Generating the Mapping Class Group: A Geometric, Algebraic, and Historical Survey

Mrinalini Sisodia Wadhwa

March 31, 2024

Contents

Acknowledgements

This thesis marks the end of what has been one of the most enriching—and unexpected—four year-journeys of studying mathematics at Columbia. My first thanks are to my advisor, Professor Francesco Lin. I credit Professor Lin's guidance and support with my decision to pursue this topic, and am especially grateful to him for his patience with answering all of my questions that came up along the way, and for his encouragement to pursue the intersections of history and mathematics in the epilogue of this project.

My next thanks is to Professor Joan Birman, who gave generously of her time to meet with me earlier this Spring to speak about her reflections on mathematical research and discovery. The beauty of her mathematics—which the reader will see in the Birman exact sequence later in the thesis—and her warmth and humility have left an indelible impression on me, and she is, and will remain, an inspiration to me and to so many others.

I came to Columbia not knowing the first thing about topology, and the fact that I pursued a thesis in the first place is thanks to a number of wonderful instructors I have had during my time in the math major. A special thanks to Professor Mikhail Khovanov, who was the first to encourage me to pursue intersections of mathematics and history and read in hyperbolic geometry; Professor Abhijit Champarnerkar, whose Modern Analysis section introduced me to the beginnings of topology and drawing pictures as a sophomore; to Professor Dusa McDuff, whose manifolds class was my first time writing a mathematics paper; and to Matthew Hase-Liu, for his wonderful section of the undergraduate mathematics seminar. Thank you also to Professor Alma Steingart in the History Department for her guidance on the history of twentieth-century algebraic topology, which made me begin to realize the historical significance of the mathematical questions explored in this thesis.

My time here has also been enriched by a number of brilliant people I am lucky to count as my friends. I don't think I would have made it through Honors Math, Differentiable Manifolds, writing two senior theses, and so much more without the camaraderie of Aiden Sagerman; nor Algebraic Topology without the warmth and friendship of Karina Dovgodko and Tony Xiao, among others. Outside of math, a special thanks to Luiza Vilanova, Suan Lee, Bryley Williams, and to the many others I have been lucky to meet through the Laidlaw Scholars Program, student organizations, and seminars.

During our conversation this February, Professor Birman—who had previously worked with two other math students from New Delhi—asked me (partly in jest), "I have the impression of all the young girls in Delhi studying math. Is that true or not?". As I think back to her question, I am left with a deep gratitude for the people in Delhi and Chandigarh who encouraged me to pursue math since my childhood, particularly Dr. Tony Muça. Most importantly, I thank my parents, whose unconditional love and support has gotten me through

Columbia, with a special thanks to my mother for encouraging me not to give up math after high school, and to my father for all of our long conversations on the beauty of this subject since I was very young.

This is for my late *dadaji* [paternal grandfather] P.C. Wadhwa, who dreamt of becoming a mathematics professor as a young boy growing up in the 1940s in a small village near Lahore. While his life took a very different turn, he never ceased to take joy in the beauty of mathematics, and to share that joy with his children and grandchildren. I hope he would have enjoyed this work.

Introduction

"The theorem that I eventually proved is now known as the Birman exact sequence. It identifies the kernel of the homomorphism $\pi_0(\mathrm{Diff}(S_{g,n})) \to \pi_0(\mathrm{Diff}(S_{g,0}))$ *as the* n*-strand braid group of* S_g *modulo its center. In the case* $n = 1$ *the kernel reduces to* $\pi_1((S_g), \cdot)$*. That was very nice, but there was still something missing: how to construct a group of diffeomorphisms of* S_g *that would be isomorphic to* $\pi_1(S_g, \cdot)$*. My pleasure in discovering the map that is now known as the point pushing map was one of those 'aha' moments in mathematics. I remember to this day where I was standing in our home when I suddenly understood how to construct, in the mapping class group of an orientable surface* S *with a marked point* $\cdot$*, an arbitrary element in a subgroup of* $\mathrm{Diff}(S_g)$ *that would be isomorphic to* $\pi_1(S_g, \cdot)$*."*
— Joan Birman in [Bir16, p.2-3]

The aim of this thesis is to examine a problem that lies in the intersection of geometry, algebra, and the history of mathematics—the generation of the mapping class group. The group is a useful algebraic invariant of geometric surfaces. Given any surface, its mapping class group is the collections of distortions of the surface (smooth or topological) that preserve orientation, boundary points, and irregularities such as punctures or marked points.

As Birman's remarks imply, the study of the mapping class group draws together concepts from geometric and algebraic topology. To understand the algebraic structure of this group, we need to understand how distortions act on a surface up to isotopy. The 'point pushing map' and Birman exact sequence, which Birman proved in her doctoral dissertation in 1968, provides a case-in-point. Her work and subsequent research in this field relies on a combination of visual intuition and algebraic machinery—the relationship between which has long preoccupied the history of mathematics.

This thesis will examine a specific aspect of this research—the generators of the mapping class group, i.e. the smallest set of distortions from which we can obtain the entire group. Our aim is to demonstrate that the mapping class group of a surface is generated by Dehn twists, which are a particular distortion obtained by cutting a surface about the neighborhood a curve, twisting the cut-out component in a full circle, and gluing the components back together again.

0.1 Technical Notes

0.1.1 On smooth versus topological categories

The mapping class group could include one of two kinds of distortions of a surface: self-diffeomorphisms or self-homeomorphims. The kind of distortion we choose depends on the category of surfaces and maps we are working within. If we are working in the smooth category, we consider self-diffeomorphisms of smooth manifolds, i.e. maps from a smooth manifold to itself that are smooth and have a smooth inverse. This is what Birman refers to when she describes $\text{Diff}(S_{g,n})$, the group of all self-diffeomorphisms of a smooth manifold $S_{g,n}$ that preserve orientation and boundary points. By contrast, if we are working in the topological category, then we instead consider self-homeomorphisms of topological manifolds, i.e. maps from a topological manifold to itself that are continuous and have continuous inverse.

Both self-diffeomorphisms and self-homeomorphisms can preserve the boundary points, orientation, and irregularities of a surface. Thus, it is possible to define and study the mapping class group in either the smooth or the topological categories. For the purpose of this thesis, we will avoid switching between categories and instead work primarily in the smooth category, defining the mapping class group as isotopy classes of self-diffeomorphisms of a smooth manifold. As all self-diffeomorphisms are also self-homeomorphisms (as all smooth maps are necessarily continuous), all results that we prove in this thesis will also work in the topological category. Readers interested primarily in the topological category can refer to the analogous results in [FM11, Ch. 4].

0.1.2 Presumed background

This thesis presumes a basic familiarity with manifold theory, hyperbolic geometry, algebraic topology, and abstract algebra. We deal primarily with hyperbolic surfaces and manifolds, and presume basic definitions of these surfaces and maps between them, as given in [CB88, Ch. 1-3]. Additionally, as Birman's remarks indicate, our work also uses concepts from algebraic topology such as fiber bundles, exact sequences, homology groups, fundamental groups, and Van Kampen's theorem, discussed more fully in [Hat01]. Finally, we apply foundational concepts from group theory and homological algebra, including the isomorphism theorems, exact sequences, quotients, stabilizers, and free abelian groups.

0.2 Chapter Outline

This thesis proceeds in four chapters and an epilogue. The first two chapters are dedicated to establishing the necessary geometric and algebraic preliminaries. Chapter 1 introduces surfaces, curves on surfaces, the relationship between homotopy and isotopy, and self-diffeomorphisms of a surface, adapting results from [FM11, Chapters 1,2] and [Mar22, Chapter 6]. Chapter 2 then formally introduces the definition of a mapping class group and a Dehn twist. We end with three motivating examples of surfaces whose mapping class groups are generated by finitely-many Dehn twists about curves in the surface: the 2-torus (a compact surface) and a twice- and thrice-punctured sphere (non-compact surfaces).

The question remains: can we generalize the results from these examples to the mapping class group of any surface? This motivates the work of the second half of the thesis, which examines the generators of the mapping class group with broader generality. Chapter 3 generalizes the example of the 2-torus by proving that the mapping class group of any compact surface is generated by Dehn twists about non-separating simple closed curves, following the arguments in [Mar22, §6.5]. Chapter 4 generalizes this further by considering the mapping class group of a punctured—and therefore non-compact—surface. Drawing on [FM11, Chapter 4] and [Bir74], we analyze how punctures affect the mapping class group, building to a proof of the Birman exact sequence. Using this sequence, we prove that in the general case, the special subgroup of the mapping class group known as the pure mapping class group is generated by finitely-many Dehn twists and isotopies about simply closed curves. This generalizes our examples of the twice-and thrice-punctured sphere.

Finally, the epilogue situates this research on the generators of the mapping class group within an important problem in the history of mathematics—the relationship between algebra and geometry, between abstraction and concrete visualization. This section is based primarily on an interview the author conducted with Professor Birman about her research on the mapping class group, which is contextualized vis-à-vis recent scholarship on the history of diagrams and visual thinking in mid-twentieth-century algebraic and geometric topology such as [Ste23, Ch. 1]. It thus presents a meta-mathematical discussion on how the different proof methods developed in this thesis—which range from proof-by-picture to homological algebra—contributed to a broader shift in 1960s-70s mathematical research from pure abstraction to what historian of mathematics Alma Steingart calls "mathematical manifestations," or concrete ways to visualize, display, and present mathematical research in [Ste15, p.49].

Chapter 1

Preliminaries I: The Geometric, Algebraic, and Topological Setup

1.1 Surfaces and Curves

We begin by providing the basic geometric setup for this thesis: surfaces, diffeomorphisms, and curves.

1.1.1 Finite-type surfaces

In general, a surface is any smooth two-dimensional manifold, as defined in [Tu07, §5.3, p.52-53]. In this thesis, we deal specifically with **hyperbolic surfaces of finite type**, which we define below.

A surface of **finite type** is one that can be obtained from a compact, oriented surface S, possibly with boundary ∂S, by removing a finite number of points. In other words, we begin with a compact surface S, possibly of genus g (i.e. with g holes) and with b boundary components. We can then remove up to p points—puncturing the surface S up to p times—for some fixed value of $p \in \mathbb{Z}_{\geqslant 0}$.

Note that finite-type surfaces need not be compact: while the original surface S must be compact, any finite-type surface obtained by puncturing S a nontrivial amount of times will be non-compact. We denote such surfaces by $S_{g,b,p}$, with g referring to its genus, b referring to the number of boundary components, and p referring to the number of punctures. An example of one such surface is illustrated in Figure 1.1.

Given a finite-type surface $S_{g,b,p}$, we can obtain a **hyperbolic surface** of finite-type by endowing $S_{g,b,p}$ with a metric so that it locally resembles the hyperbolic plane $\mathbb{H}^2$. Hyperbolic surfaces are a concept from hyperbolic geometry, a new kind of geometry that emerged in the early nineteenth-century by discarding the parallel postulate of Euclidian geometry (that two parallel lines cannot intersect). The hyperbolic plane $\mathbb{H}^2$ is key model for hyperbolic geometry, just as the Euclidian plane $\mathbb{R}^2$ in Euclidian geometry; thus a hyperbolic surface locally looks like $\mathbb{H}^2$, where a (Euclidian) surface locally looks like $\mathbb{R}^2$.

