gamma\zeta)$ of A_γ . Using (1.4) one can show that there is a parabolic element P in $\Gamma(3)$ such that the lift of L is an interval $(z_0, Pz_0) \subset (\zeta, \gamma\zeta)$. Thus $A_\gamma \wedge PA_\gamma$, as asserted.

Full details will follow in a paper written jointly with A. F. Beardon. This paper also contains the following result. Let Γ be a finitely generated fuchsian group and let $S = H^+/\Gamma$ be the associated Riemann surface. Then Theorem 1 holds for S if and only if S is of genus zero and has either three or four punctures or deleted disks.

2. We now return to the Markov Spectrum (MS); for a fuller account see [2, pp. 29–32]. In (1.1) and the following lines we defined MS to be the set of values $\{M_\nu\}$ assumed by the Markov function $M(\theta)$ in the range $M(\theta) < 3$. In order to calculate M_ν we introduce *Markov triples*. A triple of positive integers (x, y, z) is called a Markov triple if $x^2 + y^2 + z^2 = 3xyz$, $1 \leq x \leq y \leq z$. The first triples are $(1, 1, 1), (1, 1, 2), (1, 2, 5), \dots$, and the rest can be recursively generated. Order the triples by the size of z so that $1 = z_1 \leq 2 = z_2 \leq \dots \leq z_\nu \dots$. With each triple (x_ν, y_ν, z_ν) there is associated a pair of real quadratic conjugates

$$(2.1) \quad \theta_\nu, \theta'_\nu = \frac{1}{2} + y_\nu/x_\nu z_\nu \pm \frac{1}{2}(9 - 4/z_\nu^2)^{1/2}, \quad \nu \geq 1.$$

The connection of $M(\theta)$ with θ_ν is that

$$(2.2) \quad M(\theta_\nu) = M_\nu = |\theta_\nu - \theta'_\nu| = (9 - 4/z_\nu^2)^{1/2}.$$

We have $M_1 = 5^{1/2}, M_2 = 8^{1/2}, M_3 = (221)^{1/2}/5, \dots, \rightarrow 3$.

Next, introduce the equivalence relation:

$$\theta \sim \psi \text{ if and only if } \psi = (a\theta + b)/(c\theta + d) \text{ with integers } a, b, c, d \quad (2.3)$$

and $ad - bc = \pm 1$.

Then $\theta \sim \psi$ if and only if

$$(2.4) \quad M(\theta) = M(\psi).$$

Moreover, the regular continued fraction expansions of θ and ψ agree from a certain point on. Also

$$(2.5) \quad M(\theta) < 3 \Rightarrow \theta \sim \theta_\nu \text{ for some } \nu \geq 1.$$

Indeed, the definition of MS shows that $M(\theta) = M_\nu = M(\theta_\nu)$, so $\theta \sim \theta_\nu$ by (2.4).

The numbers $\{\theta_\nu\}$, $\{\theta'_\nu\}$, together with their equivalents under (2.3), are called Markov quadratic irrationalities (MQI). Theorem 2 may now be re-stated.

THEOREM 2'. $\pi(A_\gamma)$ is simple if and only if ξ_γ is equivalent to a MQI.

We can associate MS to hyperbolic elements of $\Gamma(3)$. For each ν there is a $\gamma_\nu \in \Gamma(3)$ whose fixed points are $\xi_{\gamma_\nu} = \theta_\nu, \xi'_{\gamma_\nu} = \theta'_\nu$. Namely, dropping the subscript ν , let $\zeta = 1$ if z is odd, otherwise $\zeta = 1/2$. Define

$$(2.6) \quad B = \begin{pmatrix} (N + x(2y + xz)\zeta M)2^{-1} & (2x^2z - 4xy + z)\zeta M \\ x^2z\zeta M & (N - x(2y + xz)\zeta M)2^{-1} \end{pmatrix},$$

where $M > 0$ is the smallest integral solution of the Pell equation

$$x^4(9z^2 - 4)\zeta^2 M^2 + 4 = N^2.$$

Then it can be shown that B is the $\Gamma(1)$ -primitive matrix fixing ξ, ξ' . Moreover, $B \in \Gamma(3)$ if $3|M$, otherwise $B^2 \in \Gamma(3)$. But the first case never occurs, so B^2 is the $\Gamma(3)$ -primitive matrix fixing ξ, ξ' .

By abuse of notation we say $\gamma \in \text{MS}$ if $\gamma \in \Gamma(3)$ and $\xi_\gamma \sim \theta_\nu$ for some $\nu \geq 1$. If $\gamma \in \text{MS}$ so does $V\gamma V^{-1}, V \in \Gamma(1)$, since $V\gamma V^{-1} \in \Gamma(3)$ by normality of $\Gamma(3)$ in $\Gamma(1)$ and $\xi_{V\gamma V^{-1}} = V\xi_\gamma \sim V\theta_\nu \sim \theta_\nu$. That is,

(2.7) the conjugacy class of γ in $\Gamma(1)$ belongs to MS if $\gamma \in \text{MS}$.

