
1. DEFINITION

Let $S$ be a closed orientable surface. The mapping class group of $S$ is the group of isotopy (equivalently, homotopy) classes of orientation preserving diffeomorphisms (equivalently, homeomorphisms) of $S$. It is denoted $\text{Mod}(S)$ (or sometimes $\text{MCG}(S)$) to evoke analogies with the classical modular group $\text{PSL}_2(\mathbb{Z})$, which is closely related to the mapping class group $\text{Mod}(T^2) \cong \text{SL}_2(\mathbb{Z})$ of the torus. In other words

$$\text{Mod}(S) = \text{Diff}^{+}(S)/\text{Diff}_0(S)$$

is the quotient of the group of orientation preserving diffeomorphisms by the component of the identity. More generally, one allows $S$ to have boundary and punctures (distinguished points) and requires all diffeomorphisms and isotopies to fix these pointwise.

In part, the beauty and the richness of the subject comes from its pervasiveness in modern mathematics. Mapping class groups record monodromies of families of curves in algebraic geometry, classify surface bundles, and hold keys to the understanding of symplectic 4-manifolds and hyperbolic 3-manifolds.

2. EARLY HISTORY: DEHN AND NIELSEN

Max Dehn and his (co)student Jakob Nielsen were the early pioneers in the study of mapping class groups in the 1920s and 1930s (see [18, 67, 68]). In retrospect, their work was nothing short of seminal. Here is a brief description of their accomplishments.

A Dehn twist is a homeomorphism $T_a$ of $S$ supported on a regular neighborhood $A$ of a simple closed curve $a$ in $S$ defined as follows. Fix a homeomorphism $A \cong S^1 \times [0,1]$, and define $T_a(z,t) = (ze^{2\pi it},t)$ (and $T_a = \text{id}$ outside $A$). Dehn and Nielsen showed the following (for simplicity, assume $S$ is closed of genus $g > 1$):

Dehn proved that a finite number of Dehn twists in nonseparating curves generate $\text{Mod}(S)$. The proof was later greatly simplified by Lickorish [53] (and Humphries [40] found a minimal generating set consisting of Dehn twists). Dehn points out the lantern relation among seven Dehn twists on the sphere with four holes. He also shows that the natural homomorphism $\text{Mod}(S) \to \text{Sp}_{2g}(\mathbb{Z})$ induced by looking at the action in homology is not injective (the kernel is now known as the Torelli group of $S$).

Dehn introduced coordinates on the space of sets of isotopy classes of pairwise disjoint essential simple closed curves. Start with a maximal such collection $\mathcal{P}$ so that complementary components are spheres with three holes (or pairs of pants) and for each curve system record the intersection number and the twist with each curve in $\mathcal{P}$, after minimizing the number of intersection points. Dehn proves the

---

2010 Mathematics Subject Classification. Primary 54C40, 14E20; Secondary 46E25, 20C20.
statements about the generating set by looking at the action of Mod(S) on the set of curve systems in these coordinates.

Nielsen, attributing a part of the argument to Dehn, proves that Mod±(S) ∼= Out(π1(S)), where Mod±(S) is an index 2 extension allowing orientation reversing homeomorphisms. The hard part is to show that a homotopy equivalence h : S → S is homotopic to a homeomorphism. The proof runs along the following lines: Fix a hyperbolic metric on S. This identifies the universal cover ˜S with the hyperbolic plane ℍ2. A lift ˜h : ℍ2 → ℍ2 of h extends to a homeomorphism of the circle at infinity, and to every closed geodesic α in S one can associate the closed geodesic h∗(α) obtained by projecting the geodesic line connecting the ˜h-images of the endpoints of a lift of α. Now apply this construction to two disjoint collections α1, . . . , αk and β1, . . . , βm that cut the surface into polygons, and argue that the image curves h∗(αi) and h∗(βj) cut the surface in the same combinatorial way. Using these decompositions of S as a guide, find a homeomorphism homotopic to h. More topological arguments were given later by Seifert [71] and Mangler [55]. (Seifert proves the stronger statement that any map f : S → S′ between closed orientable surfaces with χ(S′) < 0 and with |χ(S)| = |deg(f)||χ(S′)| is homotopic to a covering map.) For the injectivity of Mod±(S) → Out(π1(S)), one needs to know that homotopic homeomorphisms are isotopic, which is a theorem of Baer [3] (for a more modern treatment, see [22]), and the statement that Mod±(S) ∼= Out(π1(S)) has become known as the Dehn–Nielsen–Baer theorem.

Nielsen analyzed the dynamics of the homeomorphism of the circle at infinity induced by a lift of a homeomorphism of a hyperbolic surface. He came tantalizingly close to discovering the classification of mapping classes, due to Thurston in the late 1970s (for a detailed discussion of Nielsen’s work see [30][63]). It should also be added that this study was the motivation for Nielsen’s work on fixed point theory.