A fuller discussion of the representations of the hyperbolic plane and of hyperbolic surfaces is provided in [CB88, Ch. 1-2]. Throughout this thesis, the term ‘surface’ should be

Figure 1.1: The finite-type surface $S_{3,2,2}$, obtained from removing 2 points from a surface of genus 3 with 2 boundary components.

understood as a hyperbolic surface of finite-type.

1.1.2 Self-diffeomorphisms

Given a surface $S_{g,b,p}$, a **self-diffeomorphism** of the surface is a smooth map from the surface to itself that has a smooth inverse. We now introduce an important property of all surfaces that is invariant under diffeomorphism.

Definition 1.1. *Given a finite-type surface $S_{g,b,p}$, its Euler characteristic is given by*

$$\chi(S_{g,b,p}) = 2 - 2g - b - p$$

.

To see that $\chi(S_{g,b,p})$ is invariant under diffeomorphism, note that any self-diffeomorphism must send boundary points to boundary points and punctures to punctures. Thus, the values of g, b, and p—and therefore of $\chi(S_{g,b,p})$—remain unchanged.

Given a finite-type surface $S_{g,b,p}$, let $\text{Diffeo}^+(S_{g,b,p})$ be the group of all orientation-preserving self-diffeomorphisms of the surface that fix its boundary point-wise. It is clear to see that this is a group with the group operation of function composition: it contains the identity, its inverses, and is closed under composition. We now introduce an important result regarding this group's topological structure.

Theorem 1.1. *Given a finite-type surface $S_{g,b,p}$, suppose $(g, b, p) \neq (0, 0, 0)$ (the 2-sphere S^2), $(0, 0, 1)$ (the plane $\mathbb{R}^2$ or the disc D^2), $(1, 0, 0)$ (the 2-torus T^2), $(0, 1, 0)$ (the closed annulus), or $(0, 0, 2)$ (the once-punctured plane $\mathbb{R}^2 - x$ or disc $D^2 - x$). Then $\text{Diffeo}^+(S_{g,b,p})$ is simply connected.* □

A full proof is provided in [Gra73, Thm 1, p.54-66]. Observe that Theorem 1.1 yields certain topological properties of the group $\text{Diffeo}^+(S_{g,b,p})$: in nearly all cases, its fundamental group $\pi_1(\text{Diffeo}^+(S_{g,b,p}))$ is trivial. This result will prove useful for our proof of the Birman exact sequence in Chapter 4.

1.1.3 Curves on surfaces

We now follow [Mar22, §6.3] to introduce different kinds of curves on a surface, building to the definition of a non-separating simple closed curve.

Definition 1.2. *A **curve** on a surface $S_{g,b,p}$ is a smooth map $\gamma : I \to S_{g,b,p}$ defined on some interval I. A **closed curve** is a smooth map $\gamma : S^1 \to S_{g,b,p}$.*

Definition 1.3. *A curve is **simple** if it is an embedding (i.e. its derivative is everywhere injective and it is a homeomorphism onto its image).*

Definition 1.4. *A simple closed curve on a surface is **trivial** if it bounds a disc.*

This thesis will primarily deal with simple closed curves that are not necessarily trivial. Observe that the derivative of a simple closed curve is nowhere vanishing, as it must always be injective—and therefore nontrivial. In other words, any simple closed curve is also **regular**. We now define cutting a surface along a simple closed curve following [FM11, §1.3.1], which is the precursor to our definition of a Dehn twist in Chapter 2.

Definition 1.5. *By **cutting** a surface $S_{g,b,p}$ by a simple closed curve α, we delete an arbitrary neighborhood* $\text{nbd}(\alpha)$ *of α from the surface. This yields a new surface that (by abuse of notation) we denote as $S_{g,b,p} - \alpha$ with two new boundary components (bounded by α) as well as a homeomorphism h between those boundary components so that:*

1. *The quotient $S_{g,b,p} - \alpha/(x \sim h(x))$ is homeomorphic to the original surface $S_{g,b,p}$ (i.e. we can glue along the new boundary to return to the original surface), and,*

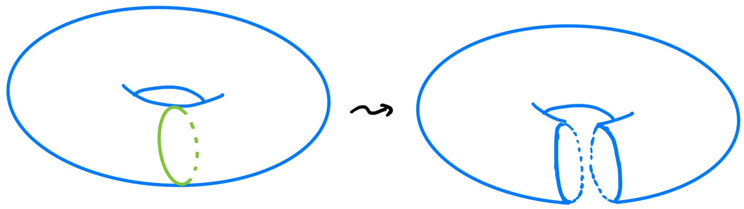
2. *The image of these two boundary components under this quotient map is the original curve α.*

We say a curve α is **separating** if $S_{g,b,p} - \alpha$ consists of multiple connected components—i.e., if cutting $S_{g,b,p}$ along α divides it into multiple connected components. Otherwise, we say a curve is **non-separating**. An illustration of these two kinds of curves is provided in Figure 1.2.

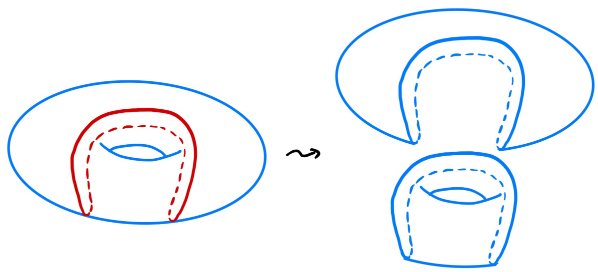
1.1.4 Intersections of curves

We now examine how curves intersect each other in order to study the homotopy classes of curves in a surface. Recall that two curves $\gamma_1, \gamma_2 : I \to S_{g,b,p}$ in a surface are **homotopic** if there exists a continuous function $F : S_{g,b,p} \times I \to S_{g,b,p}$ such that $F(x, 0) = \gamma_1(x)$ and $F(x, 1) = \gamma_2(x)$ for all $x \in S_{g,b,p}$, and that this is an equivalence relation. We need a way to characterize the intersection of curves that depends only on the curve's homotopy class. To this end, we introduce the concepts of transversality and the geometric intersection number.

Recall that two curves are in **transverse position** if at any point in their intersection their tangent spaces at that point generate the tangent space of the entire surface—locally like the hyperbolic plane $\mathbb{H}^2$. A generalized version of this definition for any two sub-manifolds of a manifold (rather than the specific case of curves on a surface) is provided in [GP78, §5, p.29-30]. Now we define the geometric intersection number accordingly.



(a) A non-separating simple closed curve



(b) A separating simple closed curve

Figure 1.2: Separating and non-separating simple closed curves on the 2-torus T^2.

Definition 1.6. *Let γ_1 and γ_2 be two simple closed curves in a surface $S_{g,b,p}$. Then their* ***geometric intersection*** *number $i_(\gamma_1, \gamma_2)$ is the minimum number of intersections of two transverse simple closed curves γ_1' and γ_2' homotopic to γ_1 and γ_2, respectively.*

Observe that by definition, the geometric intersection number depends only on the homotopy classes of curves γ_1, γ_2. We now introduce a mechanism to classify pairs of curves based on how they intersect.

Definition 1.7. *Two simple closed curves γ_1 and γ_2 in a surface $S_{g,b,p}$ are in* ***minimal position*** *if they intersect transversely at exactly $i(\gamma_1, \gamma_2)$ points.*

Our next step is to find a means of determining whether two transverse simple closed curves in a surface are in minimal position. Following [Mar22, §6.3.7], observe that the complement of two transverse simple closed curves is a finite disjoint union of open sets with polygonal boundaries. One such possibility is a **bigon**, a disc bounded with two sides, illustrated in Figure 1.3.

Theorem 1.2 (Bigon criterion)**.** *Two transverse simple closed curves γ_1 and γ_2 in a surface $S_{g,b,p}$ are in minimal position if and only if their complement does not form any bigons.*

Proof Sketch of Theorem 1.2. Let γ_1 and γ_2 be two transverse simple closed curves in a surface. For the leftwards direction, we prove the contrapositive: suppose the complement of γ_1 and γ_2 forms a bigon. Then we can apply homotopies to transform γ_1 and γ_2 into two curves that intersect at two fewer points, as demonstrated in Figure 1.4. This proves that γ_1 and γ_2 are not in minimal position.

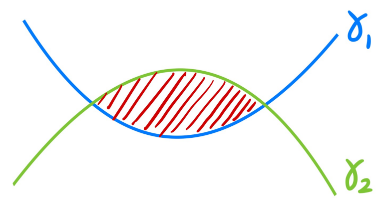


Figure 1.3: The complement of γ_1 and γ_2 can form a bigon.

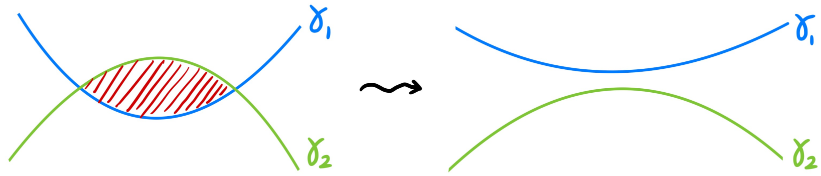


Figure 1.4: Homotopy reducing the number of points at which γ_1 and γ_2 intersect

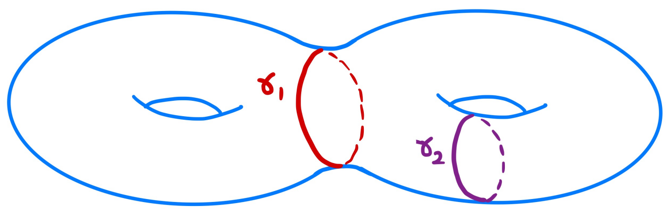
The proof of the rightwards direction is more involved, relying on casework and results from hyperbolic geometry. The reader can find the full proof in [Mar22, Thm 6.3.10]. This serves as an important reminder that while hyperbolic geometry appears to fade into the background, the hyperbolic structure of the surface is essential to defining its geometric properties and introducing the mapping class group. □

Our ability to classify curves in a surface based on how they intersect allows us to determine whether or not one curve can be mapped to another via a self-diffeomorphism of the surface. This result is formalized in the **change of coordinates principle**: a collection of curves can be mapped to another collection of curves in the same surface $S_{g,b,p}$ via an orientation-preserving diffeomorphism that fixes the boundary point-wise (i.e. some $\varphi \in \text{Diffeo}^+(S_{g,b,p})$) if they have the same intersection pattern and their complements have the same number of connected components. Here, intersection pattern refers to a topological description of how the curves in each collection intersect, based on their pairwise geometric intersection numbers (as given in Definition 1.6). This principle is given in [FM11, §1.3, p.38-41] for the topological category and in [Mar22, §6.3.4, p.169] for the smooth category. It follows directly from the classification of surfaces, given in [Mar22, Thm. 6.1.6].

In particular, this implies that there is only *one* non-separating simple closed curve in a surface up to diffeomorphism (since its complement always consists of one connected component, and that there can be no self-diffeomorphism from a non-separating to a separating curve (since the first has a complement of one connected component, while the second has a complement of more than one connected component). This is illustrated in Figure 1.5.

This principle has various useful implications for our study of curves in surfaces, some of which we will use in Chapter 4. In particular, it allows us to transform a collection of arbitrary non-separating simple closed curves into a simpler one with the same intersection pattern and complement via a self-diffeomorphism that lies in $\text{Diffeo}^+(S_{g,b,p})$. An example

(a) Both γ_1 and γ_2 are non-separating, and thus there is a diffeomorphism $\varphi \in \mathrm{Diffeo}^+(S_{2,0,0})$ such that $\varphi(\gamma_1) = \gamma_2$.



