

The Correlation of Length Spectra of Two Hyperbolic Surfaces

Richard Schwartz¹ and Richard Sharp²

¹ M.S.R.I., 1000 Centennial Drive, Berkeley, CA 94720, USA

² School of Mathematical Sciences, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, United Kingdom

Received July 2, 1992; in revised form November 10, 1992

Abstract. We derive an asymptotic formula for the number of free homotopy classes on a closed surface which have (approximately) the same length with respect to two different hyperbolic structures on the surface. The growth rate in the asymptotic formula is described in terms of the thermodynamic formalism for the geodesic flow.

1. Introduction

Let S be a smooth closed surface of genus ≥ 2 . A hyperbolic structure on S is an equivalence class of pairs (X, ϕ) such that X is a surface of constant negative curvature -1 and $\phi: S \rightarrow X$ is a diffeomorphism. Two such pairs (X, ϕ) and (X_1, ϕ_1) are equivalent if there is an isometry $\tau: X \rightarrow X_1$ such that $\tau \circ \phi$ is isotopic to ϕ_1 .

The map ϕ induces a one to one correspondence between free homotopy classes on S and closed (but not necessarily simple) geodesics on X . The closed geodesic $g \in X$ corresponds to the free homotopy class $[\gamma]$ represented by the closed curve $\phi^{-1}(g) \in S$. The structure (X, ϕ) induces a *length function* l on free homotopy classes of S . We define $l[\gamma]$ to be the length of g on X . This length function is independent of the choice of pair (X, ϕ) representing the hyperbolic structure.

The image of l is called the length spectrum, and it is natural to ask how it is distributed in \mathbf{R}^+ . One answer is the following asymptotic result [H]: For fixed “tolerance” $\varepsilon > 0$,

$$\#\{[\gamma]: l[\gamma] \in (x, x + \varepsilon)\} \sim (e^\varepsilon - 1) \frac{e^x}{\varepsilon}.$$

Here the notation $f(x) \sim g(x)$ means that $f(x)/g(x) \rightarrow 1$ as $x \rightarrow \infty$.

In this paper, we will consider the correlation between the spectra of two length functions l_1 and l_2 induced by the two hyperbolic structures (X_1, ϕ_1) and (X_2, ϕ_2) . Fixing a tolerance $\varepsilon > 0$, we ask if it is possible to find a free homotopy class $[\gamma]$ of S such that $|l_1[\gamma] - l_2[\gamma]| < \varepsilon$. Moreover, we ask for the asymptotic number of

such free homotopy classes as a function of the (common) length. These questions are partially answered by the following.

Theorem 1. *Let l_1 and l_2 be two length functions associated to two distinct hyperbolic structures on S , and let $\varepsilon > 0$ be fixed. Then there are constants $C = C(\varepsilon, l_1, l_2) > 0$ and $M = M(l_1, l_2) \in (0, 1)$ such that*

$$\#\{\gamma\} : l_i[\gamma] \in (x, x + \varepsilon), i = 1, 2\} \sim C \frac{e^{Mx}}{x^{3/2}}.$$

The proof of Theorem 1 contains a precise description of the exponential growth rate $M(l_1, l_2)$ in terms of familiar dynamical quantities associated to the geodesic flows on hyperbolic surfaces. However it seems very difficult to write down an explicit formula for the function $M(*, *)$ in terms of the two hyperbolic structures on which it depends. The only information we have about M is that it is continuous in its arguments, and $M(l_1, l_2) \rightarrow 1$ as the two structures (X_1, ϕ_1) and (X_2, ϕ_2) converge to a single structure. Similarly, for fixed ε , $C(\varepsilon, l_1, l_2) \rightarrow \infty$ as the two structures converge to a single structure. A question which we would like to have answered, but couldn't, was whether or not M is a bounded away from 0 as its arguments range over all hyperbolic structures on S .

The proof of Theorem 1 falls naturally into two parts. In the first part, Sect. 2, we describe the geodesic flow, and prove a technical "independence lemma" concerning length functions. In Sect. 3, we use this independence result to show how our Theorem 1 may be deduced from a theorem of S. Lalley concerning the distribution of orbits of the geodesic flow on a hyperbolic surface.

We would like to thank Mark Pollicott, who suggested to us that Lalley's theorem might be useful in studying the length spectra of two surfaces simultaneously. We would also like to thank the I.H.E.S. for providing support during the preparation of this paper.