We now prove Theorem 2. Suppose $\pi(A_\gamma)$ is nonsimple; then by Theorem 1 there is a δ conjugate to γ in $\Gamma(3)$ for which $A_\delta \wedge S^3 A_\delta$, i.e., $|\xi_\delta - \xi_{\delta'}| > 3$. By a translation in $\Gamma(1)$ we may assume $-1 < \xi'_\delta < 0$; then $\xi_\delta > \xi'_\delta + 3 > 1$. Thus ξ_δ is "reduced" [4, p. 73] and the regular continued fraction of ξ_δ is pure periodic; also ξ'_δ . Let $\xi_\delta = (\overline{b_0, b_1, \dots, b_{k-1}})$ for $k \geq 1$; then $-1/\xi'_\delta = (\overline{b_{k-1}, \dots, b_0})$ [4, p. 76]. Here $b_{n k + \nu} = b_\nu$ for $0 \leq \nu < k, n \geq 0$. Set

$$m_\mu = (\overline{b_\mu, b_{\mu+1}, \dots, b_{\mu+k-1}}) + (0, \overline{b_{\mu-1}, b_{\mu-2}, \dots, b_{\mu-k}}), \quad \mu \geq k.$$

By periodicity $m_\mu = m_{\mu+k}$. Moreover, $M(\xi_\delta) = \overline{\lim}_{\mu \rightarrow \infty} m_\mu$ [2, p. 29].

Therefore, for all $\varepsilon > 0$ and $n > N$,

$$3 < \xi_\delta - \xi'_\delta = m_k = m_{nk} < \overline{\lim}_{\mu \rightarrow \infty} m_\mu + \varepsilon < M(\xi_\delta) + \varepsilon,$$

implying

$$M(\xi_\gamma) = M(\xi_\delta) > 3,$$

as asserted.

Conversely, assume $\pi(A_\gamma)$ is simple. Then certainly $|\xi_\gamma - \xi'_\gamma| \leq 3$, otherwise $A_\gamma \wedge S^3 A_\gamma$. Since $\pi(A_\gamma)$ is simple if and only if $\pi(VA_\gamma) = \pi(A_{V\gamma V^{-1}})$ is simple for all $V \in \Gamma(1)$ —because $\Gamma(3) \triangleleft \Gamma(1)$ —we have

$$(*) \quad |V\xi_\gamma - V\xi'_\gamma| \leq 3, \quad V \in \Gamma(1).$$

Assuming $\gamma \notin \text{MS}$ we shall produce a $V \in \Gamma(1)$ that contradicts (*).

At this point we observe that $M(\xi_\gamma) \neq 3$ for any $\gamma \in \Gamma(1)$. Indeed, ξ_γ is a quadratic irrationality and $M(\theta)$ is never 3 if θ is a quadratic irrational [2, p. 32]. It follows that $\gamma \notin \text{MS}$ implies $M(\xi_\gamma) > 3$, that is,

$$|\xi_\gamma - p_n/q_n| < 1/(3+h)q_n^2, \quad (p_n, q_n) = 1,$$

for some $h > 0$, on a sequence $q_n \rightarrow \infty$. Write $V_n = (q'_n, -p'_n : q_n, -p_n) \in \Gamma(1)$. Then with $\xi_\gamma = \xi$, $\xi'_\gamma = \xi'$,

$$\begin{aligned} |V_n \xi - V_n \xi'| &= \frac{|\xi - \xi'|}{q_n^2 |\xi - p_n/q_n| |\xi' - p_n/q_n|} > \frac{(3+h)|\xi - \xi'|}{|\xi' - p_n/q_n|} \\ &\geq \frac{(3+h)|\xi - \xi'|}{|\xi' - \xi| + |\xi - p_n/q_n|} > \frac{3+h}{1 + 1/3q_n^2 |\xi - \xi'|} > 3, \end{aligned}$$

for $n \geq n_0$. For $V = V_{n_0}$ we have a contradiction to (*).

We close with a comment on Corollary 1. The existence of long simple geodesics on $H^+/\Gamma(3)$ is not hard to prove topologically. The feature of Corollary 1 is that the lengths are known explicitly: they are

$$\text{length } A_{B_\nu^2} = 2 \log \frac{t_\nu + \sqrt{t_\nu^2 - 4}}{2}, \quad t_\nu = \text{trace } B_\nu^2.$$

REFERENCES

1. H. Huber, *Zur analytischen Theorie hyperbolischen Raumformen und Bewegungsgruppen*, Math. Ann. **138** (1959), 1–26.
2. J. F. Koksma, *Diophantische Approximationen*, Chelsea.
3. D. Zagier, *On the number of Markov numbers below a given bound*, Math. Comp. **39** (1982), 709–723.
4. O. Perron, *Die Lehre von den Kettenbrüchen*, Vol. 1, Teubner, Leipzig, 1954.

SCHOOL OF MATHEMATICS, THE INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY 08540

DEPARTMENT OF MATHEMATICS, BARUCH COLLEGE, CITY UNIVERSITY OF NEW YORK, NEW YORK, NEW YORK 10010