(b) Here γ_1 is separating while γ_2 is non-separating, so there is no self-diffeomorphism of $S_{2,0,0}$ sending γ_1 to γ_2.

Figure 1.5: Conditions for the application of the change of coordinates principle, illustrated on the two-holed torus $S_{2,0,0}$.

of this is illustrated in Figure 1.6.

1.2 Isotopies of curves

Our next step is to introduce the notion of an isotopy between curves in a surface, extending our discussion of the homotopy classes of curves and intersections of curves from the previous section. This provides the final conceptual tool needed to define the mapping class group as isotopy classes of $\mathrm{Diffeo}^+(S_{g,b,p})$.

1.2.1 Definitions

Having defined homotopy in §1.1.4, we now introduce the notion of an isotopy as a special kind of homotopy that is everywhere a diffeomorphism.

Definition 1.8. *Given a finite-type surface $S_{g,b,p}$ and two homeomorphisms $f, g : S_{g,b,p} \circlearrowleft$, an* ***isotopy*** *from f to g is a homotopy $F : S_{g,b,p} \times I \to S_{g,b,p}$ where for each $t \in I$, the function $F(x,t) : S_{g,b,p} \times \{t\} \to S_{g,b,p}$ is a homeomorphism.*

Recall from the preceding section that curves within a surface are given by maps from the unit circle S^1 into the surface. Thus, we can substitute f and g in the definition above with two curves within a surface. This motivates our definition of an isotopy between two curves in a surface.

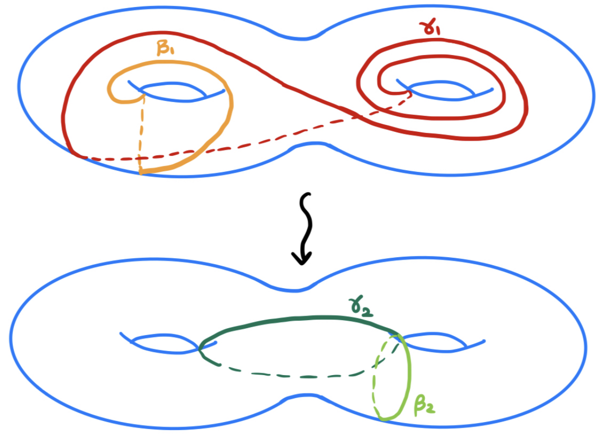


Figure 1.6: A pair of curves γ_1, β_1 that intersects once can be sent via a self-diffeomorphism of the surface to a simpler pair of curves γ_2, β_2 that likewise intersect once.

Definition 1.9. *Let α and β be two simple closed curves in a finite-type surface $S_{g,b,p}$. Then α and β are* ***isotopic*** *if there is a homotopy $H : S^1 \times I \to S_{g,b,p}$ from α to β such that for each $t \in I$, the closed curve $H(S^1 \times \{t\})$ is simple.*

As the conditions for an isotopy are stronger for those of a homotopy, it extend an isotopy of curves to an isotopy of an entire surface. We formalize this notion in the below proposition-*cum*-definition regarding ambient isotopies (i.e. an isotopy of the environment of the curves, namely the surface itself).

Proposition 1.3. *Given any finite-type surface $S_{g,b,p}$, let $F : S^1 \times I \to S_{g,b,p}$ be a smooth isotopy of simple closed curves as given in Definition 1.9. Then there exists an isotopy $H : S_{g,b,p} \times I \to S_{g,b,p}$ such that $H\big|_{S_{g,b,p}\times\{0\}}$ is the identity and $H\big|_{F(S^1\times\{0\})\times I} = F$. We call H an* ***ambient isotopy*** *of $S_{g,b,p}$.*

1.2.2 When does homotopy imply isotopy?

As is clear from the preceding subsection, all isotopies are homotopies, but not all homotopies are isotopies. The question becomes, when does homotopy imply isotopy? We will show that for hyperbolic surfaces of finite-type—i.e. for all surfaces considered in this thesis—these two conditions are interchangeable. This will be useful to us in defining the mapping class group Chapter 2, allowing us to link homotopy theory to our study of isotopy classes of $\text{Diffeo}^+(S_{g,b,p})$.

We first prove that homotopy and isotopy are interchangeable for a few special cases before moving to the general case, beginning with the 2-disc D^2. Here we briefly depart from this thesis's general focus on the smooth category over the topological category. When we work in the topological category (taking an isotopy to mean a homotopy that is everywhere a homeomorphism, rather than a diffeomorphism), then there is a relatively straightforward proof that isotopies and homotopies are interchangeable in D^2 known as the Alexander trick, given below.

Theorem 1.4 (The Alexander trick)**.** *Any boundary-preserving self-homeomorphism φ of the 2-disc D^2 is isotopic (in the topological category) to the identity homeomorphism i through homeomorphisms that are the identity on the boundary ∂D^2.*

Proof of Theorem 1.4 (The Alexander trick). We identify the 2-disc D^2 with the closed unit disc that lies in $\mathbb{R}^2$. Suppose $\varphi : D^2 \circlearrowleft$ is an orientation-preserving self-homeomorphism of D^2 that fixes its boundary point-wise. Then $\varphi\big|_{\partial D^2}$ is the identity $\mathrm{id}_{\partial D^2}$. We now define a map F given by

$$F(x,t) = \begin{cases} (1-t)\varphi(\frac{x}{1-t}), & 0 \leqslant |x| < 1-t \\ x, & 1-t \leqslant |x| \leqslant 1 \end{cases}$$

for $t \in [0,1)$, and we set $F(x,1)$ to be id_{D^2}. This yields the desired isotopy from φ to the identity on D^2. Observe that by construction, the homeomorphism $F(-,t)$ is the identity everywhere on the boundary ∂D^2 for all $t \in [0,1]$. This concludes the proof. □

Following [Mar22, §6.4], we observe that this proof cannot be directly extended to the smooth category: the isotopy F defined above is *not* smooth at $(0,0)$. While we could approximate F by smooth functions—since it is a homeomorphism and therefore continuous for all t—we would then lose the injectivity of $F(x,t)$ at given values of t, so our modified version of F is no longer an isotopy in the smooth sense. Nonetheless, Stephen Smale proved in the 1950s that the result does apply to the smooth case. Smale's proof that any two self-diffeomorphisms of D^2 that coincide on the boundary ∂D^2 are isotopic is given in [Mar22, §6.4, p.181], and draws on results from differential and hyperbolic geometry.

Using Smale's result, we can also conclude that homotopy and isotopy are interchangeable for the 2-sphere S^2.

Theorem 1.5. *Two self-diffeomorphisms $\varphi, \psi \in \mathrm{Diffeo}^+(S^2)$ are isotopic if and only if they are homotopic.*

Proof sketch of Theorem 1.5. Let φ, ψ be any two self-diffeomorphisms of S^2 that lie in $\mathrm{Diffeo}^+(S^2)$. Suppose they are homotopic. Pick any disc $D \subset S^2$, and observe that both φ and ψ send D to some disc in S^2. Now we can apply the Cerf-Palais Theorem, stated and proved in [Mar22, Thm 1.1.14, p.13], which gives us that any two orientation-preserving embeddings from the disc D into the interior of an connected and oriented manifold are ambiently isotopic. In particular, we can obtain an isotopy that relates the restrictions of φ and ψ to D, and therefore assume without loss of generality that φ and ψ coincide on D.

Then observe that the closed complement of D is another disc. Now we can conclude by our analogous results for D^2 in the smooth case, which gives an isotopy from φ to ψ on this disc that can be extended to the entirety of S^2. The reader can find a full proof in [Mar22, Thm 6.4.5, p.184]. □

We now claim that these results for D^2 and S^2 can in fact be generalized to any surface considered in this thesis.

Theorem 1.6. *Given any surface $S_{g,b,p}$, two self-diffeomorphisms $\varphi, \psi \in \mathrm{Diffeo}^+(S_{g,b,p})$ are isotopic if and only if they are homotopic.*

Proof sketch of Theorem 1.6. Take any surface $S_{g,b,p}$ and any self-diffeomorphism $\varphi \in \text{Diffeo}^+(S_{g,b,p})$. As with our proof of this statement for D^2, it suffices to show that φ is isotopic to the identity diffeomorphism $i \in \text{Diffeo}^+(S_{g,b,p})$ to conclude. But this is equivalent to showing that φ is the identity on all curves in $S_{g,b,p}$.

Thus, we pick any two sets of curves (called **multicurves**) and cut the surface along these curves. For surfaces of genus $g > 1$, this subdivides the surface into polygons formed by cutting open topological 'pairs of pants,' as is illustrated in Figure 1.7.

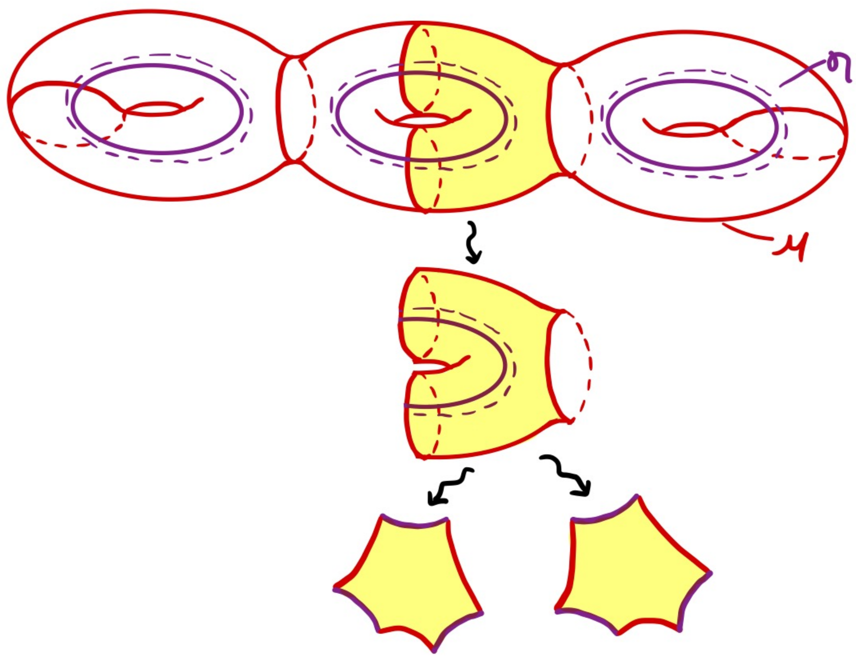


Figure 1.7: Cutting the surface $S_{3,0,0}$ along two multicurves μ (indicated in red, including the outer curve of the entire surface) and η (indicated in purple) to decompose it into polygons.

Using this decomposition of the complement of the two sets of curves in $S_{g,b,p}$, we then apply the bigon criterion from §1.1.4 and the earlier results regarding isotopies and homotopies on the disc D^2 to conclude. A full proof of the compact case (i.e. where $p = 0$) is given in [Mar22, Thm 6.4.7], and a more generalized version of this result for imbeddings (including all self-diffeomorphisms) is given in [Pat67, Thm 3, p.223]. □

Chapter 2

Preliminaries II: The Mapping Class Group and Dehn Twists

We now introduce the two key concepts for this thesis: the mapping class group and Dehn twists. This chapter defines the mapping class group and examines how cutting, puncturing, twisting, and otherwise modifying curves in a surface affects its mapping class group's algebraic structure. In doing so, it lays the groundwork for studying how Dehn twists generate the mapping class group in Chapters 3 and 4.