2. The Geodesic Flow

2.1. The Geodesic Flow. Let X be a fixed closed surface of constant curvature -1 . The unit tangent bundle of X is the manifold of pairs (x, v) , where $x \in X$ and v is a unit tangent vector to X at x . We will denote the unit tangent bundle either by $UT(X)$, or by UT , when the dependence on X is clear. Locally, UT is the product of a disk with a circle. UT has a natural Riemannian metric g , which, at each point (x, v) , is the product of the hyperbolic metric at $x \in X$ and arc length along the fibre over x evaluated at v . The metric g induces a measure μ on UT which is pointwise the product of hyperbolic area on X with arc length along the fibre.

The geodesic flow $\varrho_t : UT \rightarrow UT$ is defined as follows. For $(x, v) \in UT$, let $g(*)$ denote the unique geodesic on X (parametrized by arc length) having $g(0) = x$ and $\dot{g}(0) = v$. Now set $\varrho_t(x, v) = (g(t), \dot{g}(t))$. This flow is measure preserving and ergodic with respect to μ . Indeed, μ is the unique measure of maximal entropy for the geodesic flow.

There is a one to one correspondence between the periodic orbits of ϱ and the closed geodesics on X . If X is part of a hyperbolic structure (X, ϕ) on S , then via ϕ there is a one to one correspondence between closed orbits of ϕ and free homotopy classes on S . We will use $[\tau]$ to denote the free homotopy class on S corresponding

to the closed orbit τ . If $\lambda(\tau)$ is the least period of τ , then $\lambda(\tau) = l[\tau]$. Here l is the length function corresponding to (X, ϕ) .

2.2. *The Boundary Correspondence.* For a continuous function $f: UT \rightarrow \mathbf{R}$, we define $\lambda(f, \tau)$ to be the integral of f (with respect to arc length) around the closed orbit τ . In particular, if $f \equiv 1$, then $\lambda(f, \tau) = \lambda(\tau)$.

Suppose that (X_2, ϕ_2) is a second hyperbolic structure on S , and that l_2 is the corresponding length function. We will use the well known canonical homeomorphism $B: UT(X) \rightarrow UT(X_2)$ to produce a Hölder continuous function $\psi: UT \rightarrow \mathbf{R}^+$ such that $\lambda(\psi, \tau) = l_2[\tau]$ for each closed orbit τ .

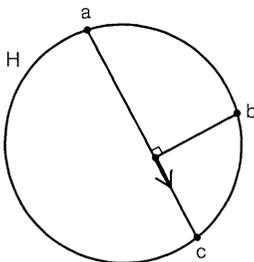


Fig. 2.2

Here is a description of B . For more details, see [BC]. Let \mathbf{H} denote the hyperbolic plane. Let $f = \phi_2 \circ \phi^{-1}: X \rightarrow X_2$. Then f lifts to a map $\tilde{f}: \mathbf{H} \rightarrow \mathbf{H}$, which covers f . The map \tilde{f} has a Hölder continuous extension $b: \mathbf{S}^1 \rightarrow \mathbf{S}^1$, which is called the boundary correspondence. A unit tangent vector of \mathbf{H} can be given by an ordered triple of distinct points on \mathbf{S}^1 , as shown in Fig. 2.2. The map $(x, y, z) \rightarrow (b(x), b(y), b(z))$ induces a map $\tilde{B}: UT(\mathbf{H}) \rightarrow UT(\mathbf{H})$. Since b is equivariant with respect to the two groups of deck transformations, \tilde{B} descends to give a Hölder continuous map $B: UT(X) \rightarrow UT(X_2)$.

The map B has the property that it takes orbits of the geodesic flow on $UT(X)$ to orbits of the geodesic flow on $UT(X_2)$. Let δ be a number which is smaller than the length of the shortest closed geodesic on either X or X_2 . Given a point $(x, v) \in UT(X)$, let $o_\delta(x, v)$ be the segment in $UT(X)$ consisting of the orbit of (x, v) up to time δ under the geodesic flow. Define $\psi(x, v)$ to be $1/\delta$ times the length of $B(o_\delta(x, v))$. (Note that we are *not* taking the limit $\delta \rightarrow 0$, which does not necessarily exist.)

