GEOMETRIC AND ALGEBRAIC TOPOLOGY
BANACH CENTER PUBLICATIONS, VOLUME 18
PWN - POLISH SCIENTIFIC PUBLISHERS
WARSZAWA 1986

ALGEBRAIC PROPERTIES OF MAPPING CLASS GROUPS OF SURFACES*

N. V. IVANOV

Leningrad, U.S.S.R.

This paper is an extended version of the author's report at the Semester on Topology at the Stefan Banach International Mathematical Center in June 1984. The aim of the paper is to survey author's recent results concerning the structure of mapping class groups of surfaces. To expose the matter in perspective, many other results are mentioned, although our discussion is without doubt incomplete. The paper is adressed to non-experts, and therefore we have included necessary preliminaries from the topology of surfaces and the theory of Teichmüller spaces. Proofs are omitted, but motivations, ideas of proofs and technical tools will be explained in detail.

Almost all results presented in the paper were already announced in [21] and at the Warsaw Congress [20] but some of them are new; among them the non-closed case of Theorem 12, Theorems 11 and 16 and, partly, the material of Section 3.

It is a pleasant duty of the author to express his gratitude to Professor H. Toruńczyk for his hospitality during the author's visit to Banach Center.

CONTENTS

1. Introduction

Let X be a smooth compact orientable surface, possibly with boundary. The main object of our interest is the group of orientation-preserving diffeomorphisms $X \to X$ considered up to isotopy. More rigorously, this group is defined as the factorgroup $\mathrm{Diff}^+(X)/\mathrm{Diff}_0(X)$, where $\mathrm{Diff}^+(X)$ is the

* This paper is in final form and no version of it will be submitted for publication elsewhere.

group of all orientation-preserving diffeomorphisms $X \to X$ and $\mathrm{Diff}_0(X)$ is the subgroup of diffeomorphisms isotopic to the identity. Topologists usually call it the mapping class group of X. This group is also known to analysts as the Teichmüller modular group. The first term is nearly self-explanatory. The origin of the second one will be explained later, for it is intimately related to our methods of investigation of this group. That is why this group will be denoted Mod_X throughout the text, in accordance with the second term.

If X is a closed surface, then every diffeomorphism $f: X \to X$ gives rise to a 3-manifold by the well-known Heegaard construction. The diffeomorphism class of this 3-manifold depends only on the isotopy class of f. Apparently, this observation explains the interest of low-dimensional topologists to mapping class groups. Isotopy classes of surface diffeomorphisms appear naturally in many other problems of 3-manifold topology, e.g. in the study of fibrations over the circle and of fibered knots. As for the surface topology, these isotopy classes are the main objects of study. The mapping class groups play a fundamental role not only in low-dimensional topology, but also in the theory of Riemann surfaces. These groups are intimately related to the moduli spaces of Riemann surfaces. However, the mapping class groups have been studied by topologists much more intensively than by analysts. The present paper is not an exception.

Many interesting questions about the mapping class groups can be stated in terms of surface topology as well as of the theory of Teichmüller spaces (to be defined below). But we shall restrict ourselves to the investigation of purely algebraic properties of mapping class groups. If X is a 2-dimensional torus, then Mod_X is isomorphic to $SL_2(\boldsymbol{Z})$ (the isomorphism is given by the action of diffeomorphisms on $H_1(X, \boldsymbol{Z}) \approx \boldsymbol{Z}^2$). The group $SL_2(\boldsymbol{Z})$ occurs in almost all branches of mathematics and has been thoroughly studied from many points of view. Its higher dimensional generalizations, such as the groups $SL_n(\boldsymbol{Z})$ and, more generally, arithmetic groups, have been also treated in detail in a large number of papers. The groups Mod_X can be considered as higher genus ("multi-handled") generalizations of $SL_2(\boldsymbol{Z})$, and hence provide an alternative and a partial counterpart to the study of $SL_n(\boldsymbol{Z})$ and arithmetic groups (I owe this remark to A. M. Vershik).

The above definition of mapping class groups is the modern differential version of the classical topological definition. Namely, Mod_X can be defined as $\mathrm{Homeo}^+(X)/\mathrm{Homeo}_0(X)$, where $\mathrm{Homeo}^+(X)$ is the group of orientation-preserving homeomorphisms $X \to X$ and $\mathrm{Homeo}_0(X)$ is the subgroup of homeomorphisms isotopic to the identity. This definition is equivalent to the previous one because every homeomorphism $X \to X$ is isotopic to a diffeomorphism and if two diffeomorphisms can be connected by a path of homeomorphisms then they can be connected by a path of diffeomorphisms, too. (Note that higher-dimensional analogues of these statements are false.)

There is also a homotopy-theoretical description of Mod_X, which we shall state for simplicity only for closed surfaces. Namely, if X is a closed surface, then Mod_X can be defined as the group of homotopy classes of all homotopy equivalences $X \to X$ which act trivially on $H_2(X, \mathbf{Z})$. The equivalence of this definition to the previous ones is a classical result of Dehn and Nielsen. Since X is either the 2-sphere or an Eilenberg–MacLane space, this homotopy definition allows us to compute Mod_X in terms of $\pi_1(X)$ by applying elementary obstruction theory. It turns out that Mod_X is isomorphic to $\mathrm{Aut}(\pi_1(X))/\mathrm{Inn}(\pi_1(X))$ where $\mathrm{Aut}(\pi)$ denotes the group of all automorphisms of a group π and $\mathrm{Inn}(\pi)$ denotes the subgroup of inner automorphisms, that is, automorphisms of the form $g \mapsto hgh^{-1}$, $h \in \pi$. The group $\mathrm{Aut}(\pi)/\mathrm{Inn}(\pi)$ is called the outer automorphisms group of π and is denoted by $\mathrm{Out}(\pi)$ (the subgroup $\mathrm{Inn}(\pi)$ is always normal in $\mathrm{Aut}(\pi)$). Thus Mod_X is isomorphic to $\mathrm{Out}(\pi_1(X))$; of course the isomorphism is given by the action of diffeomorphisms (or homeomorphisms or homotopy equivalences) on $\pi_1(X)$, which is well-defined up to inner automorphisms. Using the well-known presentation of $\pi_1(X)$ by generators and relations, we obtain a purely algebraic definition of Mod_X for closed X. There is also a similar definition in the non-closed case. To give such a definition, it is necessary to express (firstly homotopically and then) algebraically the fact that every diffeomorphism preserves the boundary of X. It is rather surprising that this algebraic definition plays a very small role in the study of algebraic properties of Mod_X.

2. Teichmüller spaces

One of the most effective and beautiful ways, which permits to understand the structure of a group, is to study its actions on suitable geometrical objects. For mapping class groups the most important objects of this type are, beyond doubts, the Teichmüller spaces. In the following definition of the Teichmüller spaces, we restrict ourselves for simplicity to closed surfaces.

Thus, let X be a smooth closed orientable surface. Fix an orientation of X. Roughly speaking, the Teichmüller space of X is the space T_X of complex structures on X, i.e. structures of 1-dimensional complex manifold on X, which are consistent with the smooth structure and orientation and considered up to isotopy. To make this definition rigorous it is necessary, first of all, to introduce a topology in the set of complex structures on X. A complex structure on X is determined by the automorphism of the tangent bundle TX given by the formula $v \mapsto \sqrt{-1}\,v$. Hence, we can identify the set of complex structures on X with a certain set of automorphisms $TX \to TX$ and, consequently, with a certain set of smooth sections of the vector bundle $\mathrm{End}\,TX$. Now, we can equip the set of complex structures with the topology induced by some natural topology on the space of sections (e.g. the C^∞-topology). We shall denote the space of complex structures on X by S_X.

(If an automorphism $J\colon TX \to TX$ is associated with a complex structure, then $J \circ J = -\mathrm{id}$. Conversely, every automorphism $J\colon TX \to TX$ which covers id_X and satisfies $J \circ J = -\mathrm{id}$ arises from some complex structure. This is a classical result going back to Gauss and Riemann. In modern language it is usually stated as follows: every 1-dimensional almost complex structure is integrable. It is well-known that this is no longer true in higher dimensions.)