Chapter 2 proceeds in three sections. First, we define the mapping class group, discuss how it acts on the first homology group of a surface, and develop two preliminary examples—the mapping class groups of the 2-disc and 2-sphere, both of which, as we will see, are trivial. Second, to begin studying nontrivial elements of the mapping class group, we introduce Dehn twists and explore their relevant algebraic properties. Finally, we use Dehn twists to compute the mapping class groups of two non-compact surfaces (the 2-sphere with two and three punctures, respectively) and one compact surface (the standard 2-torus). These examples will provide the motivation for the generation theorems in the subsequent chapters of this thesis.

This chapter draws on [FM11, Chapters 2, 3] and [Mar22, Chapter 6], proving additional results to fill in the proofs and adapt them to the smooth category.

2.1 The Mapping Class Group

First, we define the mapping class group, examine its group actions, and compute two preliminary examples: the 2-disc and the 2-sphere.

2.1.1 Defining and topologizing the mapping class group

The mapping class group of a surface is a quotient of the group of all orientation-preserving self-diffeomorphisms of the surface that fix its boundary point-wise. We formalize this definition below.

Definition 2.1. *Given a finite-type surface $S_{g,b,p}$, its **mapping class group** is given by*

$$\mathrm{MCG}(S_{g,b,p}) := \mathrm{Diffeo}^+(S_{g,b,p})/\sim,$$

where, we will recall, $\text{Diffeo}^+(S_{g,b,p})$ *is the group of all orientation preserving self-diffeomorphisms of* $S_{g,b,p}$ *that fix the boundary point-wise, and two such maps* φ *and* ψ *are equivalent under the relation* $\sim$ *if they are connected by an isotopy that fixes the boundary point-wise at every level.*

Thus, elements of the mapping class group are isotopy classes of boundary- and orientation-preserving self-homeomorphisms of the surface $S_{g,b,p}$. These elements are called **mapping classes**, from which the name 'mapping class group' is ostensibly derived.

Remark 2.1. Because isotopy and homotopy are interchangeable for the surfaces considered in this thesis by §1.2.2, we could equivalently define the mapping class group to be the collection of homotopy classes of elements of $\text{Diffeo}^+(S_{g,b,p})$—which is none other than the 0th homotopy group $\pi_0(\text{Diffeo}^+(S_{g,b,p}))$.

We endow the mapping class group with the $\mathcal{C}^\infty$ topology, i.e., the intersections of the $\mathcal{C}^k$ topologies for all k. Note that this differs from the topology we give the mapping class group when working in the topological category, where the mapping class group is defined to be isotopy (equiv. homotopy) classes of $\text{Homeo}^+(S_{g,b,p})$. In the topological category, we endow the mapping class group with the compact-open topology, as discussed in [FM11, §2.1, p.45]. In the smooth category, we need additional structure to preserve smoothness when working with non-compact (i.e. punctured) surfaces, and therefore use the $\mathcal{C}^\infty$ topology. For a fuller discussion of the compact-open versus the $\mathcal{C}^\infty$ topologies, see [Hir76, Ch.2, §1].

2.1.2 Group action and induced homomorphism

We begin investigating the algebraic properties of the mapping class group by looking at how it acts on the first homology group of a surface. What emerges is an interesting connection to algebraic topology: how an algebraic invariant of a surface (its mapping class group) relates to a homotopy invariant (its first homology group). This will provide an insight into the algebraic structure of the mapping class group, enabling us to compute the motivating examples later in this chapter.

Recall that homology is a homotopy invariant, so it is also an isotopy invariant by Definition 1.8. Then the mapping class group $\text{MCG}(S_{g,b,p})$ acts on the first homology group of the surface with integer coefficients $H_1(S_{g,b,p}, \mathbb{Z})$, as homotopic (and therefore also isotopic) functions induce the same maps on homology. More precisely, let $\varphi : S_{g,b,p} \circlearrowleft$ be a representative of some mapping class $h \in \text{MCG}(S_{g,b,p})$. Then given any element $\alpha \in H_1(S_{g,b,p}, \mathbb{Z})$, we send the pair $(h, \alpha) \in \text{MCG}(S_{g,b,p}) \times H_1(S_{g,b,p}, \mathbb{Z})$ to the element $\varphi_*(\alpha) \in H_1(S_{g,b,p}, \mathbb{Z})$, where φ_* is the map induced by φ on the first homology group. By homotopy invariance of homology, this process does not depend on the choice of representative φ in the mapping class h, so the group action is well defined.

Observe that this group action yields a natural group homomorphism from the mapping class group to the group of all automorphisms of its first homology group that sends h to φ_*. We claim that φ_* is in fact orientation-preserving, and therefore an element in $\text{Aut}^+(H_1(S_{g,b,p}, \mathbb{Z}))$, the group of all orientation-preserving automorphisms of the first homology group. This follows from a more general claim: elements of the mapping class group preserve the intersection form of pairs of elements in the first homology group. The proof of this claim uses cup products in cohomology and the concept of an intersection form

$\omega \in H_1(S_{g,b,p}, \mathbb{Z}) \times H_1(S_{g,b,p}, \mathbb{Z})$ from differential geometry to show that $\varphi_*(\omega) = \omega$, and thus that φ_* preserves the orientation of elements of $H_1(S_{g,b,p})$. For a fuller explanation of the mapping class group's actions on homology, see [FHV21, §1-§2] and [FM11, §6.3]. Thus we in fact have a natural group homomorphism $\xi : \mathrm{MCG}(S_{g,b,p}) \to \mathrm{Aut}^+(H_1(S_{g,b,p}, \mathbb{Z}))$ given by $\xi(h) = \varphi_*$.

Following [Mar22, §6.5], we can develop this observation further. For $n = 2g + \max\{b + p - 1, 0\}$, the group $\mathrm{Aut}^+(H_1(S_{g,b,p}, \mathbb{Z})) \cong \mathrm{Aut}^+(\mathbb{Z}^n)$, the group of orientation-preserving automorphisms of $\mathbb{Z}$ times itself n times. But this can be simplified even further. The group of automorphisms of $\mathbb{Z}^n$ is equivalent to the general linear group $\mathrm{GL}_n(\mathbb{Z})$, consisting of $n \times n$ invertible matrices. Of these, the subgroup of orientation-preserving automorphisms $\mathbb{Z}^n$ is equivalent to the special linear group $\mathrm{SL}_n(\mathbb{Z})$, consisting of $n \times n$ invertible matrices with determinant 1. Thus, we have that $\mathrm{Aut}^+(\mathbb{Z}^n) \cong \mathrm{SL}_n(\mathbb{Z})$, and thus we can consider ξ to be a group homomorphism from $\mathrm{MCG}(S_{g,b,p})$ to $\mathrm{SL}_n(\mathbb{Z})$.

Remark 2.2. In general, we cannot say if ξ is injective or surjective. However, we will see in the motivating examples at the end of this chapter that there is a specific case in which ξ is actually an isomorphism—allowing us to apply existing knowledge about the first homology group of a surface to compute its mapping class group.

Remark 2.3. The kernel of ξ is called the **Torelli group** of $S_{g,b,p}$.

2.1.3 Two preliminary examples

We compute two preliminary examples, relying on results from §1.2.2 to demonstrate that the mapping class groups of the 2-disc and 2-sphere are trivial.

Example 2.4 (Mapping class group of the disc). Consider the 2-disc D^2, which is diffeomorphic to the once-punctured 2-sphere $S_{0,0,1}$ under the stereographic projection. We claim its mapping class group $\mathrm{MCG}(D^2)$ is trivial.

By the smooth analogue of Theorem 1.4 (The Alexander trick), we have that any boundary-preserving self-diffeomorphism of D^2 is isotopic to the identity via an isotopy that preserves the boundary ∂D^2. Then $\mathrm{MCG}(D^2)$ is trivial, containing only the equivalence class of the identity.

Example 2.5 (Mapping class group of the two-sphere). Consider the 2-sphere $S^2 = S_{0,0,0}$. We claim that its mapping class group $\mathrm{MCG}(S^2)$ is trivial.

Observe that there are only two self-diffeomorphisms of S^2 up to homotopy: the identity i and reflection over the z-axis r (i.e., restricting the symmetry of $\mathbb{R}^3$ around the plane $z = 0$). By Corollary 1.5, these are also the only two self-diffeomorphisms of S^2 up to isotopy. However, r is not orientation-preserving, and therefore not an element of $\mathrm{Diffeo}^+(S^2)$. Then once again the identity i is the only self-diffeomorphism of S^2 up to isotopy, so $\mathrm{MCG}(S^2)$ is trivial.

2.2 Dehn twists

The preliminary examples of the 2-disc and 2-sphere raise the question: what do nontrivial elements of the mapping class group look like? This motivates our discussion of Dehn twists, which are nontrivial elements of the mapping class group of a surface that have rich geometric qualities and algebraic structure. While there is a vast literature on the properties of Dehn twists, we restrict our focus to properties that are relevant to understanding how Dehn twists generate elements of the mapping class group, providing the foundation for Chapters 3 and 4.

We proceed in three sections: we introduce the Dehn twist, the braid relation (which reveals the algebraic relationship between two Dehn twists), and finally the cutting homomorphism (which uses Dehn twists to understand how cutting a finite-type surface along a curve affects its mapping class group).

2.2.1 Defining and visualizing Dehn twists

We first define the Dehn twist and discuss how it is a well-defined element of the mapping class group of a finite-type surface.

Definition 2.2. *Given a finite-type surface $S_{g,b,p}$, let γ be a non-trivial simple closed curve along the interior of the surface. Then a **Dehn twist** along γ is the element $T_\gamma \in \mathrm{MCG}(S_{g,b,p})$ defined as follows.*

Pick a tubular neighborhood of γ that is orientation-preservingly diffeomorphic to $S^1 \times [-1,1]$, where γ lies in the middle as $S^1 \times \{0\}$. Let $f : [-1,1] \to \mathbb{R}$ be given by

$$f(t) = \begin{cases} 0, & -1 \leqslant t \leqslant -\frac{1}{2} \\ 2\pi \cdot \phi(t + \frac{1}{2}), & -\frac{1}{2} \leqslant t \leqslant \frac{1}{2} \\ 2\pi, & \frac{1}{2} \leqslant t \leqslant 1, \end{cases}$$

where ϕ is the bump function defined on the unit interval $[0,1]$. Then f is a smooth function that is 0 on $[-1,-\frac{1}{2}]$ and 2π on $[\frac{1}{2},1]$.

Then let $T_\gamma : S_{g,b,p} \circlearrowleft$ be a self-diffeomorphism of the surface $S_{g,b,p}$ that is given by $T_\gamma(e^{i\alpha}, t) = (e^{i(\alpha+f(t))}, t)$ on the tubular neighborhood and is the identity on its complement (i.e., everywhere else on $S_{g,b,p}$).

This process is illustrated in Figure 2.1.