The Hölder continuity of ψ follows from the Hölder continuity of B . To prove that $\lambda(\psi, \tau) = l_2[\tau]$, we will consider a completely analogous situation, which is simpler to treat. Suppose $T = \mathbf{R}/\lambda\mathbf{Z}$ and $T_2 = \mathbf{R}/\lambda_2\mathbf{Z}$ are two different circles with the obvious induced metrics, and $f: T \rightarrow T_2$ is an orientation preserving homeomorphism. Define $p: T \rightarrow \mathbf{R}^+$ by the formula

$$p(x) = \frac{f(x + \delta) - f(x)}{\delta} .$$

This difference is taken mod λ_2 , so that it is always positive. We will show that

$$\int_T p(x) dx = \lambda_2 .$$

To prove this, we may assume by approximation that f is in fact differentiable, with derivative $g = f'$. Then

$$p(x) = \frac{1}{\delta} \int_0^\delta g(x + y)dy .$$

So

$$\int_T p(x)dx = \frac{1}{\delta} \int_T \int_0^\delta g(x + y)dy dx = \frac{1}{\delta} \int_0^\delta \int_T g(x + y)dx dy = \frac{1}{\delta} \int_0^\delta \lambda_2 dy = \lambda_2 .$$

The function ψ is related to the Weil-Petersson metric on the Teichmuller space of hyperbolic structures on the surface S . In particular, the integral

$$\int_{UT} \psi d\mu$$

is the same as the quantity called the “ S length of an R geodesic” in [W]. (For us, $R = X$ and $S = X_2$.) Two salient features of this quantity are:

1. The integral doesn't depend on the choice $\delta > 0$.
2. As noted in [W], the integral is strictly greater than 1.

2.3. *Independence of ψ .* A unit complex valued function $u: UT \rightarrow S^1$ is said to be continuously differentiable with respect to the geodesic flow ρ if the limit

$$u'(y) = \lim_{t \rightarrow 0} \frac{u(\rho_t(y)) - u(y)}{t}$$

exists everywhere and is continuous.

Two functions $f, g: UT \rightarrow \mathbf{R}$ are said to be independent if the following implication

$$(*) \quad af + bg = \frac{1}{2\pi i} u'/u, \quad a, b \in \mathbf{R} \Rightarrow a = b = 0$$

holds (possibly vacuously) for each continuously differentiable function $u: UT \rightarrow S^1$. In particular, f is said to be independent if f and the constant function $g \equiv 1$ are independent. The remainder of this section is devoted to showing that ψ is independent.

Let l_1 denote the length function corresponding to the structure (X, ϕ) and let l_2 denote the length function corresponding to the structure (X_2, ϕ_2) . Suppose u is a continuously differentiable function which satisfies the left half of (*). The integral of u'/u over a closed orbit is an integer multiple of $2\pi i$, so the left-hand side of (*) translates into:

$$a_1 l_1[\gamma] + a_2 l_2[\gamma] \in \mathbf{Z}, \quad a_1, b_2 \in \mathbf{R}$$

for every free homotopy class $[\gamma]$.

Independence Lemma. *Let l_1 and l_2 be two distinct length functions. If $a_1, a_2 \in \mathbf{R}$ satisfy $a_1 l_1[\gamma] + a_2 l_2[\gamma] \in \mathbf{Z}$ for every free homotopy class $[\gamma]$, then $a_1 = a_2 = 0$.*

Proof. Below we will use a version of *Dehn Twisting* to prove:

Lemma 2.3. *For each free homotopy class $[\gamma_1]$ there is an infinite sequence of free homotopy classes $[\gamma_2], [\gamma_3], \dots$ and a constant $C = C(l_1, l_2, [\gamma_1])$ such that*

$$l_j[\gamma_k] + 2l_j[\gamma_1] - l_j[\gamma_{k+1}] \in \exp(-2kl_j[\gamma_1])[C^{-1}; C]; \quad j = 1, 2.$$

Let $[\alpha_1] = [\alpha]$ be a free homotopy class such that $l_1[\alpha] < l_2[\alpha]$, and let $[\alpha_2], [\alpha_3], \dots$ be the corresponding classes given by Lemma 2.3. Consider the vectors

$$v_k = (l_1[\alpha_k], l_2[\alpha_k]); \quad w_k = v_k + 2v_1 - v_{k+1}; \quad a = (a_1, a_2).$$

We have $a \cdot w_k \in \mathbf{Z}$. For large k , this dot product is small, implying that $a \cdot w_k = 0$. If $a \neq (0, 0)$ then these w_k lies in the linear subspace $A = a^\perp$.

