TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 272, Number 2, August 1982

DETERMINING THURSTON CLASSES USING NIELSEN TYPES

BY

JANE GILMAN

ABSTRACT. In previous work [3] we showed how the Thurston or Bers classifications of diffeomorphisms of surfaces could be obtained using the Nielsen types of the lifts of the diffeomorphism to the unit disc. In this paper we find improved conditions on the Nielsen types for the Thurston and Bers classes. We use them to verify that an example studied by Nielsen is a pseudo-Anosov diffeomorphism with stretching factor of degree 4. This example is of interest in its own right, but it also serves to illustrate exactly how the Nielsen types are used for verifying examples. We discuss the general usefulness of this method.

1. Introduction. Let S be a compact Riemann surface of genus g with n boundary components, with $2g-2+n>0$, and let $M(S)$ be its mapping class group (also known as the Teichmüller modular group). Using measured foliations, Thurston [9] has shown that $M(S)$ contains three different types of elements. Bers has shown that this decomposition can be obtained using Teichmüller distance [1]. Subdividing one of the classes further, he obtains four classes.

If t is a diffeomorphism of S, let $T=\langle t\rangle$ be the group generated by t and $L(T)$ the group of lifts of elements of T to the unit disc, U. For h in $L(T)$ we let the pair of integers (v_h, u_h) denote the *Nielsen type* of h (see [6] or [7-I, II, and III] or [3, §4]). In [3] we showed how the Thurston or Bers classes could be defined in terms of the possible Nielsen types of the elements of $L(T)$ (see Theorems 5.2, 5.3, 15.2 and 15.3 of [3]). We also showed (Theorem 5.1 of [3]) how one could prove a decomposition theorem for $M(S)$ directly from the Nielsen theory. However, we noted in the introduction that unless the criteria for the various Nielsen types could be further simplified it seemed unlikely that it would be very useful.

In this paper we obtain better conditions on the Nielsen types for the three Thurston classes and the four Bers classes (Theorem 4). They are useful. To illustrate this we study Nielsen's example #13 (p. 347 of [7-I]) which turns out to be a pseudo-Anosov diffeomorphism with stretching factor, λ, of degree four.

While this example does serve to illustrate how the Nielsen types associated to a diffeomorphism can be used to verify that a given diffeomorphism is pseudo-Anosov, it is of interest in its own right. It was mentioned at the end of Chapter 13 of [8] that

Received by the editors April 27, 1981. Presented at the special session on Low Dimensional Topology at the American Mathematical Society meeting in Kenosha, Wisconson, November 1, 1980.

1980 *Mathematics Subject Classification*. Primary 14H99, 30F99, 32G15.

Key words and phrases. Mapping-class group, Teichmüller modular group, Riemann surface, automorphism, pseudo-Anosov diffeomorphism, foliations, Fuchsian group.

all known examples of pseudo-Anosov diffeomorphism have stretching factor, λ, of degree 2. J. Franks **[2]** has subsequently characterized all which do have degree 2. D. Fried has found examples of degree 6 and 8, and Veech **[10]** has a construction which may yield all examples of degree 4. While Veech's construction yields the foliations of the diffeomorphism from which one then is able to compute the genus of the surface, it is not clear how to find the action induced on the fundamental group. In our case, the diffeomorphism is defined by the automorphism of the fundamental group.

In §2 of this paper we define terminology. §3 contains the proof of the improved equivalence theorem and §4 contains the example. In §5 we discuss to what extent the techniques used in §4 are general.

2. Notation and terminology. We briefly recall the definition of the Nielsen type (v_h, u_h) of an h in $L(T)$. (See §4 of **[3]** or **[6]** and **[7]** for more details.)

If F is the Fuchsian group with $U/F = S$, then $L(T)$ is an extension of F and $h \in L(T)$ acts on F by conjugation. Let $N_h = \{f \in F \mid hfh^{-1} = f\}$ and let v_h be the minimal number of generators of N_h. Here $v_h = 0$ if $N_h = \text{id}$.

If $v_h \neq 1$, u_h is the number of N_h orbits of Nielsen isolated attracting fixed points of the extension of h to the limit set of F. For simplicity we also denote the extension by h.