As discussed in [Mar22, §6.5.3, p.187], there are multiple different ways to visualize a Dehn twist. One visualization is 'cutting and gluing.' After selecting a simple closed curve γ, we cut the surface $S_{g,b,p}$ along γ, a process that yields two new boundary components. We take a neighborhood of one of these boundary components—the tubular neighborhood of γ referenced in 2.2—and twist it in a full circle (i.e., by an angle of 2π) to the left. Finally, we re-glue, obtaining a boundary-preserving self-homeomorphism of the surface $S_{g,b,p}$.

In fact, Dehn twists about simple closed curves in a surface as described in Definition 2.2 are well-defined, nontrivial elements of its mapping class group.

Proposition 2.1. *Given a non-trivial simple closed curve $\gamma \in S_{g,b,p}$, the element $T_\gamma \in \mathrm{MCG}(S_{g,b,p})$ is well-defined and depends only on the isotopy class of γ.*

Proof sketch of Proposition 2.1. The proof of this statement reduces to demonstrating that the Dehn twist about γ depends only on choice of γ. In other words, T_γ must be independent of the choice of tubular neighborhood of γ and smooth function f from Definition 2.2.

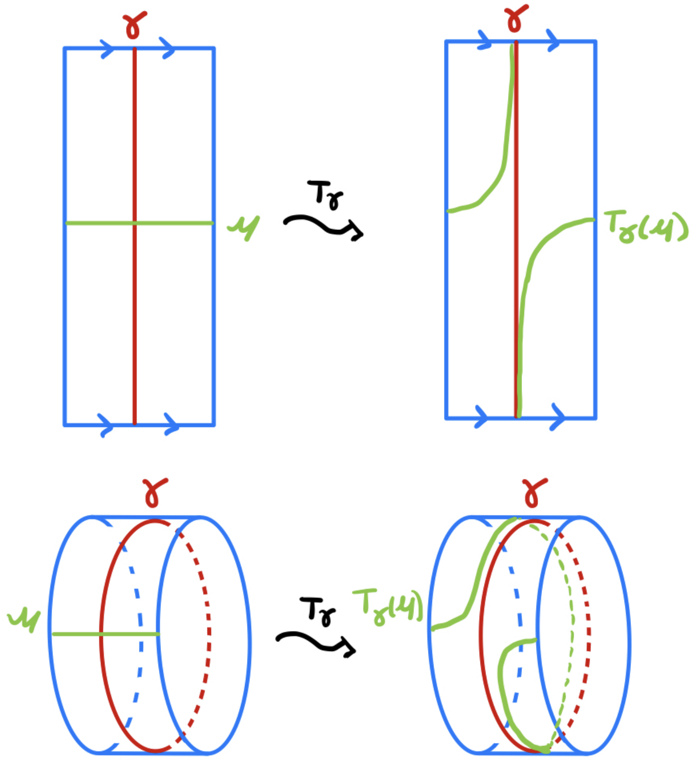


Figure 2.1: A Dehn twist along the curve γ maps the arc μ onto an arc that makes a left turn.

These facts follow from results in Chapter 1 and in [Mar22, Chapter 1]. Tubular neighborhoods are ambiently isotopic, and functions like f with fixed extremes are also ambiently isotopic. It follows that T_γ is well-defined, depending only on the isotopy class of γ. Clearly T_γ is orientation-preserving and fixes the boundary of $S_{g,b,p}$ pointwise, so we conclude. □

From now onwards, we refer to the Dehn twist about an *isotopy class* a of non-trivial simple closed curves instead of a single curve. Given any representative $\gamma \in a$, T_a will be taken to be T_γ, which is well-defined regardless of choice of γ by Proposition 2.1 above.

Remark 2.6. Observe that T_γ relies on the orientation of the finite-type surface $S_{g,b,p}$, but not on an orientation for γ. In particular, T_γ is unaffected if we reverse the orientation of γ, i.e. $T_\gamma = T_{\gamma^{-1}}$.

Thus, the inverse of a Dehn twist T_γ is *not* given by taking a Dehn twist about the inverse of γ. Rather, the inverse T_γ^{-1} transforms every curve μ crossing γ via a complete right-turn—as opposed to a complete left-turn. We call this a **negative Dehn twist**.

2.2.2 The braid relation

We now turn to the algebraic relationship between two Dehn twists, which motivates our discussion of the braid relation. This concept will be useful for our study of how Dehn twists generate the mapping class group in subsequent chapters. While a full discussion of the braid relation falls outside the scope of this thesis, here we investigate whether or not particular configurations of Dehn twists commute. As we shall see, given two isotopy classes of non-separating simple closed curves a and b, this largely depends on their geometric intersection number $i(a, b)$.

First observe that Dehn twists along disjoint isotopy classes of curves commute: for isotopy classes of simple closed curves a, b such that $i(a, b) = 0$, it follows that $T_a T_b = T_b T_a$. This can be verified directly by computation.

We now move on to the more interesting case regarding Dehn twists along two isotopy classes of simple closed curves that *do* intersect, i.e. isotopy classes a, b such that $i(a, b) \geqslant 1$. By the change of coordinates principle, it suffices to check for the case where $i(a, b) = 1$, which greatly simplifies the necessary computations. There is in fact a braid relation between intersecting Dehn twists, given in [FM11, Prop 3.11, 77-78]. We will prove a modified version of the braid relation that deals only with the product of two Dehn twists.

Theorem 2.2. *If a and b are isotopy classes of simple closed curves that satisfy $i(a, b) = 1$, then $T_a T_b(a) = b$.*

Proof of Theorem 2.2. Our proof is largely pictorial. First, by the change of coordinates principle, it suffices to check that Theorem 2.2 holds for two isotopy classes of simple closed curves a and b with geometric intersection $i(a, b) = 1$. Let α be a representative curve of a and β be a representative curve of b. Then we check via a computation that $T_a T_b(\alpha) = \beta$ for all such α and β, illustrated in Figure 2.2.

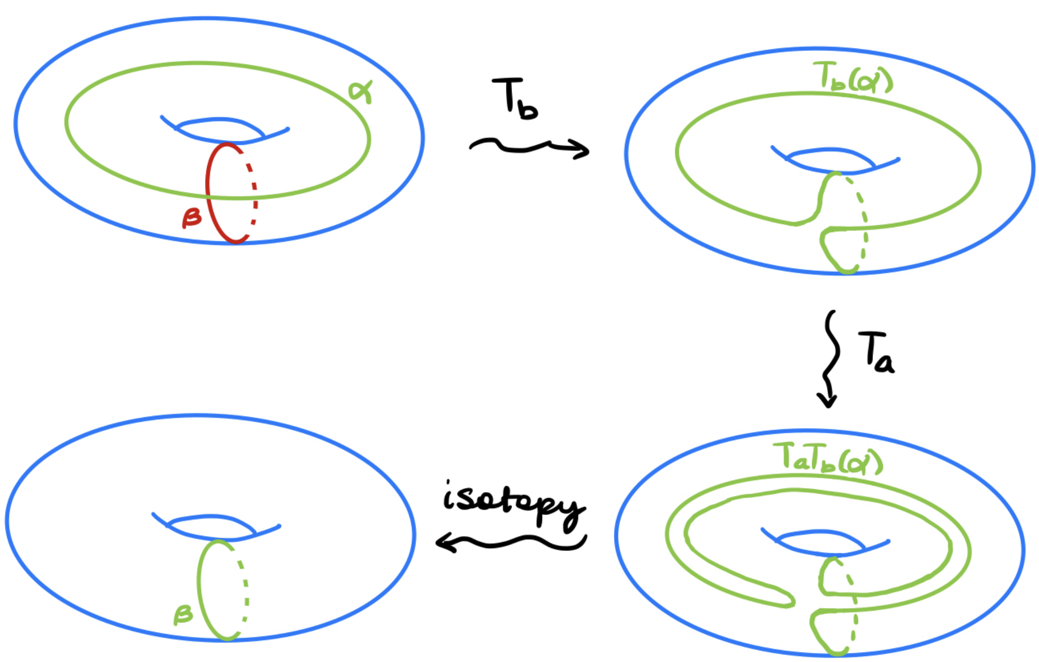


Figure 2.2: Computation demonstrating that $T_a T_b(\alpha) = \beta$

From this we can conclude that $T_a T_b(a) = b$ as desired. □

2.2.3 The cutting homomorphism

Finally, we introduce the cutting homomorphism, which allows us to use Dehn twists to analyze how cutting a surface along a simple closed curve α affects its mapping class group $\mathrm{MCG}(S_{g,b,p})$.

The cutting homomorphism is one of three major homomorphisms on the mapping class group we can analyze, the other two being the inclusion homomorphism (given by including the surface $S_{g,b,p}$ into another surface, possibly without punctures) and the capping homomorphism (given by 'capping' one boundary component of a surface to obtain a modified surface with one less boundary component). The cutting, capping, and inclusion homomorphisms are discussed in full in [FM11, §3.6, p.82-88]. We restrict the discussion to the cutting homomorphism, which is directly relevant to our work in subsequent chapters.

First we provide some visual intuition. Given a simple closed curve α in a surface $S_{g,b,p}$, let a be the isotopy class of α, and suppose we cut $S_{g,b,p}$ along α. We want to relate the stabilizer of a in the mapping class group $\mathrm{MCG}(S_{g,b,p})$ to the mapping class group of the modified surface $S_{g,b,p}-\alpha$. Set $\mathrm{MCG}(S_{g,b,p}, a) := \mathrm{Stab}(a)$, the stabilizer of a in $\mathrm{MCG}(S_{g,b,p})$. We take the most straightforward homomorphism from $\mathrm{MCG}(S_{g,b,p}, a) \to \mathrm{MCG}(S_{g,b,p} - \alpha)$: for any isotopy class h in $\mathrm{MCG}(S_{g,b,p}, a)$, let φ be some representative element. By definition, φ must fix a. Then we can restrict the homeomorphism φ to $S_{g,b,p} - \alpha$ to obtain a new homeomorphism that is an element of $\mathrm{MCG}(S_{g,b,p} - \alpha)$. Call this map ξ.

We will now show that ξ is well-defined, which reduces to showing that it descends to the quotients on both the domain and co-domain (i.e. up to isotopy, which by §1.2.2 is the same as up to homotopy).

Theorem 2.3 (Modified version of the cutting homomophism)**.** *Let $S_{g,b,p}$ be any finite-type surface. Let α be a non-separating simple closed curve in $S_{g,b,p}$ and let a be its isotopy class. Then there is a well-defined homomorphism*

$$\xi : \mathrm{MCG}(S_{g,b,p}, a) \to \mathrm{MCG}(S_{g,b,p} - \alpha)$$

with kernel $< T_\alpha >$.

Proof sketch of Theorem 2.3. It is clear that the map ξ is a homomorphism provided that it is well-defined. Thus, our task reduces to showing that ξ is well-defined, and that its kernel is generated by the Dehn twist about α.

Let N a tubular neighborhood around α and consider its complement in $S_{g,b,p}$, given by $S_{g,b,p}-N$. We can include $S_{g,b,p}-N \to S_{g,b,p}$, which induces a homomorphism on the mapping class groups (specifically, the inclusion homomorphism) $\varphi_1 : \mathrm{MCG}(S_{g,b,p}-N) \to \mathrm{MCG}(S_{g,b,p})$. By [FM11, Thm 3.18, p.84], the kernel of this homomorphism K_1 is generated by the product of Dehn twists $T_{\alpha^+}T_{\alpha^-}$, where α^+ and α^- are the two boundary components of N isotopic to α in $S_{g,b,p}$.