Let e_1 and e_2 be the coordinate vectors of \mathbf{R}^2 . The unit vectors $w_k/\|w_k\|$ all lie in A , and by Lemma 2.3, $w_k/\|w_k\| \rightarrow e_2$. Therefore $A = \mathbf{R}e_2$, which in turn implies that the first coordinate of w_k is zero for sufficiently large k , a contradiction to Lemma 2.3. \square

Proof of Lemma 2.3. It suffices to prove Lemma 2.3 for the single length function l , induced by the hyperbolic structure (X, ϕ) . Let $[\gamma_1]$ be an arbitrary free homotopy class and let $[\gamma_2]$ be any free homotopy class which intersects $[\gamma_1]$ and which is not a multiple of $[\gamma_1]$. Let γ_1 and γ_2 be two closed curves on S , representing these classes, which have the minimum number of intersections. Let $\xi \in \gamma_1 \cap \gamma_2$. Let γ_k ($k > 2$) be the curve on S given by the following procedure: start at ξ ; wind k times clockwise around γ_1 ; wind once clockwise around γ_2 ; then wind k times clockwise around γ_1 .

The minimal intersection property guarantees that there is an isotopy of X which carries $\phi(\gamma_1)$ and $\phi(\gamma_2)$ to the two closed geodesics g_1 and g_2 representing the classes $[\gamma_1]$ and $[\gamma_2]$. (See [BC].) Thus there is a point $x \in g_1 \cap g_2$ corresponding to ξ , and a piecewise geodesic curve h_k on X corresponding to γ_k . In particular, h_k represents the free homotopy class $[\gamma_k]$. Finally, let g_k be the closed geodesic on X representing $[\gamma_k]$.

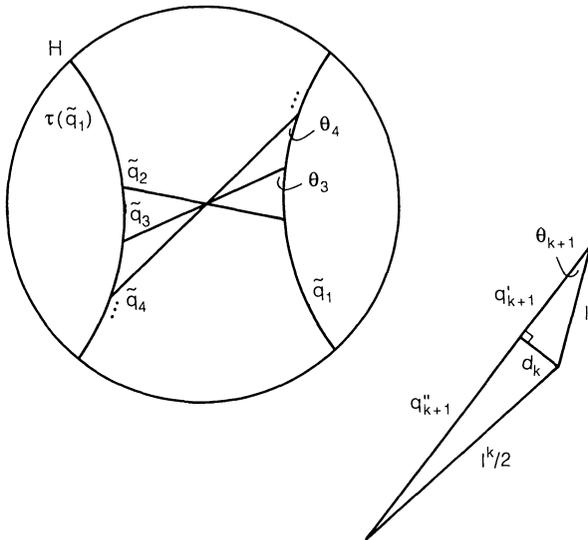


Fig. 2.3

The hyperbolic plane \mathbf{H} universally covers X . We will denote the lift of $w \in X$ to \mathbf{H} by \tilde{w} . Figure 2.3 shows the lifts to \mathbf{H} of the various curves on X . Here τ is the deck

transformation whose axis is \tilde{g}_2 and whose translation length is $l[\gamma_2]$. The segment of \tilde{g}_k lying between \tilde{g}_1 and $\tau(\tilde{g}_1)$ covers g_k exactly once under the quotient map $\mathbf{H} \rightarrow X$. Indeed, by symmetry, the deck transformation identifying the endpoints of this segment correctly matches up the tangent vectors to produce the closed geodesic g_k on the quotient. Thus $l[\gamma_k]$ is exactly the length of the portion of \tilde{g}_k lying between \tilde{g}_1 and $\tau(\tilde{g}_1)$. Set $l^k = l[\gamma_k]$.

Below we will use the notation $A \approx B$ iff $A \in (B - C, B + C)$ and $A \asymp B$ iff $A \in (C^{-1}B, CB)$. Here C is a fixed but unspecified constant. We will also use the following hyperbolic trigonometry formulae [B] for a right-angled hyperbolic triangle having hypotenuse C and angle β opposite the side B .

1. $\sinh C = \sinh B / \sin \beta$ (Law of Sines);
2. $\cosh C = \cosh B \cosh A$ (Law of Cosines).