3. Subsequent years: presentation, cohomology

Dehn and Nielsen never considered the action of Mod(S) on Teichmüller space T(S). This is the space of hyperbolic (or equivalently complex) structures on S modulo isotopy. More precisely, it is the space of pairs (g, X) where X is a hyperbolic surface and g : S → X is an orientation preserving homeomorphism, with the equivalence (g, X) ∼ (g′, X′) if there is an isometry F : X → X′ such that Fg is isotopic to g′. The homeomorphism g is called a marking.

Teichmüller space was first considered globally by Fricke and Klein [28], who viewed it as the space of hyperbolic polygons canonically assigned as fundamental domains to the universal cover of the hyperbolic surface. Teichmüller [23], from the point of view of complex structures, gave the modern definition of T(S) in terms of markings from which it is clear that Mod(S) (sometimes called the Teichmüller modular group) acts, proved that T(S) is topologically an open ball, introduced a metric (the Teichmüller metric), and showed that any two points are joined by a unique geodesic. Fenchel and Nielsen [27], from the hyperbolic geometry point of view, introduced coordinates on T(S) associated to a pair of pants decomposition. For a modern introduction to Teichmüller theory, see [39] (and its sequel). Fricke had proved earlier that Teichmüller space is an open ball in [28]; see [49].

For example, if f is a homeomorphism of S and fp is isotopic to the identity with p prime, then f acts on T(S) as a homeomorphism of order p. So by the elementary fixed point theory, f fixes a point of T(S). This easily implies that
$f$ is isotopic to an isometry $g$ of a suitable hyperbolic metric on $S$; in particular to a homeomorphism of order $p$. The same statement holds when $p$ is not prime, but one has to argue that the fixed point sets in $T(S)$ are contractible (and this follows from Teichmüller’s work, since the fixed set is convex, or alternatively, the fixed point set is the Teichmüller space of the quotient orbifold). Fenchel gave this argument in [25, 26]. Nielsen claimed the result using different methods, but in some cases his argument was incorrect. The conjecture that any finite subgroup of $\text{Mod}(S)$ fixes a point in $T(S)$ became known as the Nielsen realization problem.

Earle and Eells used Teichmüller space and the theory developed by Ahlfors and others [1, 2] to prove that the identity component $\text{Diff}_0(S)$ of the diffeomorphism group of $S$ is contractible [20, 21]. Briefly, they argue that the map $M(S) \to T(S)$ from the space of complex structures $M(S)$ is a principal $\text{Diff}_0(S)$-bundle. Since $M(S)$ and $T(S)$ are contractible, the statement follows.

A consequence of the Earle–Eells theorem is that surface bundles $S \hookrightarrow E \to X$ are classified by conjugacy classes of homomorphisms $\pi_1(X) \to \text{Mod}(S)$. The search for characteristic classes for such bundles naturally led to an attempt to understand the cohomology of $\text{Mod}(S)$.

A finite presentation of $\text{Mod}(S)$ would at least guarantee that $H^2(\text{Mod}(S); \mathbb{Z})$ is finitely generated. Besides, investigating whether given groups are finitely generated or presented was a natural question even much earlier (e.g., Nielsen found a finite presentation of $\text{Aut}(F_n)$). This was first established by McCool [60] by an argument later clarified and extended by Culler and Vogtmann [17], whose version we will now outline. They define Outer space $X_n$, an analog of the Teichmüller space for the group $\text{Out}(F_n)$. A point of $X_n$ is represented by a marked metric graph with volume 1. There is a cocompact spine $K_n \subset X_n$ which can be represented as the union of compact contractible subcomplexes (stars of roses), one for every marking of the wedge of $n$ circles. They prove contractibility of $X_n$ and $K_n$ by carefully ordering the set of all stars of roses $S_1, S_2, \ldots$ so that $S_k \cap (S_1 \cup S_2 \cup \cdots \cup S_{k-1})$ is contractible for $k > 1$. The same proof shows that if $\alpha$ is a fixed conjugacy class in $F_n$ and one restricts to the stars of roses $S_i$ for which $\alpha$ has the minimal possible length, then this union is also contractible and the stabilizer of $\alpha$ acts on it cocompactly. Now using the relative version of the Dehn–Nielsen–Baer theorem, one knows that the mapping class group of a once punctured surface $S$ can be identified (up to index 2) with the subgroup of $\text{Out}(\pi_1(S-\{p\}))$ that fixes the conjugacy class corresponding to a loop around $p$. This argument shows that the mapping class group of a punctured surface virtually has finite classifying space. (McCool’s argument constructs essentially the 2-skeleton of the Culler–Vogtmann complex.)