The group $\mathrm{Diff}^+(X)$ acts naturally on S_X. This natural action is the restriction of the action of $\mathrm{Diff}^+(X)$ on the space of automorphisms of TX given by the formula: $(f, J) \mapsto Tf \circ J \circ Tf^{-1}$, where $f \in \mathrm{Diff}^+(X)$, $J\colon TX \to TX$ and Tf is the tangent map of f. If we equip $\mathrm{Diff}^+(X)$ with a suitable topology (e.g. if we equip $\mathrm{Diff}^+(X)$ and S_X with C^∞-topology) then this action becomes continuous. The factor space $S_X/\mathrm{Diff}^+(X)$ is called the moduli space of complex structures on X or the moduli space of Riemann surfaces (of genus g, where g is the genus of X) and is denoted by M_X. The moduli space M_X can be regarded as the set of Riemann surfaces (i.e. 1-dimensional complex manifolds) diffeomorphic to X and considered up to isomorphism. Indeed, two complex structures on X represent the same point in M_X iff they are isomorphic: an isomorphism is an orientation-preserving diffeomorphism taking one structure into the other. We can consider not only $M_X = S_X/\mathrm{Diff}^+(X)$ but also the factor-space $T_X = S_X/\mathrm{Diff}_0(X)$ which is called the Teichmüller space of X. Two complex structures on X represent the same point in T_X iff they are isomorphic and the diffeomorphism taking one structure into the other can be chosen isotopic to id_X. It is clear that $\mathrm{Mod}_X = \mathrm{Diff}^+(X)/\mathrm{Diff}_0(X)$ acts on T_X and $T_X/\mathrm{Mod}_X = M_X$.

To understand the structure of M_X is one of the main problems in the theory of Riemann surfaces and in the theory of complex algebraic curves (which is another side of the same subject). This problem is usually called the moduli problem. It has been attracting much attention since Riemann's times up to nowadays. By introducing T_X this problem can be divided into two: the study of T_X and the study of Mod_X and its action on T_X.

T_X turns out to be homeomorphic to $\boldsymbol{R}^{6g-6}$ where g is the genus of X. The group Mod_X acts on T_X discretely. In particular, the isotropy groups of this action are finite. Points with nontrivial isotropy groups arise from complex structures on X with nontrivial automorphisms. Hence, the projection $T_X \to M_X$ is a branched covering and the branching set is the set of isomorphism classes of Riemann surfaces with symmetries. Moreover, Mod_X contains a subgroup of finitie index, which acts on T_X freely. For every such subgroup Γ the projection $T_X \to T_X/\Gamma$ is a genuine (unbranched) covering. Since T_X is contractible, we see that T_X/Γ is an Eilenberg–MacLane space and Γ is its fundamental group. For Γ suitably chosen the space T_X/Γ

can be interpreted as the moduli space of Riemann surfaces with an additional structure. E.g., if Γ is the kernel of the natural map $\mathrm{Mod}_X \to \mathrm{Aut}(H_1(X, \boldsymbol{Z}/p\boldsymbol{Z}))$ with $p \geqslant 3$ (the subgroup Γ acts freely on T_X), then this additional structure is a fixed element in $H_1(X, \boldsymbol{Z}/p\boldsymbol{Z})$. As for Mod_X, it is "almost" the fundamental group of M_X and M_X is "almost" an Eilenberg–MacLane space.

If X is a torus, then T_X can be naturally identified with the upper half-plane $H = \{z \in \boldsymbol{C}\colon \operatorname{Im} z > 0\}$. Under this identification, the action of Mod_X on T_X corresponds to the well-known action of the modular group $SL_2(\boldsymbol{Z})$ on H:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot z = \frac{az+b}{cz+d}.$$

These observations explain why the term "Teichmüller modular group" stands for the group Mod_X.

In fact, T_X and M_X have much richer structures than just the structure of a topological space. T_X is a metric space and a complex manifold. The metric structure had been introduced by Teichmüller. Teichmüller also claimed to have proved that T_X is a complex manifold, but the first complete proof was given by Ahlfors [7] (this result was already anticipated in 19-th century). Mod_X preserves both these structures and for both of them it is the full automorphism group. This remarkable theorem is due to Royden [37]. M_X is not a complex manifold, but it is an analytic space and the projection $T_X \to M_X$ is an analytic map. Moreover, M_X is, in a natural way, a quasiprojective algebraic variety over $\boldsymbol{C}$. This result is due to Mumford. However, we will not use these additional structures.

3. Cohomological dimension and finiteness properties

The title of this section is the same as that of Section 1 of the well-known paper of Serre [38]. It was chosen to remind the general theory and its applications to arithmetic groups considered in Serre's paper. Nevertheless, we do not assume that the reader is familiar with this theory and, in particular, with Serre's paper.

Let X be a closed surface. Let Γ be a subgroup of finite index in Mod_X acting freely on T_X. Then T_X/Γ is an Eilenberg–MacLane space and a topological manifold (actually, a smooth and even a complex one). This manifold is noncompact because M_X is known to be noncompact. One can pose a natural question: is T_X/Γ the interior of a compact manifold with boundary? The answer is positive, at least for some Γ. Moreover, it is possible to glue a boundary to all manifolds of the form T_X/Γ simultaneously. Namely, Harvey [18] had constructed a manifold with boundary $\bar{T}_X$, such that:

(i) T_X is the interior of $\bar{T}_X$;

(ii) the action of Mod_X on T_X extends by continuity to an action of Mod_X on $\bar{T}_X$ and this extended action is discrete;

(iii) there is a subgroup of finite index in Mod_X, which acts freely on $\bar{T}_X$;

(iv) the factorspace $\bar{T}_X/\mathrm{Mod}_X$ is compact.

Note that T_X and its boundary $\partial \bar{T}_X$ are noncompact. It follows from (iv) that $\bar{T}_X/\Gamma$ is compact for every Γ of finite index in Mod_X. A priori it may happen that T_X/Γ is a manifold and $\bar{T}_X/\Gamma$ is not. In any case, this does not take place at least for some Γ (by (iii)) and this is all we need for our applications. If $\bar{T}_X/\Gamma$ is a manifold then $\bar{T}_X/\Gamma$ is finitely triangulable. Triangulability of $\bar{T}_X/\Gamma$ follows from the fact that $\bar{T}_X$ has a natural Mod_X-invariant structure of a smooth manifold with corners (the definition of manifolds with corners can be found in [6], which is the most appropriate reference in our context). Hence, $\bar{T}_X/\Gamma$ is a smooth manifold with corners and every such manifold is triangulable (see [6]); moreover, every triangulation is finite because $\bar{T}_X/\Gamma$ is compact.

As an application, we can prove all finiteness properties of Mod_X. Let Γ be a subgroup of finite index in Mod_X, acting freely on $\bar{T}_X$. Since $\bar{T}_X/\Gamma$ is obtained by attaching a boundary to T_X/Γ, it is homotopy equivalent to T_X/Γ and hence is a $K(\Gamma, 1)$-space (because T_X/Γ is $K(\Gamma, 1)$-space, see Section 2). Moreover, $\bar{T}_X/\Gamma$ is a finite complex, as was mentioned above. Thus there is a finite complex which is a $K(\Gamma, 1)$-space. Therefore, Γ is finitely presented, the cohomological dimension of Γ is finite, and Γ is of type (FL). Recall that the cohomological dimension $\mathrm{cd}\, G$ of a group G is defined to be the supremum over all integers n such that $H^n(G, A) \neq 0$ for some G-module A. It is possible that $\mathrm{cd}\, G = \infty$. Further, G is of type (FL) if the trivial G-module $\boldsymbol{Z}$ admits a resolution of finite length consisting of finitely generated free modules. If G is of type (FL) then $\mathrm{cd}\, G < \infty$. Since Γ is finitely presented and of finite index in Mod_X, Mod_X is finitely presented, too. But since Mod_X contains elements of finite order, $\mathrm{cd}\,\mathrm{Mod}_X = \infty$ and Mod_X is not of type (FL). The fact that Mod_X contains a subgroup Γ of finite index with $\mathrm{cd}\,\Gamma < \infty$ and of type (FL) is usually expressed as follows: Mod_X virtually has finite cohomological dimension and is a group of type (VFL).