From Figure 2.3 it is clear that $l^k \approx 2kl^1$. Using the Law of Sines, with $C = l^k/2$, $B \asymp 1$, and $\beta = \theta_k$, we obtain $\theta_k \asymp \sin \theta_k \asymp \exp(-kl^1)$. Using the Law of Sines again with $C = g'_{k+1} \asymp 1$, $B = d_k$, and $\beta = \theta_k$, we obtain $d_k \asymp \sinh d_k \asymp \exp(-kl^1)$. Since d_k is small, we have $g''_{k+1} \asymp l^k/2$. To summarize, we have

$$d_k \asymp \exp(-kl^1); \quad l(g'_{k+1}) \asymp l^1 \asymp 1; \quad l(g''_{k+1}) \approx l^k/2 \approx kl^1.$$

We also have the equation:

$$l^1 + \frac{1}{2}l^k - \frac{1}{2}l^{k+1} = (l^1 - l(g'_{k+1})) + (l^k/2 - l(g''_{k+1})).$$

We may write the Law of Cosines as $\cosh C - \cosh A = \cosh A (\cosh B - 1)$. If A and C are pinched away from 0, $A \approx C$, and $B < 1$ then $C - A \asymp \cosh B - 1 \asymp B^2$. The first of these estimates follows from the expansion properties of \cosh on $[A, C]$, and the second follows from Taylor's series. Taking $A = l(g'_{k+1})$, $B = d_k$, and $C = l^1$ gives $l^1 - l(g'_{k+1}) \asymp \exp(-2kl^1)$. Taking $A = l(g''_{k+1})$, $B = d_k$, and $C = l^k/2$ gives $l^k/2 - l(g''_{k+1}) \asymp \exp(-2kl^1)$. Adding these estimates proves the lemma. \square

We conclude from the Independence Lemma that ψ is independent with respect to the geodesic flow on $UT(X)$.

3. Proof of Theorem 1

We will deduce Theorem 1 from a result of Lalley [L] concerning the distribution of periodic orbits of a hyperbolic flow. We begin by introducing some ideas from the thermodynamic formalism. A good reference for these ideas is [PP].

Recall that $UT = UT(X)$ is the tangent bundle to the hyperbolic surface X , and ϱ_t is the geodesic flow on UT . Given any invariant ergodic measure ν on UT , let $h(\nu)$ denote the measure-theoretic entropy of ϱ with respect to ν .

Given a Hölder continuous function $f: UT \rightarrow \mathbf{R}$, we define the *pressure* of f to be

$$P(f) = \sup_{\nu} \left\{ h(\nu) + \int_{UT} f d\nu \right\}.$$

Here the supremum is taken over all invariant probability measures. There is a unique such measure μ_f , for which this supremum is attained. μ_f is called the equilibrium state for f . In particular, the equilibrium state for $f \equiv 0$ is the measure μ , described in Sect. 2.1. This is just a restatement of the fact that μ is the unique probability measure of maximal entropy.

Similar to Sect. 2.3, we say that a function $v:UT \rightarrow \mathbf{R}$ is continuously differentiable with respect to ϱ if the limit

$$v'(y) = \lim_{t \rightarrow 0} \log \frac{v(\varrho_t(y)) - v(y)}{t}$$

exists everywhere, and is continuous. Two Hölder continuous functions f and g are said to be *cohomologous* if $f - g = v'$ for some continuously differentiable v . Two Hölder continuous functions f and g have the same equilibrium state if and only if $f - g$ is cohomologous to a constant (function), say c . In this situation $P(f) = P(g) + c$.

Let f be a Hölder continuous function, we shall be interested in the real variable function $t \rightarrow P(tf)$. If f is *not* cohomologous to a constant function, then this function is analytic and strictly convex. Furthermore,

$$P'(tf) \stackrel{\text{def}}{=} \frac{d}{dt} P(tf) = \int_{UT} f d\mu_{tf}$$

holds for each value of $t \in \mathbf{R}$. Let $J(f)$ denote the open interval of values $P'(tf)$. If $a \in J(f)$, let $t_a \in \mathbf{R}$ be the unique real number for which $P'(t_a f) = a$.

Lalley's result concerns Hölder continuous functions on UT which are independent in the sense of Sect. 2.3. It is easy to see that an independent function cannot be cohomologous to a constant.

Proposition [L]. *Let $f : UT \rightarrow \mathbf{R}$ be an independent Hölder continuous function and let $a \in J(f)$. Then, for fixed $\varepsilon > 0$, there is a constant $C = C(f, \varepsilon)$ such that*

$$\#\{\tau : \lambda(\tau) \in (x, x + \varepsilon), \lambda(f, \tau) \in (ax, ax + \varepsilon)\} \sim C \frac{\exp(h(\mu_a)x)}{x^{3/2}}.$$

Here we have set $\mu_a = \mu_{t_a f}$.