If $v_h = 1$, we assume that N_h is generated by a primitive element f of F. We let Ax_f be its axis and V_f and U_f its attracting and repelling fixed points on the boundary of U. In this case $u_h =$ the number of N_h orbits of attracting Nielsen isolated fixed points of the extension of h to the limit set of F, except we exclude V_f and U_f from this set.

We further subdivide type $(1,0)$ as follows: $(1,0)^{**}$ means that V_f and U_f are alternately attracting and repelling fixed points for the extension of h while $(1,0)^*$ means that they are neutral fixed points. We use the notation $(1,0)^{**}\partial$ and $(1,u)\partial$ to denote the fact that the axis of f projects onto a boundary curve of S.

We remind the reader that Thurston shows that a diffeomorphism is either (a) isotopic to one of finite order, (b) isotopic to a reducible one or (c) isotopic to a pseudo-Anosov one. Further in the infinite case, either (b) or (c) occurs but not both. A reducible diffeomorphism fixes a partition, a set of disjoint, nonhomotopic simple closed curves, none of which is homotopic to a boundary curve. Each reducible diffeomorphism fixes a maximal partition. The curves in such a partition divide the surface into a finite number of components which t permutes. Each component is fixed by some smallest power of t and the restrictions of these powers to the components are called the *component maps*. Infinite reducible diffeomorphisms (reduced by a maximal partition) either have all component maps finite up to isotopy in which case they are called *parabolic* or have at least one pseudo-Anosov component map (up to isotopy) in which case they are called *pseudohyperbolic*.

We adopt the Riemann surface theory convention that a mapping t is called finite, reducible, pseudo-Anosov, etc., if its mapping class, denoted $[t]$, contains such an element. Finite, of course, means of finite order up to homotopy and infinite means not of finite order up to homotopy.

In [**3**] we proved

THEOREM 1 (NIELSEN TETRACHOTOMY). *Given $[t] \in M(S)$, then precisely one of the following holds*:

(i) *For each $h \in L(T)$, $h \neq \mathrm{id}$, h is either of type $(1,0)^{**}$ or of type $(0,0)$.*

(ii) *For each $h \in L(T)$, $h \notin F$, $v_h = 0$ unless h is of type $(1,0)^{**}\partial$ or $(1,u)\partial$ and $\exists h \in L(T)$ either of type $(1,u)\partial$ with $u \neq 0$ or $(0,u)$ with $u \geqslant 2$.*

(iii) *$\exists h \in L(T)$, $h \notin F$ with $v_h \neq 0$ and $\exists h' \in L(T)$ with $u_{h'} \neq 0$ and either we may assume that $h = h'$ and h is not of type $(1,u)\partial$ or if $h \neq h'$, h' is either of type $(1,u)\partial$ or $u_{h'} \geqslant 2$ and h is not of type $(1,0)^{**}\partial$ or $(1,u)\partial$.*

(iv) *$\forall h \in L(T)$, $u_h = 0$ unless h is of type $(0,1)$ and $\exists h \in L(T)$, $h \neq \mathrm{id}$, with either $v_h \geqslant 2$ or h of type $(1,0)^*$, where $h = g^2$, $g \in L(T)$ of type $(0,0)$.*

THEOREM 2 (EQUIVALENCE THEOREM). (i) *$[t]$ is elliptic (of finite order) $\Leftrightarrow$* 1 (i) *holds.*

(ii) *$[t]$ is hyperbolic (pseudo-Anosov) $\Leftrightarrow$* 1 (ii) *holds.*

(iii) *$[t]$ is pseudohyperbolic (infinite reducible with at least one pseudo-Anosov component map) $\Leftrightarrow$* 1 (iii) *holds.*

(iv) *$[t]$ is parabolic (infinite reducible with finite component maps) $\Leftrightarrow$* 1 (iv) *holds.*

THEOREM 3. *Given $[t]$ of infinite order, $[t]$ is reducible if and only if $\exists h \in L(T)$ with $v_h \neq 0$ and h is not of type $(1,0)^{**}\partial$ or $(1,u)\partial$.*

Finally recall that the *index of h*, $j(h) = 1 - v_h - u_h$ for $h \in L(T)$.