If a group G contains a subgroup of finite index of finite cohomological dimension, then the cohomological dimension is the same for all such subgroups, this common cohomological dimension being called the *virtual cohomological dimension* of G and denoted by $\mathrm{vcd}\, G$. The natural problem is to compute $\mathrm{vcd}\,\mathrm{Mod}_X$. Of course, $\mathrm{vcd}\,\mathrm{Mod}_X = \mathrm{cd}\,\Gamma$ where Γ is as above. Let g be the genus of X. It is easy to show that $3g-3 \leqslant \mathrm{vcd}\,\mathrm{Mod}_X \leqslant 6g-7$. To prove the first inequality it is sufficient to exhibit a free abelian subgroup of rank $3g-3$ in Γ (because $\mathrm{cd}\,\boldsymbol{Z}^m = m$). Such a subgroup can be generated by suitable powers of Dehn twists along $3g-3$ disjoint and pairwise nonisotopic circles on X. To prove the second inequality note that $\bar{T}_X/\Gamma$ is a $K(\Gamma, 1)$-space and $\dim \bar{T}_X/\Gamma = 6g-6$. Hence $\operatorname{cd}\Gamma \leqslant 6g-6$. Moreover, since $\partial(\bar{T}_X/\Gamma)$ is nonempty, $\bar{T}_X/\Gamma$ is homotopy equivalent to a $(6g-7)$-dimensional complex and hence $\operatorname{cd}\Gamma \leqslant 6g-7$. It had been conjectured (see [17]) that $\operatorname{vcd}\operatorname{Mod}_X = 3g-3$, but this turned out to be false.

The inequality $\operatorname{cd}\Gamma \leqslant 6g-7$ can be proved by using only T_X/Γ without recourse to $\bar{T}_X/\Gamma$. However, the above proof can be generalized and this generalization leads to the computation of $\operatorname{vcd}\operatorname{Mod}_X$. The following general theorem is essentially due to Bieri and Eckmann [3].

THEOREM 1. *Let G be a group acting freely on a contractible topological manifold V with boundary. Suppose that V/G is compact and triangulable. If ∂V is $(n-1)$-connected, then* $\operatorname{cd} G \leqslant \dim V-n-1$. *If, moreover, ∂V is homotopy equivalent to a CW-complex of dimension n (and hence is homotopy equivalent to a bouquet of n-spheres), then* $\operatorname{cd} G = \dim V-n-1$.

The above proof of $\operatorname{vcd}\operatorname{Mod}_X \leqslant 6g-7$ is in fact an application of the case $n = 0$ of Theorem 1 (a space is (-1)-connected iff it is nonempty). To compute $\operatorname{vcd}\operatorname{Mod}_X$ by means of this theorem, we must study the homotopy type of $\partial\bar{T}_X$. The first step in this direction was made by Harvey [18]. He proved that $\partial\bar{T}_X$ is homotopy equivalent to the geometric realization of a certain simplicial complex which is called the complex of curves of X and is denoted by $C(X)$. We now define it.

Recall that a simplicial complex is a set with a family of its finite subsets called its simplices. Every subset of a simplex is required to be a simplex, too. The elements of this set are called vertices of the simplicial set; two-element simplices are called edges and $(n+1)$-element simplices are called n-dimensional simplices. The vertices of $C(X)$ are isotopy classes of simple closed curves on X (for brevity called also circles) which do not bound a disk in X. If several circles on X are disjoint, pairwise nonisotopic and do not bound disks, then the set of their isotopy classes is a simplex in $C(X)$ and there are no other simplices.

THEOREM 2. *The geometric realization of $C(X)$ is homotopy equivalent to a bouquet of $(2g-2)$-spheres where g is the genus of X.*

COROLLARY. $\operatorname{vcd}\operatorname{Mod}_X = 4g-5$.

These results are due to Harer [15]. The connectedness of $C(X)$ was first proved by Harvey [18] and the 1-connectedness for $g \geqslant 2$ was first proved by the author [22]. In [22], the 1-connectedness of $C(X)$ was deduced from the 1-connectedness of a much more complicated complex introduced by Hatcher and Thurston [19]. The Hatcher–Thurston complex is 1-connected by the main result of their paper. Not much later the author found a direct method for studying $C(X)$. The first application of this method was the proof that $C(X)$ is 3-connected for $g \geqslant 3$. These results were presented at the Warsaw Congress. The potential of this method was completely realized after the author has learned from Harer that $\operatorname{vcd}\operatorname{Mod}_X = 4g-5$. As it turned out, the method provides the precise inequality $\operatorname{vcd}\operatorname{Mod}_X \leqslant 4g-5$. From Harer's preprint [15] it became clear that the complete result $\operatorname{vcd}\operatorname{Mod}_X = 4g-5$ can be deduced from $\operatorname{vcd}\operatorname{Mod}_X \leqslant 4g-5$ by a simple combinatorial argument due to Harer [15] (cf. [15], end of the proof of Theorem 3.5) and some general facts from the Bieri–Eckmann theory [3] (cf. [15], the remark after Lemma 4.2).

The idea of this quite simple method will be explained below. But before doing this we show how to prove that Mod_X is finitely presented without using Teichmüller spaces. An interesting feature of this proof is that we have to deal with nonclosed surfaces in order to demonstrate the claim for closed ones. Let Y be a compact surface and let X be the closed surface obtained by glueing disks to all boundary components of Y. Every diffeomorphism $Y \to Y$ can be extended to a diffeomorphism $X \to X$ and such an extension is unique up to isotopy. Hence we have a homomorphism $\operatorname{Mod}_Y \to \operatorname{Mod}_X$. This homomorphism is surjective and its kernel can be easily identified with the braid group of n strings on X, where n is the number of components of ∂Y. Recall that this group is defined as the fundamental group of the space of n-element subsets of X and is denoted by $B_n(X)$. Thus we have an exact sequence

$$(*) \qquad 1 \to B_n(X) \to \operatorname{Mod}_Y \to \operatorname{Mod}_X \to 1$$

LEMMA. *If a sequence of groups $1 \to \Gamma' \to \Gamma \to \Gamma'' \to 1$ is exact and Γ', Γ'' are finitely presented, then Γ is also finitely presented.*

The proof is elementary. Applying this lemma to the well-known exact sequences

$$1 \to B_n(X - \{\text{point}\}) \to B_n(X) \to \pi_1(X) \to 1$$

we obtain that the braid groups are finitely presented. Applying further this lemma to the sequence $(*)$, we obtain that if Mod_X is finitely presented then Mod_Y is also finitely presented.

Now assume that Mod_Y is finitely presented for all closed Y of genus $< g$. Then Mod_Y is finitely presented for all compact Y of genus $< g$. We shall prove that Mod_X is finitely presented for closed X of genus g. By induction, this will be sufficient for the proof that Mod_X is finitely presented for all X. Our strategy is to apply the following theorem to the action of Mod_X on $C(X)$.

THEOREM 3. *Let G be a group acting on a simplicial complex C. If*

(i) *the geometric realization of C is simply-connected (and in particular, connected);*

(ii) *the isotropy group of every vertex is finitely presented;*

(iii) *the isotropy group of every edge is finitely generated;*

(iv) *the number of orbits of 2-simplices is finite, then G is finitely presented.*

A special case of this theorem was used implicitly by Hatcher and Thurston [19] and was formulated in an explicit form in Laudenbach's report [24] on Hatcher and Thurston's paper. The proof of this special case can be easily adapted for the general case. In the above form this theorem was recently published by K. S. Brown [8]. He deduced it from the Bass–Serre theory of group actions on trees. In fact, finite presentations of isotropy groups of vertices and finite sets of generators of isotropy groups of edges (and the action of G on C) give rise to finite presentation of G itself.