To prove the same fact about mapping class groups of closed surfaces of genus $> 1$ (or at least the existence of a finite presentation), one can use the Birman exact sequence whose proof uses Earle–Eells (see [13]):

$$1 \to \pi_1(S) \to \text{Mod}(S, p) \to \text{Mod}(S) \to 1.$$  

Ironically, the fact that $\text{Mod}(S)$ virtually has a finite classifying space could have been deduced earlier from the work of Deligne and Mumford [19]. They showed that the quotient $T(S)/\text{Mod}(S)$ (the moduli space of curves) can be compactified to a projective variety, and in particular it is homeomorphic to a finite simplicial complex with a subcomplex removed (one has to be careful with the orbifold structure, or alternatively, do it with a finite index subgroup). This implies that $\text{Mod}(S)$ acts cocompactly on a contractible subset of $T(S)$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Dennis Johnson wrote a remarkable sequence of papers on the Torelli groups [45–48]. Building on the work of Powell [69], he shows that the Torelli group of a closed (or once punctured) surface of genus \( \geq 3 \) is finitely generated and computes the abelianization. In genus 2 Torelli group is free of infinite rank [61]. More recently, the virtual cohomological dimension has been computed and has been shown that in the top dimension, cohomology has infinite rank [12]. Perhaps the biggest classical open problem about mapping class groups is whether Torelli groups (of high genus) are finitely presented.

4. Thurston and beyond

Modern study of mapping class groups began with the work of Thurston in the 1970s. He reintroduced the point of view of hyperbolic geometry and proved a classification theorem [74] for mapping classes that Nielsen was so close to establishing. Take a typical mapping class \( f \) on a hyperbolic surface \( S \) and a simple closed geodesic \( \alpha \). Let \( \alpha_n \) be the geodesic homotopic to \( f^n(\alpha) \). Thurston observes that the sequence \( \alpha_n \) (unless periodic) converges (for \( n \rightarrow \infty \) and also for \( n \rightarrow -\infty \)) to a (measured) lamination \( \Lambda \), i.e., a closed subset of \( S \) which is a union of pairwise disjoint simple geodesics. If \( f \) is “typical” (i.e., a pseudo-Anosov mapping class), the two laminations fill the surface and one finds a “canonical form” for \( f \)—a homeomorphism \( g \) isotopic to \( f \) so that in the complement of a finite singular set in suitable coordinates \( g \) looks like \( \begin{pmatrix} 0 & 0 \\ \lambda^{-1} & \lambda \end{pmatrix} \) for a certain \( \lambda > 1 \). Combinatorially, a pseudo-Anosov mapping class can be efficiently represented via a train track map. In general, Thurston finds a canonical \( f \)-invariant curve system on \( S \) so that on each complementary component the first return map is either periodic or pseudo-Anosov.

Formally, Thurston first compactifies Teichmüller space to a closed ball, with a point on the boundary represented by a measured lamination, and then uses Brouwer’s fixed point theorem to find either an \( f \)-invariant measured lamination or an \( f \)-invariant hyperbolic structure. For proofs along these lines, see [24] and [16]. A different, more combinatorial and algorithmic argument directly producing train track maps can be found in [10].

Upon learning of Thurston’s theorem, Bers [9] promptly produced a proof entirely using Teichmüller space and its finer, geometric structure. It is fair to say that from this point on, the two subjects of mapping class groups and of Teichmüller theory merged into a single, much richer subject.

More spectacular results quickly followed. Kerckhoff [50] solved the Nielsen realization problem by showing convexity properties of length functions in Teichmüller space. A strong form of the Tits alternative was established in [15, 59] and a classification of subgroups by Ivanov in [43].

Hatcher and Thurston constructed an explicit simply connected 2-complex on which \( \text{Mod}(S) \) acts freely and from this an explicit presentation can be found [36]. Even though their presentation was complicated, a remarkable feature was that all relators were supported on a subsurface of genus \( \leq 2 \). Subsequently, the presentation was substantially simplified by Wajnryb [75] (see also [52]).

In the meantime, there were important developments in Teichmüller theory. Masur [58] identified the natural augmentation of Teichmüller space by nodal curves with the metric completion with respect to the Weil–Petersson metric, and showed that the quotient space by \( \text{Mod}(S) \) can be identified with the Deligne–Mumford
compactification of moduli space. Harvey [35], inspired by the work of Borel and Serre on the cohomology of arithmetic groups, defined a simplicial complex, analogous to the Bruhat–Tits building, that keeps track of the strata at infinity in the augmentation. This complex, the curve complex, has become indispensable in the study of Mod(S). Its vertices are isotopy classes of essential nonperipheral simple closed curves in S, and a collection bounds a simplex if it can be realized by disjoint curves.