Remarks. (i) Our statement of Lalley's result differs slightly from the original in [L]. First, our independence condition is a minor rephrasing of the one appearing in [L]. Second, the original is stated for C^∞ functions instead of Hölder functions, and only for values of a in a small neighborhood of $\int f d\mu$. However, these more stringent conditions are not necessary. (See [S].)

(ii) The constant $C = C(f, \varepsilon)$ is related to the second derivative $P''(t_a f)$, and has the same order of magnitude as ε^2 .

(iii) The geodesic flow has positive entropy with respect to any equilibrium state, so $0 < h(\mu_a) \leq h(\mu) = 1$. Equality occurs if and only if μ_a is the measure of maximal entropy, which is to say that $t_a = 0$ and (hence) $a = \int f d\mu$.

We shall apply Lalley's result to the function ψ , constructed in Sect. 2.2, and seen to be independent in Sect. 2.3. To prove Theorem 1, we just need to show that $1 \in J(\psi)$, and that $h(\mu_1) < 1$.

As in Sect. 2.3, let l_1 and l_2 be the two length functions associated to the hyperbolic structures (X, ϕ) and (X_2, ϕ_2) .

Let $I(\psi)$ denote the set of values $\int \psi d\nu$, where ν ranges over invariant probability measures. Clearly $I(\psi)$ is a closed interval. For any closed orbit τ , $\lambda(\psi, \tau)/\lambda(\tau) \in I(\psi)$. It is standard that there exist two homotopy classes $[\gamma_1]$ and $[\gamma_2]$ such that $l_1[\gamma_1] > l_2[\gamma_1]$ and $l_1[\gamma_2] < l_2[\gamma_2]$. These correspond to periodic orbits τ_1 and τ_2 for which $\lambda(\psi, \tau_1)/\lambda(\tau_1) < 1$ and $\lambda(\psi, \tau_2)/\lambda(\tau_2) > 1$. Therefore $(1 - 2\varepsilon, 1 + 2\varepsilon) \in I(\psi)$ for small $\varepsilon > 0$.

The following line of reasoning is adapted from [MT]. From the above, we have $1 \pm \varepsilon \in I(\psi)$, and hence

$$t + |t|\varepsilon \in tI(\psi) \quad \forall t.$$

It is clear from the definition of pressure that $y \in I(\psi)$ implies $P(t\psi) \geq ty$. Hence

$$P(t\psi) \geq \sup tI(\psi) \quad \forall t.$$

Combining the last two inequalities gives

$$P(t\psi) - t \geq |t|\varepsilon \quad \forall t.$$

Consider the function $Q(t) = P(t\psi) - t$. Since $P(0) = 1$, we have $Q(0) = 1$. Also, if $|t| > 1/\varepsilon$ then $Q(t) > 1$. Therefore $Q(*)$ must have a minimum for some value in $x \in [-1/\varepsilon, 1/\varepsilon]$. At this point $Q'(x) = 0$, which is to say that $P'(x\psi) = 1$, as desired.

From Remark (iii) above, $h(\mu_1) = 1$ if and only if $\int \psi d\mu = 1$. However, this integral is strictly greater than 1, by Remark (ii) of Sect. 2.2.

This completes the proof of Theorem 1.

References

- [B] Beardon, A.: The geometry of discrete groups. Berlin, Heidelberg, New York: Springer 1983
- [BC] Bleiler, S., Casson, A.: Automorphisms of surfaces after Nielsen and Thurston. L.M.S. Student Texts, 1990
- [H] Huber, H.: Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen. II. Math. Ann. **142**, 385–398 (1961)
- [L] Lalley, S.: Distribution of periodic orbits in symbolic and Axiom A flows. Adv. in Appl. Math. **8**, 154–193 (1987)
- [MT] Marcus, B., Tuncel, S.: Entropy at a weight-per-symbol and embeddings of Markov chains. Inv. Math. **102**, 235–266 (1990)
- [PP] Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. Asterisque, 187–188 (1990)
- [S] Sharp, R.: Prime orbit theorems with multi-dimensional constraints for Axiom A flows. Monat. Math. **114**, 261–304 (1992)
- [W] Wolpert, S.: Thurston's Riemannian metric for Teichmüller space. J. Diff. Geo. **23**, 143–174 (1986)

Communicated by J.-P. Eckmann