Let S be a circle on X which do not bound a disk in X, and let $\mathrm{Mod}_{X,S}$ be the isotropy group of the corresponding vertex of $C(X)$. Let Y be the surface obtained by cutting X along S. Then Y consists of one or two components, having genus less than the genus of $X(=g)$. It is easy to check that there is an exact sequence

$$1 \to \boldsymbol{Z} \to \mathrm{Mod}_{X,S} \to \mathrm{Mod}_Y \to 1.$$

It follows from Lemma and our inductive assumption that $\mathrm{Mod}_{X,S}$ is finitely presented. Similarly, isotropy groups of edges are finitely presented (and not only finitely generated). Finally, it is easy to see that the number of orbits of 2-simplices is finite. This number is equal to the number of configurations of 3 disjoint circles on a surface of genus g considered up to diffeomorphism. Since $C(X)$ is simply connected for $g \geqslant 2$ by Theorem 2, it follows from Theorem 3 that Mod_X is finitely presented. Thus, we have proved the following theorem.

THEOREM 4. *All groups Mod_X are finitely presented.*

Now we sketch a proof of the $(2g-3)$-connectedness of $C(X)$. Actually, we shall prove only the connectedness of $C(X)$ for $g \geqslant 2$. The general assertion can be proved in the same manner. We have to show that every two vertices of $C(X)$ can be connected by a chain of edges. We shall denote by $\langle S \rangle$ the isotopy class of a circle S. Let $\langle S \rangle$ and $\langle S' \rangle$ be two vertices. Our aim is to find a chain of vertices $\langle S_0 \rangle$, $\langle S_1 \rangle, \ldots, \langle S_{n+1} \rangle$ such that $\langle S \rangle = \langle S_0 \rangle$, $\langle S_{n+1} \rangle = \langle S' \rangle$ and such that $\langle S_i \rangle$ is joined with $\langle S_{i+1} \rangle$ by an edge for $i = 0, 1, \ldots, n$. This means that certain disjoint circles T_i and T_i' are isotopic to S_i and S_{i+1}, respectively. To this end we choose two Morse functions $f, f'\colon X \to \boldsymbol{R}$ such that S (respectively S') is a component of a level set $f^{-1}(a)$ (respectively $f'^{-1}(a')$), and a, a' are regular values of f, f', respectively. Let $\{f_t\colon X \to \boldsymbol{R}\}_{t\in[0,1]}$ be a generic path of functions joining f with f': $f_0 = f$, $f_1 = f'$. Then for every t except for a finite set of values of t, f_t is a Morse function with distinct critical values. For exceptional t, either f_t has exactly one degenerate critical point of the form $(x, y) \mapsto x^3 \pm y^2$ (the birth-death point) or f_t has exactly one pair of Morse critical points with the same critical value. It is easy to check that every such f_t has a noncritical value a_t and a component S_t of $f_t^{-1}(a_t)$ which does not bound a disk in X. Since a_t is a regular value of f_t, there is an open interval U_t containing t such that a_t is a regular value of f_u for every $u \in U_t$, and, moreover, some component S_{tu} of $f_u^{-1}(a_t)$ is isotopic to S_t. We can cover $[0, 1]$ by a finite set $\{U_{t_1}, U_{t_2}, \ldots, U_{t_n}\}$ of intervals U_t. We can assume that $0 \in U_{t_1}$, $1 \in U_{t_n}$ and $U_{t_i} \cap U_{t_{i+1}} \neq \emptyset$ for $i = 1, \ldots, n-1$. Consider the chain $\langle S \rangle$, $\langle S_{t_1} \rangle, \ldots, \langle S_{t_n} \rangle$, $\langle S' \rangle$ of vertices. Every two consecutive vertices in this chain are joined by an edge, because if $u \in U_{t_i} \cap U_{t_{i+1}}$ then $S_{t_i u}$, $S_{t_{i+1} u}$ are isotopic to S_{t_i}, $S_{t_{i+1}}$, respectively, and the circles $S_{t_i u}$, $S_{t_{i+1} u}$ are either equal or disjoint. The latter statement follows from the fact that $S_{t_i u}$, $S_{t_{i+1} u}$ are components of level sets of the same function. Hence, we have proved the connectedness of $C(X)$.

Theorem 4 is due to McCool [32], who proved it by the methods of combinatorial group theory. Hatcher and Thurston [19] suggested a geometrical proof. Their proof is similar to the above one but they used a much more complicated complex formed of so-called cut systems on X (which are certain sets of g circles on X). In the proof of 1-connectedness of this complex the Morse-Cerf theory and the connection between circles and functions on surfaces were also used, but in somewhat disguised form. Indeed, it was the thought of Hatcher–Thurston's paper that led the author to the ideas sketched above. One of Hatcher and Thurston's goals was to write down an explicit presentation for Mod_X by generators and relations. It was difficult to get an explicit presentation from McCool's proof. An explicit finite presentation was known only for genus two, it was found by Birman and Hilden [4]. But the methods of Hatcher and Thurston led to a very complicated presentation and they had not written it down. Later Harer [14] found a smaller complex with the same properties as Hatcher–Thurston's one, and using this complex Wajnryb [41] wrote down a finite presentation for Mod_X. This presentation, and especially its proof, are very cumbersome. The proof sketched above also allows one to write down a presentation of Mod_X, at least in principle (cf. the remarks after Theorem 3), but its main advantage is simplicity.

Appendix. Cohomology of Mod_X. Our knowledge of cohomology properties of Mod_X is not exhausted by the value of $\mathrm{vcd}\,\mathrm{Mod}_X$. There are several other remarkable results. Powell [36] has computed H_1 and Harer [14] had computed H_2 in almost all cases. Namely,

$$H_1 = 0 \qquad \text{for } g \geqslant 3,$$
$$H_2 = \mathbf{Z}^{b+1} \qquad \text{for } g \geqslant 5,$$

where g is the genus of X and b is the number of boundary components of X.

Harer [16] proved several results on stabilization of homology of Mod_X. In particular, if $g \geqslant 3n+1$ then $H_n(\mathrm{Mod}_X)$ depends only on b. Morita [33] has found a lot of nontrivial elements in $H_*(\mathrm{Mod}_X)$. Using these nontrivial elements and the Bott vanishing theorem [7] he proved the following remarkable theorem.

THEOREM. *The natural homomorphism* π: $\mathrm{Diff}^+(X) \to \mathrm{Mod}_X$ *do not admit any section; that is, there is no homomorphism* i: $\mathrm{Mod}_X \to \mathrm{Diff}^+(X)$ *such that* $\pi \circ i = \mathrm{id}$.

Finally, in some papers (e.g. [9], [42]) homology groups of the moduli space of stable curves are studied. These homology groups turned out to be closely related with $H_*(\mathrm{Mod}_X)$.

4. The Thurston boundary of the Teichmüller space and the classification of elements of Mod_X

All results in the rest of the paper are based on the Thurston theory of diffeomorphisms of surfaces. This Section is devoted to a short survey of this theory. Complete (and, in fact, unique) exposition of this theory can be found in the Proceedings of Orsay Seminar [11].

A diffeomorphism f: $X \to X$ is called reducible, if $f(S) = S$ for some one-dimensional submanifold S of X such that every component of S neither bounds a disk in X nor is isotopic to a component of ∂X (components of S can be permuted by f). An element of Mod_X is called reducible if it contains reducible diffeomorphism. An element α of Mod_X is called periodic if $\alpha^n = 1$ for some $n \neq 0$. By a deep theorem of Nielsen, if $\alpha^n = 1$ then α contains a diffeomorphism f such that $f^n = \mathrm{id}$. Irreducible (that is, nonreducible) and nonperiodic elements of Mod_X are called pseudo-Anosov. These elements are the most interesting ones. Thurston has proved that these elements also contain some remarkable representatives, called pseudo-Anosov diffeomorphisms. We are not going to use them explicitly and so we do not give a definition here. But the main ingredient of the proof of the existence of these representatives, that is, the Thurston boundary of the Teichmüller space, will be of great importance for us.