We now cap these two boundary components with a punctured disc, obtaining a surface $\overline{S_{g,b,p} - N}$ that is homeomorphic to $S_{g,b,p} - \alpha$. As these surfaces are homeomorphic, we have the canonical isomorphism $\psi : \mathrm{MCG}(\overline{S_{g,b,p} - N}) \to \mathrm{MCG}(S_{g,b,p} - \alpha)$.

Finally, observe that we can include the surface $S_{g,b,p} - N$ into the capped surface $\overline{S_{g,b,p} - N}$. Again by [FM11, 3.18, p.84], this induces the inclusion homomorphism on their mapping class groups, given by $\varphi_2 : \mathrm{MCG}(S_{g,b,p} - N) \to \mathrm{MCG}(\overline{S_{g,b,p} - N})$ whose kernel K_2 is generated by T_{α^+} and T_{α^-} for α^+, α^- as defined above.

The final result is obtained by inspecting the following commutative diagram (which follows from the isomorphism theorems regarding the kernel of group homomorphisms). From this, we observe that $\psi \circ \varphi_1 \circ \varphi_2^{-1}$ is well defined as K_1 is a subgroup of K_2.

$$
\begin{array}{ccccccc}
 & & & & K_1 & & \\
 & & & & \downarrow & & \\
0 & \longrightarrow & K_2 & \longrightarrow & \mathrm{MCG}(S_{g,b,p} - N) & \xrightarrow{\varphi_2} & \mathrm{MCG}(\overline{S_{g,b,p} - N}) \\
 & & & & \varphi_1 \downarrow & & \psi \downarrow \\
 & & & & \mathrm{MCG}(S_{g,b,p}, a) & \xrightarrow{\xi} & \mathrm{MCG}(S_{g,b,p} - \alpha) \\
 & & & & \downarrow & & \\
 & & & & 1 & &
\end{array}
$$

We conclude by noting that this is precisely the definition of ξ, and its kernel is precisely T_α. □

2.3 Motivating examples

We now introduce some motivating examples that reflect the richness of the structure of the mapping class group. We have already introduced two preliminary examples—the 2-sphere S^2 and the 2-disc D^2—in §2.1.3. Now we draw on the technical machinery introduced in §2.2 to conduct three nontrivial computations of the mapping class group.

We first consider the 2-torus T^2 as an example of a compact (i.e. non-punctured) surface with a nontrivial mapping class group. As we shall see, its mapping class group is in fact generated by two Dehn twists along non-separating simple closed curves, which we will generalize to all compact surfaces $S_{g,b,0}$ in §3.2 of Chapter 3. Next, we consider two non-compact cases, the twice- and thrice-punctured spheres. These examples foreshadow the challenges that arise when dealing with punctured surfaces. This motivates our work in Chapter 4, where we will generalize these examples to the sphere with n punctures in §4.3.1 and then to punctured surfaces more broadly.

2.3.1 The standard 2-torus

Example 2.7 (Mapping class group of the torus). Consider $T^2 = S_{1,0,0}$, the standard 2-torus. We claim that $\mathrm{MCG}(T^2) \cong SL_2(\mathbb{Z})$ and is indeed generated by two Dehn twists about non-separating simple closed curves.

First, we fix two curves in T^2: the meridian m, and the longitude l. As illustrated in Figure 2.3, m and l are both non-separating simple closed curves. If we identify T^2 with the quotient $\mathbb{R}^2/\mathbb{Z}^2$, then m corresponds to the generator $(1, 0)$ and l corresponds to the generator $(0, 1)$ of its first homology group $H_1(T^2)$.

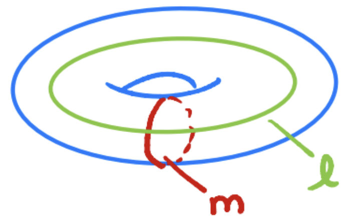


Figure 2.3: Meridian and longitudinal curves in the torus T^2

Recall from §2.1.2 that we have a group homomorphism $\xi : \text{MCG}(T^2) \to \text{SL}_2(\mathbb{Z})$ (where $n = 2 \cdot 1 + 0 = 2$). We will show that in this case, ξ is in fact an isomorphism.

Injectivity of ξ: Let φ be some representative element of an isotopy class in the kernel of ξ (i.e. in the Torelli group of T^2). By definition, φ must act trivially on the first homology group $H_1(T^2, \mathbb{Z}) = \mathbb{Z}^2$. It must send the curves m and l to two simple closed curves $\varphi(m)$ and $\varphi(l)$ that are homotopic to m and l, respectively. By §1.2.2, they must also be isotopic to m and l, respectively. Then by Theorem 1.6, φ is isotopic to the identity, so the kernel of ξ is trivial and thus ξ is injective.

Surjectivity of ξ: Note that a matrix $A \in \text{SL}_2(\mathbb{Z})$ acts linearly on $\mathbb{R}^2$, preserving the orientation and lattice structure of $\mathbb{Z}^2$. Then A must descend to a self-diffeomorphism of the quotient $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, which is mapped to A under ξ. Thus, ξ is also surjective. We conclude that $\text{MCG}(T^2) \cong \text{SL}_2(\mathbb{Z})$.

Generating $\text{MCG}(T^2)$: There is more we can say about the structure of the mapping class group of T^2, which will relate to our previous discussion of Dehn twists while also offering a flavor for the investigation of the mapping class group's generators in subsequent chapters. We will show that $\text{MCG}(T^2)$ is generated by Dehn twists about the two isotopy classes of non-separating simple closed curves of T^2—namely, T_m and T_l.

To prove this, we will use our identification of $\text{MCG}(T^2)$ with $\text{SL}_2(\mathbb{Z})$. Note that $\text{SL}_2(\mathbb{Z})$ is generated by the matrices $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. We claim that in fact the Dehn twists are

$$T_m = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, T_l = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Now we check that this result holds via computation. Recall that we have identified m and l with generators $(1, 0)$ and $(0, 1)$ respectively of $T^2 \cong \mathbb{R}^2/\mathbb{Z}^2$. We can write these as column vectors to compute how these elements of the mapping class group act on homology:

$$T_m(m) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = m$$

$$T_m(l) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = l - m$$

$$T_l(l) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = l$$

$$T_l(m) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = l + m.$$

Thus we conclude.

The 2-torus provides our first example of a compact surface with a nontrivial mapping class group. As we have see, the algebraic machinery from §2.1.2 regarding the group action of the mapping class group on the first homology group enabled us to understand the structure of the mapping class group up to isomorphism. We have also seen that for T^2, not only are Dehn twists *elements* of the mapping class group but the group is in fact generated by two Dehn twists along the non-separating simple closed curves m and l, which correspond to the generators $(1, 0)$ and $(0, 1)$ of T^2. Chapter 3 will generalize this result to for any compact surface. For now, we turn to the other interesting case—the mapping class group of a non-compact surface.

2.3.2 The twice- and thrice-punctured 2-sphere

This section will take up one of the simplest examples of the mapping class group of a punctured surface, namely the twice- and thrice-punctured 2-sphere. Recall that by §2.1.3, both the 2-sphere $S^2 = S_{0,0,0}$ and the once-punctured 2-sphere $D^2 \cong S_{0,0,1}$ have trivial mapping class groups. The situation changes once we begin to add additional punctures, revealing how puncturing a surface—and rendering it non-compact—complicates our analysis of the structure and generators of its mapping class group.

As these examples are primarily intended to provide motivation, we will use this section to introduce some techniques for dealing with the punctured case and offer sketches of the computations of their mapping class groups.

First, by [Bir74] and [FM11], we can think of puncturing a finite-type surface as the same thing as marking a point on the surface: both actions have the same effect on the mapping class group, which must send punctures to other punctures (or marked points to other marked points). We now state a useful lemma.

Lemma 2.4. *We regard the thrice-punctured sphere $S_{0,0,3}$ as the 2-sphere with three marked points. Then any two essential proper arcs in $S_{0,0,3}$ with the same endpoints are isotopic. Moreover, any two essential proper arcs that start and end at the same marked point are isotopic.*

Here, an essential proper arc is a properly-embedded arc α with endpoints on the boundary of the surface that is non-trivial (i.e., α is not isotopic to a single boundary point) and has no parallel components (i.e., α is not isotopic to an arc contained entirely within the boundary, and must therefore interact with the interior of the surface).

Proof sketch of Lemma 2.4. This proof draws upon a number of results from Chapter 1. The idea is roughly as follows: let α and β be two arcs in $S_{0,0,3}$ with endpoints on the same two distinct marked points. We recall that the once-punctured sphere is diffeomorphic to the plane $\mathbb{R}^2$ (as well as the disc D^2) via the stereographic projection. Then, we can regard one of the marked points—corresponding to one of the punctures—as the point at infinity from which we project downwards, and then regard α and β as arcs between two marked points (the remaining two punctures) on a plane.

Either α and β are disjoint, or they bound a disc, in which case we can apply the bigon criterion from Theorem 1.2 to isotope them until they are in minimal position and have disjoint interiors. We now cut the surface along $\alpha \cup \beta$, which yields a disc (with the two marked points on the boundary) and another disc (with the remaining third marked point). Then, by results from §1.6, α and β bound an embedded disc and must therefore be isotopic.

The proof is roughly analogous for the case where the two essential proper arcs start and end at the same marked point, and is discussed further in [FM11, Prop 2.2, p.49]. $\square$

We now apply this lemma to compute the mapping class group of the thrice-punctured 2-sphere.

Example 2.8 (Mapping class group of the thrice-punctured 2-sphere). Consider $S_{0,0,3}$, the sphere with three punctures. We claim $\mathrm{MCG}(S_{0,0,3}) \cong S_3$, the symmetric group of degree 3 containing all permutations of the three-element set $\{1, 2, 3\}$.

Observe firstly that $\mathrm{MCG}(S_{0,0,3})$ acts on the elements of S_3 by permutation: mapping classes in $\mathrm{MCG}(S_{0,0,3})$ must send punctures to other punctures, and thus acts on the symmetric group by permuting the three punctures. This group action yields a natural homomorphism $\varphi : \mathrm{MCG}(S_{0,0,3}) \to S_3$. Clearly, this homomorphism is surjective. We claim that it is in fact an isomorphism, from which we will conclude that $\mathrm{MCG}(S_{0,0,3}) \cong S_3$ as desired.

It remains to show that φ is injective, i.e., that any representative self-diffeomorphism ψ of $S_{0,0,3}$ that fixes all three marked points, say x_1, x_2, x_3 (i.e., is mapped to the identity element in S_3) is isotopic to the identity in $\mathrm{MCG}(S_{0,0,3})$. Given ψ as defined above, let α be an arc in the surface $S_{0,0,3}$ that has endpoints at two distinct marked points: without loss of generality, say at x_1 and x_2. Since $\varphi(\psi) = 1$, i.e., ψ fixes all marked points, it follows that the endpoints of the arc $\psi(\alpha)$ are also x_1 and x_2. Then Lemma 2.4 implies that the arcs α and $\psi(\alpha)$ are isotopic. Then by [FM11, Prop 1.11], ψ must be isotopic to a map that fixes the arc α point-wise.