3. The improved equivalence theorem. Recall that t is elliptic if $[t]$ is of finite order. Our main lemma is

LEMMA 1. *Let t be any nonelliptic diffeomorphism of a compact Riemann surface. Then t is reducible $\Leftrightarrow \exists h \in L(T)$ with $v_h \neq 0$ and $j(h) < 0$.*

PROOF. If t is reducible, then either there is a component map which is finite or all component maps are pseudo-Anosov. Following the notation of §11 of [**3**], let B be a component of K_F, the Nielsen kernel of F. If the first case occurs, then $\exists f \in L(T)$ with $v_{f,B} = \operatorname{rank} F_B \geqslant 2$. Then $v_f \geqslant 2$. If the latter occurs, by Corollary 11.3 of [**3**], the component map satisfies Theorem 1 (iii) on B. Thus either $u_{f,B} \geqslant 2$ or $(v_{f,B}, u_{f,B}) = (1, u_B)\partial$ and $u_B \neq 0$. The image of B is a surface with boundary. On each boundary component the foliations of a pseudo-Anosov map must have a singularity and near the nonsingular points, the boundary is a leaf (see [**9**]). In terms of the geodesic laminations, this means that there exists an f of type $(1, u_B)\partial$, $u_B \neq 0$. (For a detailed discussion see [**5**].) Viewed as an element of $L(T)$, f then has $v_f \neq 0$ and $u_f \neq 0$, but f is not of type $(v_f, u_f)\partial$ because some boundary of B is not a boundary of S. The backward implication is just Theorem 8.1 of [**3**].

DEFINITION. $L_0(T) =$ *the essential lifts of* $T = \{h \in L(T) - \mathrm{id} \mid j(h) < 0\}$.

As a corollary to Lemma 1 we need only look at $L_0(T)$ to determine the classification of $[t]$. We obtain

THEOREM 4 (EQUIVALENCE THEOREM). *Given* $[t] \in M(S)$, *then*

(i) $[t]$ *is elliptic* $\Leftrightarrow L_0(T) = \varnothing$.

(ii) $[t]$ *is pseudo-Anosov* $\Leftrightarrow \forall h \in L_0(T)$, $v_h = 0$ *unless* h *is of type* $(1, u)\partial$ *and* $\exists h \in L_0(T)$ *with* $u_h \neq 0$.

(iii) $[t]$ *is pseudohyperbolic* (*infinite reducible with at least one pseudo-Anosov component map*) $\Leftrightarrow \exists h \in L_0(T)$ *with* $u_h \neq 0$ *and* $v_h \neq 0$, *but* h *is not of type* $(1, u_h)\partial$.

(iv) $[t]$ *is parabolic* (*infinite reducible with finite component maps*) $\Leftrightarrow \forall h \in L_0(T)$ $u_h = 0$ *and* $\exists h \in L_0(T)$ *with* $v_h \neq 0$.

PROOF. Part (ii) is merely Theorem 2 (ii) translated into the language of $L_0(T)$. The proof of Lemma 1 together with the translation to the language of $L_0(T)$ actually shows that Theorem 2 (iii) holds if and only if Theorem 4 (iii) holds and that Theorem 2 (iv) holds if and only if Theorem 4 (iv) holds. It is clear that if $[t]$ is elliptic, $L_0(T) = \varnothing$. To see the converse, assume $L_0(T) = \varnothing$. Then for each lift $h \in L(T)$, h is either of type $(0,0)$, $(1,0)^{**}$ or $(1,0)^*$. If there is a lift of type $(1,0)^*$, then the case of Theorem 1 part (i) does not occur. Thus either 1 (ii), 1 (iii) or 1 (iv) occurs. In each of these cases, we have shown that 4 (ii), (iii), or (iv) occurs. Thus in each of these cases $L_0(T) \neq \varnothing$.

REMARK. It is possible that leaving 4 (i) in the form of 1 (i) is more informative.

REMARK. Theorem 4 still looks unwieldly because one must know (v_h, u_h) for all h in $L_0(T)$. In fact it suffices to know the type for a finite number of lifts as will be seen in §4. The difficulty is to choose the finite number of lifts propitiously (see §5).