As in Section 2, we shall confine ourselves to closed surfaces. If X is a closed orientable surface of genus g, then T_X is homeomorphic to an open ball of dimension $6g-6$. Thurston has constructed a compactification $\mathcal{T}_X$ of T_X which is homeomorphic to a closed ball of dimension $6g-6$. The difference $\mathcal{T}_X \setminus T_X$ is called the Thurston boundary of T_X and is denoted by $\partial\mathcal{T}_X$; it is homeomorphic to S^{6g-7}. This compactification is natural in the following sense: the action of Mod_X on T_X can be extended by continuity to $\mathcal{T}_X$. Much earlier Teichmüller had constructed a compactification of T_X which is also homeomorphic to a closed ball and which is very natural from the viewpoint of geometry of Teichmüller spaces, but recently Kerckhoff [23] showed that the action of Mod_X on T_X cannot be extended to the Teichmüller boundary. The construction of $\mathcal{T}_X$ is rather difficult. The points of $\partial\mathcal{T}_X$ are certain equivalence classes of special geometric objects, namely, singular foliations with transverse measure. $\partial\mathcal{T}_X$ is another natural geometric object on which Mod_X acts. The study of this action will help us to get new results on Mod_X.

In [11] we encounter the following description of the action of pseudo-Anosov elements on $\partial\mathcal{T}_X$. Each element of this type has exactly two fixed points in $\partial\mathcal{T}_X$. One of them is called the attracting fixed point and is denoted by A_f, the other is called the repelling fixed point and is denoted by R_f. If a point $x \in \partial\mathcal{T}_X$ is not fixed, then

$$\lim_{n \to \infty} f^n(x) = A_f, \qquad \lim_{n \to -\infty} f^n(x) = R_f.$$

We need a more precise description of the action of pseudo-Anosov elements, which is given by the following theorem. This theorem was proved independently by J. McCarthy [29], Papadopoulos [35] and the author [21] and certainly was known to Thurston.

THEOREM 5. *Let f be a pseudo-Anosov element of* Mod_X. *Let K be a compact subset and U an open subset of $\partial\mathcal{T}_X$. If $A_f \in U$ and $R_f \notin K$ then there is an integer $N > 0$ such that*

$$f^n(K) \subset U$$

for $n \geqslant N$.

The theory outlined above is applicable to surfaces with a boundary. Actually, surfaces with boundary arise in the analysis of reducible elements. If $f\colon X \to X$ is a reducible diffeomorphism and $f(S) = S$ for some S as above, then by cutting X along S we obtain a new surface Y with nonempty boundary and a new diffeomorphism $f'\colon Y \to Y$. The isotopy class of f' is again either reducible or periodic or pseudo-Anosov; components of Y have genus smaller than the genus of X. Considerations of this kind lie in the background of inductive reasoning.

Periodic elements are often a source of difficulties. Nevertheless, in many cases we can avoid these difficulties in the following way. An element of Mod_X is called *aperiodic* if either it is pseudo-Anosov or it contains a diffeomorphism f such that $f(S) = S$ for some S as above (so f is reducible) and such that $f'\colon Y \to Y$ obtained from f by cutting X along S preserves each component of Y and induces on each component either identity (up to isotopy), or a pseudo-Anosov element. Let $\Gamma_X(p)$ be the kernel of the natural map

$$\mathrm{Mod}_X \to \mathrm{Aut}(H_1(X, \mathbf{Z}/p\mathbf{Z})).$$

THEOREM 6. *If $p \geqslant 3$, then all elements of $\Gamma_X(p)$ are aperiodic.*

Note that the index of $\Gamma_X(p)$ in Mod_X is finite.

Theorem 5 can be extended to the case of aperiodic elements. This extension is too cumbersome to be presented here, but it is crucial for applications.

5. Subgroups of Mod_X

One of guiding principles in the study of Mod_X is an analogy between Mod_X and linear groups, i.e. subgroups of $GL_n(k)$ for some field k. Actually, not only certain results, but also some basic methods of the theory of linear groups can be extended to Mod_X. For example, the results of Section 3 are analogous to those of Serre [38] and Borel–Serre [6] on arithmetic groups which constitute a certain special class of linear groups over $\boldsymbol{C}$ (for more details about relations with arithmetic groups, see Section 7). The results of this section coresspond to some central results in the theory of linear groups. The analogy between Mod_X and linear groups leads to the natural question, whether the groups Mod_X are isomorphic to certain linear groups. The answer is very likely to be negative but our knowledge of Mod_X and of linear groups (!) is unsufficient to answer this question.

THEOREM 7. Mod_X *contains a subgroup of finite index without torsion.*

This theorem is due to Serre [39]. It is a consequence of Theorem 6. Indeed, $\Gamma_X(p)$ for $p \geqslant 3$ is such of subgroup. Actually, Theorem 6 is a natural strengthening of Theorem 7 in the setting of Thurston's theory. Serre used algebraic geometry in his proof of Theorem 7 (and stated it in other terms). Several topological proofs are known by now.

THEOREM 8. Mod_X *is a residually finite group; this means that, for every $g \in \mathrm{Mod}_X$, $g \neq 1$, there is a homomorphism φ: $\mathrm{Mod}_X \to G$ on a finite group G such that $\varphi(g) \neq 1$.*

This theorem is due to Grossman [13]. Her proof is based on rather involved arguments from the combinatorial group theory. A more conceptual proof was suggested in a very interesting paper of Bass and Lubotzky [2]. Unfortunately, their proof does not cover the genus 2 case. Algebraic ideas of Bass and Lubotzky can be combined with Theorem 6 and this leads to a very clear proof of Theorem 8. It is interesting to note that, unlike all other algebraic results on Mod_X, all proofs of Theorem 8 known to the author use the algebraic definition of Mod_X.

THEOREM 9. *Let g be the genus of X and let b be the number of boundary components of X. If G is a subgroup of Mod_X then either $G \cap \Gamma_X(p)$ is a free abelian group of rank $\leqslant 3g-3+b$ or $G \cap \Gamma_X(p)$ contains a free group with two generators.*

This theorem was proved by the author [21] and, independently, by McCarthy [29, 30] in a somewhat weaker form. It is an analogue of a famous Tits' theorem [40] on linear groups.

TITS' THEOREM. *Let G be a finitely generated linear group. Then either G contains a solvable subgroup of finite index, or it contains a free group with two generators.*

The key point in the proof of Tits' Theorem, and of Theorem 9 as well, is the following lemma, which is applied to an appropriate action of G.

LEMMA. *Let G be a group acting on a set P. Let U_1, U_2, V_1, $V_2 \subset P$ and $f_1, f_2 \in G$. If*

$$f_i(P \setminus U_i) \subset V_i, \quad f_i^{-1}(P \setminus V_i) \subset U_i, \quad i = 1, 2,$$

$$U_i \cap V_i = \emptyset, \quad i = 1, 2,$$

$$(U_1 \cup V_1) \cap (U_2 \cup V_2) = \emptyset, \quad P \setminus (U_1 \cup U_2 \cup V_1 \cup V_2) \neq \emptyset,$$

then f_1 and f_2 are the generators of a free subgroup of G.

This lemma goes back to Klein, who applied it to the action of Kleinian groups on CP^1. Its proof is quite easy. Let us prove, for example, that $f_1 f_2^{-1} f_1 f_2 \neq 1$. Take $x \in P \setminus (U_1 \cup U_2 \cup V_1 \cup V_2)$. Then $f_2(x) \in V_2$, $f_1 f_2(x) \in V_1$, $f_2^{-1} f_1 f_2(x) \in U_2$ and $f_1 f_2^{-1} f_1 f_2(x) \in V_1$. Therefore, $f_1 f_2^{-1} f_1 f_2(x) \neq x$ and $f_1 f_2^{-1} f_1 f_2 \neq 1$.