Now, cut $S_{0,0,3}$ along α, which will yield a 2-disc that has one marked point (the remaining point x_3) and a boundary (α). We will reduce this to the case of the 2-disc from Example 2.4. Observe that ψ induces a self-homeomorphism ψ' of the disc that fixes the boundary point-wise. But then ψ' lies in $\mathrm{MCG}(D^2)$, which is trivial by Example 2.4, so there is an isotopy from ψ' to the identity. This yields an isotopy from ψ to the identity, proving that the kernel of φ is trivial and thus that φ is injective. Thus we conclude.

Finally, we consider the twice-punctured 2-sphere.

Example 2.9 (Mapping class group of the twice-punctured 2-sphere). Consider $S_{0,0,2}$, the sphere with two punctures. We claim $\mathrm{MCG}(S_{0,0,2}) \cong S_2$, the symmetric group of degree 2 containing all permutations of the two-element set $\{1, 2\}$.

Observe how $\mathrm{MCG}(S_{0,0,2})$ likewise acts on $S_2 \cong \mathbb{Z}/2\mathbb{Z}$ by permuting its two elements. We can then follow the outline of the proof offered in Example 2.8 to conclude.

These two examples highlight the complexity of the non-compact case, which we will return to in Chapter 4.

Chapter 3

Generating the Mapping Class Group: A Special Case

The motivating examples of surfaces with nontrivial mapping class groups raise the question: what can we say about the generators of the mapping class group? For the standard 2-torus $T^2 = S_{2,0,0}$, we saw in Example 2.7 that its mapping class group was generated by two Dehn twists along non-separating simple closed curves in the surface. This chapter will generalize this result to any compact surface, which is a special case of a broader result we will prove in Chapter 4 for all (possibly punctured) surfaces.

Our objective here is to prove the following statement:

Theorem 3.1. *For $g \geqslant 0$, the mapping class group $g = \mathrm{MCG}(S_{g,b,0})$ is generated by Dehn twists about non-separating simple closed curves of the surface $S_{g,b,0}$.*

This is, in fact, a modified version of what is known as the Dehn-Lickorish theorem.

Theorem 3.2 (Dehn-Lickorish)**.** *For $g \geqslant 0$, the mapping class group $g = \mathrm{MCG}(S_{g,0,0}$ is generated by finitely many Dehn twists about non-separating simple closed curves.*

As discussed in [FM11, p.89], Theorem 3.2 was developed separately by Max Dehn (1878-1952) and William Lickorish (1938-) over the course of the twentieth century. Dehn first proved that the mapping class group of $S_{g,0,0}$ is generated by $2g(g-1)$ Dehn twists in 1938 in [Deh87, The Group of Mapping Classes, §9]. In 1967, David Mumford developed this result further to show that it was only necessary to take Dehn twists along non-separating simple closed curves. Simultaneously, Lickorish—evidently not aware of Dehn's work on the generators of the mapping class group—independently proved that the mapping class group of $S_{g,0,0}$ is generated by Dehn twists along $3g-1$ non-separating simple closed curves.

We will set aside proving the full statement of Theorem 3.2—specifically showing that the mapping class group of a compact surface is *finitely* generated—until Chapter 4, where we introduce more advanced machinery to deal with both the compact and the non-compact cases. In this chapter, our aim is to use slightly simpler machinery to prove the statement of Theorem 3.1.

We will work towards this proof in two stages. In the first section, we introduce the concept of relatedness of curves and arcs in a finite-type surface $S_{g,b,p}$ and prove two important lemmas regarding the relatedness of curves and arcs within the interior of compact surfaces $S_{g,b,0}$. In the second section, we apply these results to prove Theorem 3.1 through casework and induction on g and b. Subsequently, we discuss some issues with generalizing this result to the non-compact case, which will provide the motivation for our investigation of the general case in Chapter 4.

3.1 The Concept of Relatedness

In this section, we introduce the notion of 'relatedness' of curves and arcs and prove two useful lemmas regarding the curves and arcs in the interior of $S_{g,b,0}$.

3.1.1 Relatedness of curves

We begin with the concept of relatedness of two non-separating simple closed curves in the interior of $S_{g,b,0}$.

Definition 3.1. *Two non-separating simple closed curves α, β in the interior of $S_{g,b,0}$ are* ***related*** *if there is a combination of isotopies and Dehn twists f_i transforming one into the other.*

Remark 3.1. Relatedness of curves is an equivalence relation. Any curve α is clearly related to itself under the identity, showing reflexivity. For symmetry, observe that at any given t-value, isotopies and Dehn twists are diffeomorphisms onto their image and therefore invertible. Thus, if α is related to β by a sequence g_i of Dehn twists and isotopies, it is possible to construct an inverse sequence f_i from β to α. Finally, transitivity follows from the fact that the composition of the final element of one sequence of Dehn twists and isotopies with the first element of another will always be either a Dehn twist or isotopy. Thus, given a sequence f_i from α to β and g_k from β to γ, it is possible to compose the final f_i with the first g_k on their shared domain β to obtain a sequence h_j from α to γ.

We will now prove an important lemma regarding the non-separating simple closed curves in the interior of $S_{g,b,0}$.

Lemma 3.3. *Any two non-separating simple closed curves in the interior of $S_{g,b,0}$ are related.*

Proof of Lemma 3.3. Let α, β be two non-separating simple closed curves in the interior of $S_{g,b,0}$. We can apply a sequence of isotopies to α and β to obtain curves α' and β' (related to α and β, respectively) such that α' and β' are in transverse position.

Now suppose that α' and β' intersect at k points for some $k \in \mathbb{Z}_{\geqslant 0}$. Note that the value of k depends solely on the original curves α and β, and not on the choice of isotopy to obtain α' and β', as the transversal intersection of any two curves is unique up to homotopy (and therefore isotopy) by §1.1.4 and §1.2.2. To conclude that α is related to β, it suffices to show that α' is related to β' by transitivity of relatedness established in Remark 3.1.

There are precisely three cases: either $k = 0$, $k = 1$, or $k \geqslant 2$.

Case 1: Suppose $k = 0$, i.e. that α' and β' do not intersect. Then because α, β (and therefore α', β') are non-separating, by the change of coordinates principle and Theorem 2.2

there exists another curve, say γ, such that the intersection number $i(\alpha', \gamma) = i(\beta', \gamma) = 1$. Then α' and β' are each related to γ. By transitivity and symmetry of relatedness from Remark 3.1, it follows that α' and β' are related.

Case 2: Suppose $k = 1$. Then α' and β' are related by a sequence of two Dehn twists. First we take the Dehn twist of β' along the curve α' to obtain a curve $T_{\alpha'}(\beta) = \beta''$ that is related to β'. Subsequently, we take the Dehn twist of this new curve β'' along β', which yields our second curve $\alpha' = T_{\beta'}(T_{\alpha'}(\beta'))$. This sequence of two Dehn twists transforming β' into α' is visualized in Figure 3.1.

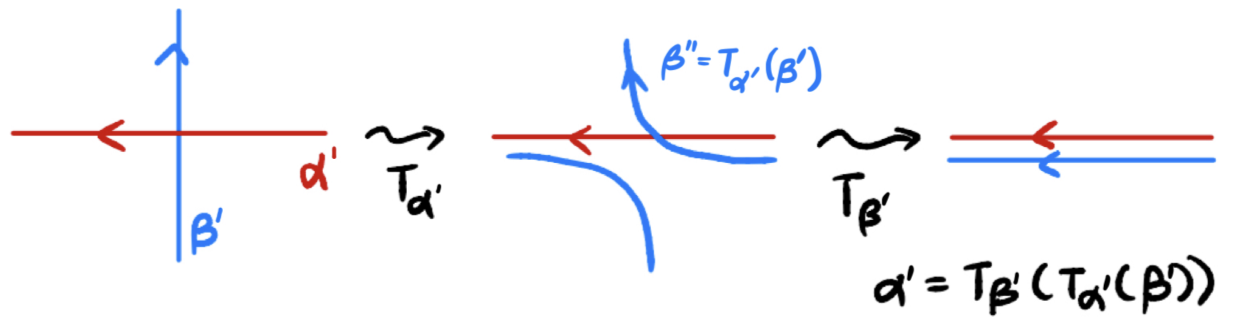


Figure 3.1: Sequence of two Dehn twists in the $k = 1$ case

Case 3: Suppose $k \geqslant 2$. Our aim is to reduce this to the $k \leqslant 1$ cases that we have already dealt with above. We first select any two consecutive intersection points of the curves α', β'. We can find a curve γ intersecting α', β' transversely in $< k$ points. Then there are two sub-cases: either γ intersects β' once, or it does not intersect β' at all.

Sub-case (a): If γ intersects β' once, then the curve γ is also non-separating.

Sub-case (b): If γ does not intersect β', then we have two possibilities for γ, say γ_1 and γ_2. Because β' is non-separating, it follows that one of the γ_i must also be non-separating. These two sub-cases are visualized in Figure 3.2.

Thus, in all possible cases, we have a non-separating simple closed curve γ that intersects α', β' at $< k$ points. To conclude, we can induct on k until we have $k \leqslant 1$ and apply the results from Cases 1 and 2 above to show that α', β' are each related to γ. It follows by transitivity of relatedness of curves from Remark 3.1 that α', β' are related. This concludes the proof.

□

3.1.2 Relatedness of arcs

We will see that Lemma 3.3 is very useful in thinking about curves within the interior of a surface $S_{g,b,0}$. However, as $S_{g,b,0}$ can have a nontrivial number of boundary components, we need an analogous concept to think about what lies on its boundary components. This provides the motivation for our next step in the proof of Theorem 3.1: introducing and establishing the analogous result for arcs.

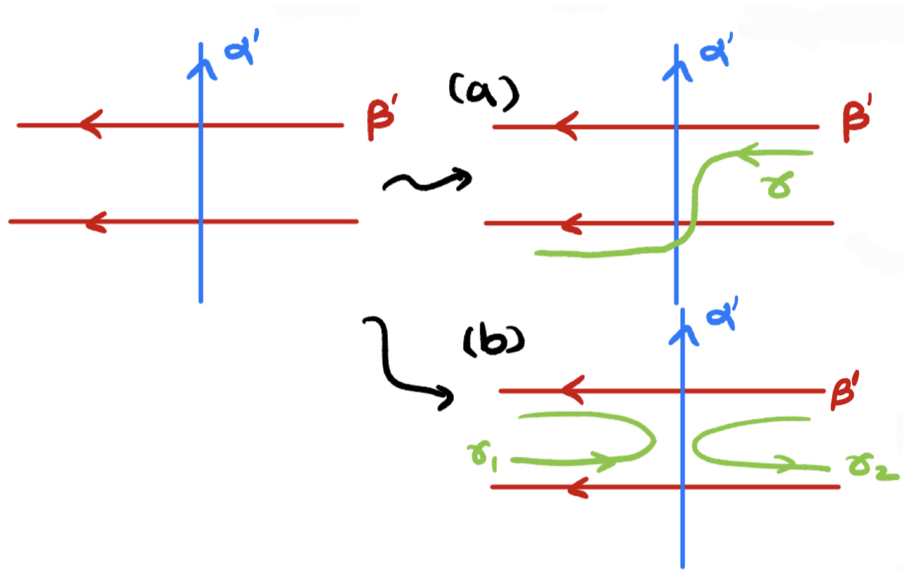


Figure 3.2: Two possibilities for γ in the $k \geqslant 2$ case

Consider any surface $S_{g,b,0}$ with at least two distinct boundary components, i.e. with $b \geqslant 2$. Fix any two points r, q in two distinct boundary components of $S_{g,b,0}$, and consider all properly embedded arcs in $S_{g,b,0}$ with endpoints r, q. We offer a definition of relatedness of arcs that closely follows the definition for relatedness of non-separating simple closed curves in Definition 3.1.