4. Nielsen's example #13. In this section we assume that S is a compact Riemann surface of genus 2 and that F has presentation

$$F = \langle a, b, c, d; [a, b][c, d] = 1 \rangle.$$

We define t by requiring that some h in $L(T)$ induce the automorphism of F given by

$$\text{(I)} \qquad \begin{aligned} &a \to c^{-1}a^{-1}, \qquad &&b \to b^{-1}a^{-1}, \\ &c \to b^{-1}a^{-1}d, \qquad &&d \to c^{-1}. \end{aligned}$$

Nielsen (see pp. 347–348 of [7-I]) computes that h is of type $(0,0)$ with h^4 of type $(0,4)$. He also shows (pp. 73–74 of [7-II]) that ch is of type $(0,0)$ with $(ch)^4$ of type $(0,4)$. Moreover, by looking at the boundary fixed points of h^4 and $(ch)^4$ it is clear that they are not conjugate in $L(T)$ by an element of F. (Elements of $L(T)$ which are conjugate by elements of F are said to be *equivalent* and equivalent lifts have the same type.)

LEMMA 2. *If f in $L(T)$ has negative index, then f is either of type* $(0,2)$ *or* $(0,4)$. *When the latter occurs, f is equivalent to a power of either h^4 or $(ch)^4$ which are not themselves equivalent.*

PROOF. From area considerations Nielsen derives (p. 70 of [7-II])

$$\text{(II)} \qquad \sum_{(s = \text{equivalence classes of lifts of } t^4)} 2(v_s - 1) + u_s \leq 4(g - 1).$$

Here g is of course the genus. In our particular case, since $(ch)^4$ and h^4 are not equivalent and $g = 2$, any lifts in $L_0(T)$ must be equivalent to $(ch)^4$ or h^4 or be of type $(0, 2)$ which contributes nothing to the sum.

COROLLARY 1. *The t defined by equation* (I) *is pseudo-Anosov.*

PROOF. Apply Lemma 2 and Theorem 4 (ii).

COROLLARY 2. *The stretching factor, λ, of t satisfies $x^4 + 2x^3 + x^2 + 2x + 1 = 0$ which is irreducible. The fixed foliations of t each have two four-pronged singularities which t permutes cyclically.*

PROOF. The geodesic laminations determined by the principal region (see [**3**], [7-II], [**6**] or [**5**] for the definitions) of an element of type $(0, 4)$ will correspond to a four-pronged singularity of the unstable foliation, denote it by $\mathfrak{F}_u$, on S. (For a detailed discussion of the relationship between the principal regions and the prongs of the singularities of the foliations see [**5**].) Type $(0, 2)$ determines no singularity. Type $(0, 0)$ determines no singularity directly. (An h of type $(0, 0)$ has a power either of type $(0, 2)$ or $(0, 4)$.) Thus the foliation $\mathfrak{F}_u$ has precisely two four-pronged singularities. If $\hat{S}$ is the two-sheeted cover of S branched along the singularities which make $\mathfrak{F}_u$ transversely oriented, then λ is an eigenvalue of the induced matrix representation of the action of $\hat{t}$ on homology. This is the remark following Theorem 6 of (one of the versions) of [**9**] and $\hat{t}$ is the lift of t to $\hat{S}$. Since $\mathfrak{F}_u$ has only even pronged singularities, $\hat{S} = S$. Thus λ is merely an eigenvalue for the matrix of the action on homology of the automorphism given by (I). We compute that λ satisfies the above polynomial and that is irreducible. Note also that since h^4 and $(ch)^4$ are not equivalent, we know that they determine distinct singularities.

Nielsen's computations of the fixed points of h show exactly that t acts on the prong points by sending the 4-tuple (P_1, P_3, P_5, P_7) into the 4-tuple (P_3, P_5, P_7, P_1). (See pp. 347–349 of [7-I].)

5. Discussion of the general method for verification. One would like to understand to what extent Nielsen's results yield a general method for verifying examples beginning with the assumption that one is given the action of the automorphism on a set of generators for the fundamental group. It turns out that except for two or three points in Nielsen's calculations the method is general and useable.

The most interesting thing about example #13 is the theory behind Nielsen's computation of the type of h^4 and $(ch)^4$. Given a fixed set of generators for F, a point on the limit set of F corresponds to an infinite word in these generators, called the *development of the point*. (See §7 of [7-I].) Fixed points of elements of F have periodic developments while other limit points have nonrepeating developments. These developments are, therefore, referred to as *rational and irrational developments* or rational and irrational points on the limit set. Nielsen actually computes the developments of the eight boundary fixed points of h^4.