In the proof of Tits' Theorem, this lemma is applied to the action of subgroups of $GL_n(k)$ on the projective space kP^n. In the proof of Theorem 9, the projective space is replaced by the Thurston boundary $\partial \mathcal{T}_X$. As an illustration, we shall outline the proof of the following assertion: if f, g are pseudo-Anosov elements of Mod_X and f, g cannot be presented as powers of the same element of Mod_X, then f^m, g^n are generators of a free group for all sufficiently large m, n. It is easy to check that for such f, g all points A_f, R_f, A_g, R_g are distinct. Choose disjoint neighbourhoods V_f, U_f, V_g, U_g of the points A_f, R_f, A_g, R_g, respectively. It follows from Theorem 5 that $f^n(\partial \mathcal{T}_X \setminus U_f) \subset V_f$ for all sufficiently large n. Similarly,

$$f^{-n}(\partial \mathcal{T}_X \setminus V_f) \subset U_f, \quad g^n(\partial \mathcal{T}_X \setminus U_g) \subset V_g, \quad g^{-n}(\partial \mathcal{T}_X \setminus V_g) \subset U_g$$

for all sufficiently large n. Hence we can complete the proof of our assertion by an application of the Lemma. Thus, if G is generated by pseudo-Anosov elements, Theorem 9 is proved. But the main difficulties that arise in the proof come from reducible elements. By Theorem 6, we may confine ourselves to aperiodic elements only. These elements can be treated by means of the generalization of Theorem 5 mentioned at the end of Section 4. The argument sketched above was known to the author for a long time (but not Theorem 5). However, only after the paper of Birman, Lubotzky and McCarthy [5] it become clear that the complete analogue of Tits' Theorem is within reach. In [5] it was proved that every solvable subgroup of Mod_X contains an abelian subgroup of finite index.

Recall that a subgroup H of a group G is called maximal if $H \neq G$ and if for every subgroup K satisfying $H \subseteq K \subseteq G$, either $H = K$ or $K = G$.

THEOREM 10. *Let G be a finitely generated subgroup of* Mod_X. *Then either G contains a maximal subgroup of infinite index in G, or G contains an abelian subgroup of finite index in G.*

In particular, Mod_X contains a maximal subgroup of infinite index. For $SL_n(\boldsymbol{Z})$ this was proved quite recently by Margulis and Soifer. Theorem 10 is an analogue of the following theorem of Margulis and Soifer [28].

MARGULIS–SOIFER'S THEOREM. *Let G be a finitely generated linear group. Then either G contains a solvable subgroup of finite index in G, or G contains a maximal subgroup of infinite index in G.*

The proof of Theorem 10 follows that of Margulis and Soifer as a pattern and deals with the action of Mod_X on $\partial\mathscr{T}_X$. The idea is like this. Assume that G does not contain any abelian subgroup of finite index. Then it is possible to find an infinite set $Z \subset G$ such that: (i) Z intersects all residue classes modulo all subgroups of finite index in G; (ii) Z is the set of free generators of a free subgroup of G. Since Z is infinite, the group F generated by Z is not equal to G, which is finitely generated by assumption. Therefore, there is a maximal subgroup H of G containing F. It follows from (i) that H is of infinite index. Z being suitably chosen, one can verify (ii) by an application of the Lemma to the action of G on $\partial\mathscr{T}_X$.

This method allows us to derive some additional conclusions. For example, in Mod_X there is a maximal subgroup of infinite index which contains $\boldsymbol{Z} \times \boldsymbol{Z}$. If G is a finitely generated subgroup of Mod_X and if G contains a maximal subgroup of infinite index then G contains an uncountable set of such subgroups.

The next theorem is not a purely algebraic assertion on Mod_X. To state this theorem we need a definition. A subgroup G of Mod_X is called *reducible* if there is a 1-dimensional submanifold S of X with properties as in the definition of reducible diffeomorphisms and such that every element of G can be represented by a diffeomorphism $f\colon X \to X$ with $f(S) = S$. It is possible to study reducible subgroups in the same manner as reducible diffeomorphisms, i.e. by cutting X along S. A subgroup of Mod_X is called *irreducible* if it is not reducible.

THEOREM 11. *Every infinite irreducible subgroup of* Mod_X *contains an irreducible and nonperiodic (i.e. a pseudo-Anosov) element.*

Conversely, if a subgroup contains an irreducible element then, obviously, it is irreducible. Note that the conclusion of Theorem 11 is false for finitie subgroups. Gilman [12] has constructed finite irreducible subgroups of Mod_X consisting of reducible elements only.

6. Automorphisms of Mod_X and of related groups

In this section we consider not only Mod_X but also the extended mapping class group $\overline{\mathrm{Mod}}_X$ and its subgroups $\overline{\mathrm{Mod}}_X^{\partial}$, $\mathrm{Mod}_X^{\partial}$. By definition, $\overline{\mathrm{Mod}}_X = \mathrm{Diff}(X)/\mathrm{Diff}_0(X)$; $\overline{\mathrm{Mod}}_X^{\partial} = \mathrm{Diff}^{\partial}(X)/\mathrm{Diff}_0(X)$ and $\mathrm{Mod}_X^{\partial} = \overline{\mathrm{Mod}}_X^{\partial} \cap \mathrm{Mod}_X$ where $\mathrm{Diff}(X)$ is the group of all diffeomorphisms $X \to X$ and $\mathrm{Diff}^{\partial}(X)$ is the subgroup of diffeomorphisms fixing every component of ∂X setwise. It is clear that Mod_X is a normal subgroup of index 2 in $\overline{\mathrm{Mod}}_X$ and that $\overline{\mathrm{Mod}}_X^{\partial}$, $\mathrm{Mod}_X^{\partial}$ are normal subgroups of finite index in $\overline{\mathrm{Mod}}_X$.

THEOREM 12. *Assume that X is neither a closed surface of genus 2, nor a torus with $\leqslant 2$ holes, nor a sphere with $\leqslant 4$ holes. Then all automorphisms of the groups $\overline{\mathrm{Mod}}_X$, Mod_X, $\overline{\mathrm{Mod}}_X^{\partial}$, $\mathrm{Mod}_X^{\partial}$ have the form $\alpha \mapsto \beta\alpha\beta^{-1}$ with $\beta \in \overline{\mathrm{Mod}}_X$. In particular, all automorphisms of $\overline{\mathrm{Mod}}_X$ are inner, the groups of outer automorphisms $\mathrm{Out}(\overline{\mathrm{Mod}}_X^{\partial})$, $\mathrm{Out}(\mathrm{Mod}_X^{\partial})$, $\mathrm{Out}(\mathrm{Mod}_X)$ are finite and, moreover, $\mathrm{Out}(\mathrm{Mod}_X) = \mathbf{Z}/2\mathbf{Z}$.*

The case of closed surfaces of genus 2 was examined by McCarthy [31]. His paper contains also his version of the author's proof of Theorem 12 in the case of closed surfaces. In this case Theorem 12 was conjectured by V. Turaev.

THEOREM 13. *Let X be a closed surface of genus 2. Then $\mathrm{Out}(\overline{\mathrm{Mod}}_X) \approx \mathrm{Out}(\mathrm{Mod}_X) \approx \mathbf{Z}/2\mathbf{Z} \oplus \mathbf{Z}/2\mathbf{Z}$.*

The difference between the genus 2 case and the higher genus cases in these theorems stems from the fact that the center of Mod_X is nontrivial only if the genus of X is 2. This reflects the following fact from the theory of Riemann surfaces: all Riemann surfaces of genus 2 are hyperelliptic (it is not true for surfaces of higher genus).

The proofs of these two theorems are based on a description of pairs of commuting elements of Mod_X. Recall that the support of a diffeomorphism f is the closure of the set $\{x: f(x) \neq x\}$. It is clear that two diffeomorphisms with disjoint supports are commuting and hence their isotopy classes are commuting. More generally, if $f_1, \dots, f_n$: $X \to X$ are diffeomorphisms with disjoint supports then all diffeomorphisms of the form $f_n^{m_1} \circ \dots \circ f_n^{m_n}$ commute. It turns out that these are essentially all examples of commuting elements in Mod_X. The precise assertion is given by the following theorem.