Harer [31–34] laid the foundations in the study of the homology of mapping class groups. He computed the virtual cohomological dimension and proved homological stability, e.g., the kth homology is independent of the genus as long as the genus is large. A significant topological ingredient in Harer’s work is his proof that the curve complex is homotopy equivalent to the wedge of spheres of a suitable dimension, which leads to Harer’s theorem that mapping class groups are virtual duality groups. Ivanov [41] has simplified and clarified some of Harer’s arguments. For example, Ivanov [42] showed that the thick part of Teichmüller space is a contractible manifold with corners and its boundary is homotopy equivalent to the curve complex. For more details, see the extensive survey [44].

Mumford and Morita defined classes \( \kappa_i \in H^2_i(\text{Mod}(S); \mathbb{Z}) \) [64–66], and Miller [62] and Morita show that the polynomial algebra on the \( \kappa_i \)'s (now known as Miller–Mumford–Morita classes) injects in the stable range to \( H^*(\text{Mod}(S); \mathbb{Z}) \). Mumford conjectured that rationally and stably the cohomology of Mod(S) is the polynomial algebra on the \( \kappa_i \)'s. Mumford’s conjecture was proved by Madsen and Weiss [54]. The study of the geometry of mapping class groups began with the seminal work of Masur and Minsky [56, 57]. Their work starts by showing that the curve complex, viewed as a metric space, is \( \delta \)-hyperbolic and that pseudo-Anosov mapping classes act on it as hyperbolic elements. (The proof of hyperbolicity of the curve complex has recently been dramatically simplified; see [38].) They go on to estimate distances in Mod(S) in terms of distances in the curve complexes of S and its subsurfaces. A way to succinctly phrase their result was given in [11]: Mod(S) acts with quasi-isometrically embedded orbits on a finite product of \( \delta \)-hyperbolic spaces; each of the hyperbolic spaces is quasi-isometric to a tree of curve complexes of subsurfaces of S. The Masur–Minsky theory led to geometric results about Mod(S): quasi-isometric rigidity [7], rapid decay property [8], measure rigidity [51], boundary amenability [29], finiteness of asymptotic dimension [11], the structure of asymptotic cones [5] [11], bounds on the conjugacy problem [72], and others, and it also led to a qualitative understanding of the geometry of Teichmüller space [70].

The reader can look at [23] for a recent list of open problems about mapping class groups. It is a testimony to the rapid development of the field that many of the problems on this list are now solved; however, many more remain open!

5. The book

The book arose from a graduate course given by the first author. Its goal is to introduce a motivated reader into this beautiful subject, describe its main results and examples, and cover the most important techniques.

The book starts by giving on page 1 the definitions of a mapping class group and of Teichmüller space. The two objects are really on an equal footing throughout the book.
Part 1 covers the classical aspects of the subject. It starts with a quick review of hyperbolic geometry, but the reader is expected to have a working knowledge of it. It proceeds with basic facts about isotopy of curves to minimal position and the discussion of the (geometric) intersection number between simple closed curves. Perhaps the main theorem in this part is that mapping class groups are generated by finitely many Dehn twists. They give a modern proof of this fact, by studying the action of $\text{Mod}(S)$ on the suitable version of the curve complex whose vertex stabilizers are mapping class groups of lower complexity, hence finitely generated by induction. The inductive process uses the Birman exact sequence, which is explained in detail. They give a similar argument for finite presentation, by examining the action on the arc complex, whose contractibility they establish using the Hatcher flow [37]. There are additional sections on lantern and chain relations, on the Torelli group, and the Johnson homomorphism, on the $84(g-1)$ theorem, the Dehn–Nielsen–Baer theorem, and on the braid groups (the last one surely inspired by Birman’s monograph [14]).

Part 2 is entirely dedicated to Teichmüller theory. The first section covers the basics, including the Fenchel–Nielsen coordinates, thus showing $\mathcal{T}(S)$ is an open ball. The second section explains the work of Teichmüller. It is well motivated after recalling the work of Grötzsch on conformal distortion of maps between rectangles. The last section is a discussion of moduli space, including Mumford’s compactness criterion.

Part 3 is on the Nielsen–Thurston classification. They present both the original Thurston proof and the Bers proof. They also give five different constructions of pseudo-Anosov mapping classes and discuss their dynamics.

The book displays a beautiful blend of topology, geometry, and analysis involved in the modern study of mapping class groups. The authors have found the right balance between too many and too few details. The theorems come with proofs, with technical parts omitted with references. Speaking from a personal experience as someone running reading courses, graduate students find the book easy to read with only occasional places that need a clarification. Many carefully drawn pictures greatly help the reader understand the arguments.

The book (understandably) does not cover the whole subject; most notably the Masur–Minsky theory is omitted. However, a student who completes the book should be ready to start reading current research papers. The book has become the standard introductory text, and anyone with a research interest in mapping class groups should have a copy of the book on the shelf.

REFERENCES


Mladen Bestvina
University of Utah