These computations with developments are the Nielsen analogue of Thurston's computations with train tracks. The development of a point contains a great deal of geometric information about the point: developments of any three points determine their cyclical order on the boundary of U and developments of any two points determine which elements of F map one to the other. This is used when Nielsen claims that h^4 and $(ch)^4$ are not conjugate by an element of F.

The first two stumbling blocks arise as follows:

1. Once Nielsen has the two inequivalent lifts, h^4 and $(ch)^4$, and their types, equation (II) tells him that he has found enough lifts. However, how does he know a priori that the choice of ch as a lift to have its type computed will be productive? (E.g. a lift fh for another f in F might lead to a lift already equivalent to h or to a lift not in $L_0(T)$.) A partial answer lies in Nielsen's Theorems 17 and 18, parts (a), (b) and (c) of [7-II]. These theorems contain results about how the type of fh is related to the type of h. It assumes that one knows the fixed point structure of h and the boundary locations of V_f and U_f in relation to the fixed points of h. These results, however, only tell for what choices of f, fh will definitely not be in $L_0(T)$ and for which choices fh may possibly be in $L_0(T)$. To reduce guess work, one would like to prove a stronger result than "possibly".

2. The computation of the type of a lift proceeds upon the principle that the images of a point on the limit set under powers of the lift accumulate to fixed points of the lift. One arbitrarily picks V_a, for example, and computes that the developments of $h^n(V_a)$ converge to the irrational development of a point, P_1. Similarly if one computes that the developments of $h^{-n}(V_b)$ converge from the opposite side of P_1 to the development of P_1, then one can conclude that h has no fixed points other than P_1 between V_a and V_b. Nielsen's method is to continue in this manner until he has located "enough" boundary fixed points of h. However, there is a degree of choice here, too. It may turn out, for example, that $h^n(V_c)$ and $h^n(V_d)$ converge to the same irrational fixed point, and it is not clear why Nielsen knows in computing example #13 that the choice of $h^n(V_{d^{-1}a})$ will be productive.

Aside from these two questions that need to be answered, the procedure outlined for example #13 could be used to verify the Thurston or Bers class of any diffeomorphism beginning with the knowledge of its action on generators for the fundamental group, and not only for pseudo-Anosov maps. Applying the technique described in 2 above, roughly speaking, one would recognize the reducible case by the existence of fixed points with periodic developments and the pseudo-Anosov or the pseudohyperbolic cases by the existence of those with irrational developments.

A complete determination of the type of all the elements of $L_0(T)$ is easier in the pseudo-Anosov case because we know that equation (II) is actually an equality. (This is the singularity structure of the quadratic differential explained by the equations at the bottom of p. 226 of Hubbard and Masur [**4**].) This inequality might be another stumbling block. However, in many cases a complete determination of the types of all elements of $L_0(T)$ is not actually required.

A final remark needs to be made about computing the degree of λ, the stretching factor. In example #13 the computation was simplified because $\hat{S} = S$. In the general case, there is not really more trouble. Once one knows the action of t on the prongs of the singularities, one can easily find the action of $\hat{t}$ on the homology of $\hat{S}$.

References

1. L. Bers, *An extremal problem for quasiconformal mapping and a theorem of Thurston*, Acta Math. **141** (1978), 73–98.

2. J. Franks, unpublished.

3. J. Gilman, *On the Nielsen type and the classification for the mapping-class group*, Adv. in Math. **40** (1981), 68–96.

4. J. Hubbard and H. Masur, *Quadratic differentials and foliations*, Acta Math. **142** (1979), 223–274.

5. R. Miller, *Nielsen's viewpoint on geodesic laminations*, Adv. in Math. (to appear).

6. J. Nielsen, *Surface transformation classes of algebraically finite type*, Mat.-Fys. Medd. Danske Vid. Selsk. **21** (2) (1944), 1–89.

7. ______, *Untersuchungen zur Topologie der geschlossenen zweiseitigen Flachen*. Parts I–III, Acta Math. **50** (1927), 189–358; **53** (1929), 1–76; **58** (1932), 87–167.

8. V. Poénaru et al., *Traveaux de Thurston*, Astérisque **66–67**, Société Mathématique de France, 1980.

9. W. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*. I, preprint.

10. W. Veech, *Gauss measures of transformations on the space of interval exchange maps*, Ann. of Math. (to appear).

Department of Mathematics, Newark College of Arts and Science, Rutgers University, Newark, New Jersey 07102