THEOREM 14. *If $\alpha, \beta \in \Gamma_X(p)$, $p \geqslant 3$ and $\alpha\beta = \beta\alpha$ then there are diffeomorphisms $f_1, \dots, f_n$: $X \to X$ with disjoint supports and integers $m_1, \dots, m_n$, $l_1, \dots, l_n$ such that α, β are the isotopy classes of $f_1^{m_1} \circ \dots \circ f_n^{m_n}$, $f_1^{l_1} \circ \dots \circ f_n^{l_n}$, respectively.*

This theorem is a by-product of the proof of Theorem 9. It can be used for an algebraic characterization of certain elements in Mod_X of geometric origin. As an example, we give a characterization of Dehn twists along nonseparating circles. Fix some $p \geq 3$. Let $N = [\mathrm{Mod}_X: \Gamma_X(p)]$ and let Γ'_X be the group generated by the set $\{\alpha^N: \alpha \in \mathrm{Mod}_X\}$. It is clear that $\Gamma'_X \subset \Gamma_X(p)$. We denote by $C(H)$ the center of a group H and by $C_a(H)$ the centralizer of $a \in H$ in H.

THEOREM 15. *Let X be a surface as in Theorem* 12. *An element $\alpha \in \mathrm{Mod}_X$ is a Dehn twist along a nonseparating circle iff*

(i) $C(C_{\alpha^N}(\Gamma'_X)) \approx \mathbf{Z}$;

(ii) Mod_X *contains a free abelian group A of rank $3g-3+b$ generated by α and some elements conjugated with α, where g is the genus of X and b is the number of boundary components of X.*

(iii) $\alpha \neq \beta^n$ *for all β from the centralizer of A in* Mod_X *and* $n > 1$.

It follows immediately from this theorem that every automorphism of Mod_X takes Dehn twists along non-separating circles into themselves. There are standard systems of generators of Mod_X consisting of such Dehn twists. These systems have been constructed by Lickorish [25]. It turns out that every automorphism of Mod_X takes every standard system into another standard system. In other words, every automorphism acts on standard generators as conjugation by the isotopy class of some diffeomorphism. The closed case of Theorem 12 follows immediately from this fact. The proof in the non-closed case is based on similar ideas, although it is much more complicated. In this case, it is another set of elements of geometric origin that plays the decisive role. This set is the image of the natural homomorphism of a braid group into Mod_X:

$$B_b(Y) \to \mathrm{Mod}_X.$$

Here b is the number of boundary components of X and Y is the result of glueing disks to all boundary components of X (for the definition of $B_b(Y)$ see Section 3). Surprisingly, this homomorphism can be used also in another direction, namely to get a description of automorphisms of $B_b(Y)$. If the genus of X is ≥ 2 then $B_b(Y) \to \mathrm{Mod}_X$ is a monomorphism and its image is a normal subgroup of Mod_X. In this case we shall identify $B_b(Y)$ with its image in Mod_X.

THEOREM 16. *If Y has genus ≥ 2, then all automorphisms of $B_b(Y)$ have the form $\alpha \mapsto \beta\alpha\beta^{-1}$ with $\beta \in \mathrm{Mod}_X$.*

The proof is based on a description of pairs of commuting elements in $B_b(Y)$. It is deduced from the description of pairs of commuting elements in Mod_X. This proof is fairly surprising because braid groups are considered usually as objects of much simpler nature than mapping class groups. In fact, braid groups have been intended to serve as a tool in mapping class group theory, but not conversely.

Like Theorems 7–10, Theorems 12 and 13 can be considered as analogues of certain results on linear groups, now concerning arithmetic groups. For instance, it is well known that all automorphisms of $GL_n(\boldsymbol{Z})$ are inner. More generally, it follows from Mostow's rigidity theorem [34] that $\mathrm{Out}(\Gamma)$ is finite for almost all arithmetic groups Γ (the subject of our theory is an exception: the fundamental groups of surfaces are arithmetic and their outer automorphism groups are exactly the mapping class groups of surfaces. (Cf. Section 1). If Γ' is a subgroup of finite index in an arithmetic group Γ, then Γ' is also an arithmetic group (simply by the definition of arithmetic groups). Moreover, if the theorem on the finiteness of the outer automorphism group applies to Γ, then it applies also to Γ'. By analogy, this leads to the following conjecture:

If Γ is a subgroup of finite index in Mod_X then $\mathrm{Out}(\Gamma)$ is finite.

7. Mod_X and arithmetic groups

The above results indicate a strong analogy between Mod_X and linear or arithmetic groups. This leads to a natural question, whether Mod_X is an arithmetic or at least a linear group. Actually, the question of arithmeticity of Mod_X is much older than the above results. (See Harvey [17].) The answer to this question turns out to be negative. (We do not give the definition of arithmetic groups, which is not used explicitly in our discussion. It is enough to know that arithmetic groups are a natural generalization of classical groups over $\boldsymbol{Z}$, such as $SL_n(\boldsymbol{Z})$, $Sp_{2n}(\boldsymbol{Z})$ etc.) No answer to the question about linearity is known. (See the beginning of Section 5.)

THEOREM 17. *If X is not a torus or a torus with one hole, then* Mod_X *is not isomorphic to an arithmetic group.*

If X is a torus or a torus with one hole, then Mod_X *is isomorphic to* $SL_2(\boldsymbol{Z})$, *and hence* Mod_X *is an arithmetic group.*

If Γ is an arithmetic group, then either (i) Γ contains a solvable normal subgroup, or (ii) Γ does not contain such a subgroup. In the case (i) we can check that $\Gamma \not\approx \mathrm{Mod}_X$, applying Theorems 9 and 14. (By Theorem 9 every solvable subgroup contains an abelian subgroup of finite index, and by Theorem 14 such a subgroup cannot be normal.) The case (ii) is divided into two rather different subcases: groups of rank 1 and groups of rank $\geqslant 2$. The case of groups of rank $\geqslant 2$ is treated by means of the following deep Margulis' theorem [26].

MARGULIS' THEOREM. *Let Γ be an arithmetic group of rank $\geqslant 2$ and let $\varphi\colon \Gamma \to H$ be a homomorphism. Then either* $\mathrm{Ker}\,\varphi$ *lies in the center of Γ, or* $\mathrm{Im}\,\varphi$ *is finite.*

There is a natural homomorphism $\mathrm{Mod}_X \to \mathrm{Aut}\, H_1(X, \boldsymbol{Z})$. It is well known that both the image and the kernel of this homomorphism are infinite. Since the center of Mod_X is trivial (this follows, e.g., from Theorem 14), it follows that Mod_X cannot be isomorphic to an arithmetic group of rank 2.

As to arithmetic groups of rank 1, all of them are isomorphic to the fundamental groups of complete Riemannian manifolds of finite volume of negative curvature pinched between two negative constants, that is, $C_1 \leqslant K(\sigma) \leqslant C_2$ for some $C_1 \leqslant C_2 < 0$, where K is the curvature function. Hence, this case can be treated by means of the following theorem.

THEOREM 18. *Suppose X is not a torus or a torus with a hole and M is a complete Riemannian manifold of finite volume of negative curvature pinched between two negative constants. Then* Mod_X *is not isomorphic to* $\pi_1(M)$.

Under the conditions of Theorem 18, $\pi_1(M)$ cannot contain a subgroup isomorphic to $(\boldsymbol{Z} * \boldsymbol{Z}) \times \boldsymbol{Z}$. This can be deduced from the Eberlein–O'Neill theory [10]. On the other hand Mod_X contains a subgroup isomorphic to $(\boldsymbol{Z} * \boldsymbol{Z}) \times \boldsymbol{Z}$; hence the theorem.

The main point of the proof of Theorem 17 is the existence of an infinite normal subgroup of infinite index in Mod_X. We have exhibited only one subgroup of this kind, the kernel of the homomorphism $\mathrm{Mod}_X \to \mathrm{Aut}\, H_1(X, \boldsymbol{Z})$. The image of this homomorphism is an arithmetic group. If g is the genus of X, then this image is isomorphic to $Sp_{2g}(\boldsymbol{Z})$. Thus the next question is, whether the kernel of this homomorphism is an arithmetic group. If the answer were positive, Mod_X would be composed of two arithmetic groups. But the answer turns out to be negative.

THEOREM 19. *Let X be as in Theorem 17 and 18. Then no normal subgroup of* Mod_X *can be isomorphic to an arithmetic group.*

There are two proofs of this theorem. The first one is more algebraic in nature and is based on some additional results on arithmetic groups. For this proof the normality assumption is vital. The second one is based rather on the techniques of the theory of arithmetic groups than on its results. This proof extends a certain part of this techniques to Mod_X and so it exhibits once again the analogies between Mod_X and arithmetic groups. Actually, this proof was found earlier. It leads to the conjecture that the normality assumption in Theorem 19 is superfluous. Of course, there are some arithmetic subgroups in Mod_X, e.g. free subgroups. The right conjecture is as follows.

If X is as in Theorems 17–19, then no subgroup of Mod_X can be isomorphic to an arithmetic group of rank $\geqslant 2$.

There is certain evidence in favour of this conjecture. Every arithmetic group lies naturally in a Lie group. By the Margulis superrigidity theorem, every homomorphism from one arithmetic group into another is induced, possibly after replacing the first group by some of its subgroups of finite index, by some homomorphism of the ambient Lie group (there are some exceptions, including the fundamental groups of surfaces). Since there is no natural Lie group containing Mod_X, we expect that there are no homomorphisms from arithmetic groups to Mod_X (if the analogy between Mod_X and arithmetic groups goes sufficiently far). A natural candidate for such a Lie group is the group of all (metric or complex) automorphisms of the Teichmüller space, but, surprizingly, this group is equal to Mod_X by Royden's theorem [37]. (The situation is more subtle than sketched above, because there are homomorphisms from Mod_X to arithmetic groups, e.g. $\mathrm{Mod}_X \to \mathrm{Aut}\, H_1(X, \mathbf{Z})$.)

References

[1] L. V. Ahlfors, *The complex analytic structure of the space of closed Riemann surfaces*, In: *Analytic functions*, Princeton, 1960, 45-66.

[2] H. Bass and A. Lubotzky, Automorphisms of groups and of schemes of finite type, Israel J. Math. 44 (1983), 1-22.

[3] R. Bieri and B. Eckmann, *Groups with homological duality generalizing Poincaré Duality*, Invent. Math. 20 (1973), 103-124.

[4] J. Birman and H. Hilden, *On mapping class groups of closed surfaces as covering spaces*, Ann. of Math. Studies 66 (1971), 81-115.

[5] J. Birman, A. Lubotzky and J. McCarthy, *Abelian and solvable subgroups of the mapping class group*, Duke Math. J. 50 (1983), 1107-1120.

[6] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Comment. Math. Helv. 48 (1973), 436-491.

[7] R. Bott, *On a topological obstruction to integrability*, Actes Congr. Int. Math. 1, 1970, Paris (1971), 27-36.

[8] K. S. Brown, *Presentations for groups acting on simply-connected complexes*, J. Pure Appl. Algebra 32 (1984), 1-10.

[9] R. Charney and R. Lee, *Moduli space of stable curves from a homotopy viewpoint*, Preprint, 1983.

[10] P. Eberlein and B. O'Neill, *Visibility manifolds*, Pacific J. Math. 46 (1973), 45-109.

[11] A. Fathi, F. Laudenbach, and V. Poenaru, *Travaux de Thurston sur les surfaces*, Sèminaire Orsay, Asterisque 66-67, (1979).

[12] J. Gilman, *Structures of elliptic irreducible subgroups of the modular group*, Proc. London Math. Soc. 47, (1983), 27-42.

[13] E. Grossman, *On the residual finiteness of certain mapping class groups*, J. London Math. Soc. 9 (1974), 160-164.

[14] J. Harer, *The second homology group of the mapping class group of an orientable surface*, Invent. Math. v. 72, N 2 (1983), 221-239.

[15] —, *The virtual cohomological dimension of the mapping class group of an orientable surface*, preprint, 1984.

[16] —, *Stability of the homology of the mapping class groups of orientable surface*, preprint, 1984.

[17] W. Harvey, *Geometric structure of surface mapping class groups*, LMS Lect. Notes Series 36 (1979) 255-269.

[18] —, *Boundary structure of the modular group*, Ann. of Math. Studies 97 (1981) 245-251.

[19] A. Hatcher and W. Thurston, *A presentation for the mapping class group of a closed orientable surface*, Topology 19 (1980), 221-237.

[20] N. V. Ivanov, *Free subgroups of mapping class groups*, (in Russian), Short Communications (Abstracts) of ICM Warszawa 1983, IV, p. 21.

[21] —, *Algebraic properties of the Teichmüller modular group*, Soviet Math. Dokl. 29 (1984), 288–291.

[22] —, *On the virtual cohomology dimension of the Teichmüller modular group*, Lecture Notes in Math. 1060 (1984), 306–318.

[23] S. P. Kerckoff, *The asymptotic geometry of Teichmüller space*, Topology 19 (1980), 23–41.

[24] F. Laudenbach, *Présentation du groupe de difféotopies d'une surface compacte orientable*, Exposé 15 dans [11], Astérisque 66–67 (1979), 267–282.

[25] W. B. R. Lickorish, *A finite set of generators for the homeotopy group of a 2-manifold*, Proc. Cambridge Ph. Soc. 60 (1964), 769–778.

[26] G. A. Margulis, *Finiteness of factorgroups of discrete groups* (in Russian), Funktsional. Anal. i Prilozhen 13 (3) (1979), 28–39.

[27] —, *Arithmeticity of the irreducible lattices in the semisimple groups of rank greater then* 1, Invent. Math. 76 (1984), 93–120.

[28] G. A. Margulis and G. A. Soifer, *Maximal subgroups of infinite index in finitely generated linear groups*, J. of Algebra 69, (1981), 1–23.

[29] J. D. McCarthy, *Subgroups of surface mapping class groups*, Ph. D. Thesis, Columbia Univ., 1983.

[30] —, *A "Tits alternative" for subgroups of surface mapping class groups*, preprint, 1984.

[31] —, *Automorphisms of surface mapping class groups – a recent theorem of N. Ivanov*, preprint, 1984.

[32] J. McCool, *Some finitely presented subgroups of the automorphism group of a free group*, J. Algebra 35 (1975), 205–213.

[33] S. Morita, *Characteristic classes of surface bundles*, preprint, 1983.

[34] G. D. Mostow, *Strong rigidity of locally symmetric spaces*, Ann. of Math. Stud. 78 (1973).

[35] A. Papadopoulos, *Réseaux ferroviaires, difféomorphismes pseudo-Anosov et automorphismes symplectiques de l'homologie d'une surface*, Publications Orsay, 1983.

[36] J. Powell, *Two theorems on the mapping class group of a surface*, Proc. Amer. Math. Soc. 68 (1978), 347–350.

[37] H. L. Royden, *Automorphisms and isometries of Teichmüller space*, Ann. of Math. Stud. 66 (1971), 369–383.

[38] J.-P. Serre, *Cohomologie des groupes discrets*, Ann. of Math. Stud. 70 (1971), 77–169.

[39] —, *Rigidité de foncteur d'Jacobi d'echelon $n \geqslant 3$*, Sem. H. Cartan 1960/1961, Appendix to exposé 17.

[40] J. Tits, *Free subgroups of linear groups*, J. Algebra 20 (1972), 250–270.

[41] B. Wajnryb, *A simple presentation for the mapping class group of an orientable surface*, Israel J. Math. 45 (1983), 157–174.

[42] S. Wolpert, *On the homology of the moduli space of stable curves*, Ann. of Math. 118 (1983), 491–523.

Presented to the Topology Semester
April 3 – June 29, 1984