Definition 3.2. *Take any two properly embedded arcs α, β in $S_{g,b,0}$ with endpoints r, q for r, q as defined above. We say that α and β are **related** if there exists a combination of Dehn twists and isotopies transforming α to β.*

Remark 3.2. Observe that relatedness of arcs as given in Definition 3.2 is also an equivalence relation. The proof is almost directly the same as the proof offered in Remark 3.1, as we likewise have that an arc is related to itself (showing reflexivity), that we can obtain a reverse sequence of Dehn twists and isotopies (showing symmetry), and that we can compose the first and final Dehn twists and/or isotopies of two sequences (showing transitivity).

Having established this definition, we will state and prove the analogous result to Lemma 3.3.

Lemma 3.4. *For a finite-type surface $S_{g,b,0}$ with $b \geqslant 2$, and any two points r, q that lie on distinct boundary components of $S_{g,b,0}$, the arcs in $S_{g,b,0}$ with endpoints at r, q are all related.*

As with the proof of Lemma 3.3, our proof of Lemma 3.4 will also involve casework based on the number of intersection points.

Proof of Lemma 3.4. Pick any two properly-embedded arcs α, β with endpoints r, q. Once again, we can apply a sequence of isotopies to α, β respectively to obtain arcs α', β' that are in transverse position. Observe that by definition, we have that α is related to α' and β is related to β'.

The modified arcs α' and β' intersect at their common endpoints r and q, and (possibly) also intersect transversely at k interior points for some $k \in \mathbb{Z}_{\geqslant 0}$. Observe that, as in the proof of Lemma 3.3, the value of k depends only on the initial choice of arcs α, β (and not on the sequence of isotopies to obtain α', β' since transversal intersection is unique up to isotopy. Once again, by transitivity of relatedness of arcs as an equivalence relation, given in Remark 3.2, it suffices to show that α' and β' are related to conclude that α and β are related.

There are precisely two cases: either $k = 0$ or $k > 0$. These are outlined in Figure 3.3.

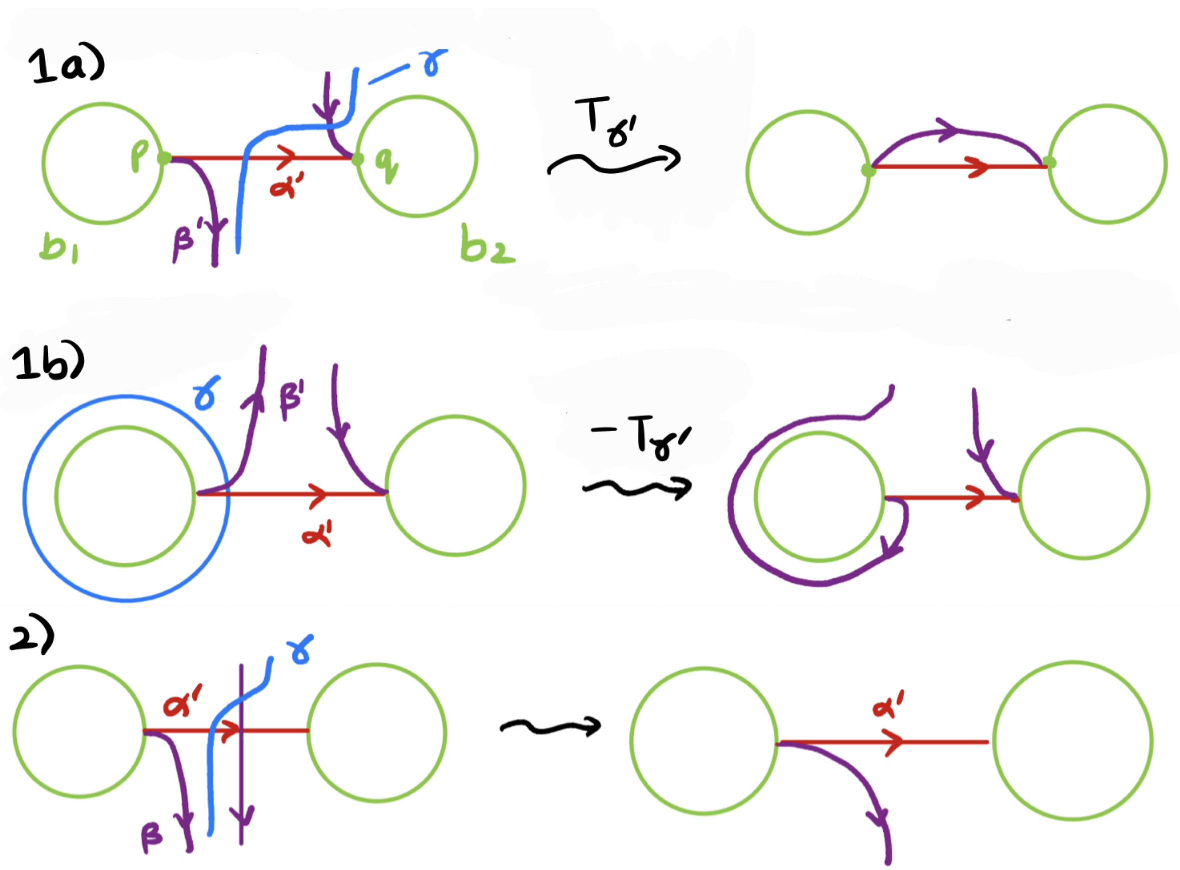


Figure 3.3: Two possible cases (with two sub-cases of the first case): either $k = 0$ or $k > 0$.

Case 1: If $k = 0$, i.e. the arcs α', β' do not intersect anywhere other than their common endpoints r, q, then there are two possible sub-cases depending on the orientation of the arcs: either the orientation of β' is coherent (i.e. when an orientation is specified at a point on β', the same orientation is specified on an arbitrary neighborhood of that point), or the orientation of β' is not coherent. This definition is elaborated in [Mor98, Chapter 1, p.46].

Sub-case (a): If the orientation is coherent, then we can obtain a non-separating simple closed curve γ in the interior of $S_{g,b,0}$ that intersects α' and β' as shown in Figure 3.3. By performing a Dehn twist of β' along this curve, we can transform β' into α', i.e. we have that $T_\gamma(\beta') = \alpha'$. Thus, β' and α'' are related.

Sub-case (b): If the orientation is not coherent, we define a nonseparating simple closed curve γ in the interior of $S_{g,b,0}$ to run parallel to one of the two boundary components. Then with a negative Dehn twist of β' along γ, we can transform β' to obtain a new curve β'' that has coherent orientation, i.e. we have $-T_\gamma(\beta') = \beta''$. Note that β'' is clearly related to β'. But now, considering β'' and α, we have precisely the same set-up as Sub-case (a), so it follows that β'' and α' are related. Then by transitivity of relatedness of arcs by Remark 3.2, it follows that β' and α' are related.

Case 2: If $k > 0$, i.e. the arcs α' and β' intersect transversely in the the interior at least once, then we look at the first of these intersection points as in Figure 3.3. Once again, there are two possibilities: either the orientation is coherent or it is not coherent.

Figure 3.3 shows the coherent case. In this case, we can obtain a non-separating simple closed curve γ in the interior of $S_{g,b,0}$ such that a Dehn twist of β' along γ removes the intersection point, thus decreasing the overall value of k by one: in other words, $T_\gamma(\beta') = \beta''$ where β'' transversely intersects α at $k-1$ points in the interior of $S_{g,b,0}$. We can repeat this process inductively until we obtain a modified arc $\beta*$ that does not intersect α' transversely, i.e. the $k = 0$ case. We then apply the results from Case 1 to conclude that α' and $\beta*$ are related. Since $\beta*$ is clearly related to β' by the inductive sequence of Dehn twists detailed above, it follows by transitivity of relatedness of arcs in Remark 3.2 that α' and β' are related as desired.

In the case where the orientation is not coherent, we can apply a similar process to Sub-case (b) of Case 1 to reduce this to the coherent case: we likewise select a non-separating simple closed curve γ in the interior of $S_{g,b,0}$ that is parallel to one of the boundary components and take a negative Dehn twist of β' along γ to obtain a new curve $\beta'' = -T_\gamma(\beta')$ that yields a coherent orientation. Then by the above discussion of the coherent case, it follows that β'' and α' are related. Since β' and β'' are clearly related, it again follows by Remark 3.2 that α' and β' are related.

Thus in all possible cases, α' and β' are related, so we conclude. □

3.2 Proving the Special Case

We are now ready to prove Theorem 3.1. This section offers a proof of Thoerem 3.1, and then moves to discuss some limitations of the proof and its findings for the general case.

In fact, Lemma 3.3 and Lemma 3.4 regarding the relatedness of curves and arcs in $S_{g,b,0}$ provide us with majority of the technical machinery we need for this proof. The proof uses casework based on the value of the genus g to reduce the special case to the cases of these two lemmas that we have already proved.

Proof of Theorem 3.1. Take a mapping class of $\mathrm{MCG}(S_{g,b,0})$, and let $\varphi : S_{g,b,0} \circlearrowleft$ be a representative element. To conclude that $\mathrm{MCG}(S_{g,b,0})$ is genereated by Dehn twists, it suffices to show that any such φ can be obtained by a sequence of Dehn twists and isotopies.

Now there are precisely two cases, depending on the genus of $S_{g,b,0}$: either $g = 0$ or $g > 0$.

Case 1: If $g = 0$, we will proceed by inducting on b. We know from Example 2.4 that $\mathrm{MCG}(S_{0,1,0})$ (i.e. the mapping class group of the disk) is trivial. Then to restrict to the nontrivial case, suppose $b \geqslant 2$, i.e. that $S_{0,b,0}$ has at least two distinct boundary components. Let p, q be two points that lie in distinct boundary components of $S_{0,b,0}$, and let α be an arc connecting p and q. Observe that φ must preserve boundary points, so $\varphi(\alpha)$ is also an arc from p to q. Then by Lemma 3.4, we have that α and $\varphi(\alpha)$ are related as they are both arcs with endpoints p, q.

Composing Dehn twists and isotopies, we can suppose that φ is the identity of α, and is thus also the identity along a tubular neighborhood of α. We proceed to cut $S_{0,b,0}$ along α to obtain $S_{0,b-1,0}$. Restricting to this new domain, we see that φ becomes a self-diffeomorphism of $S_{0,b-1,0}$ that preserves orientation and fixes the boundary point-wise. We now repeat this process inductively $b-1$ times. The new version of φ is generated by Dehn twists and isotopies. But then, so too is the original φ, yielding the desired result.

Case 2: If $g > 0$, we will induct on g to reduce this to the $g = 0$ case resolved above. Let α be a non-separating simple closed curve in the interior of $S_{g,b,0}$. Note that by definition, $\varphi(\alpha)$ must also be a non-separating simple closed curve in the interior of $S_{g,b,0}$. Then by Lemma 3.3, α and $\varphi(\alpha)$ are related. Once again, we can suppose up to isotopies and Dehn twists that φ is the identity on α. We proceed to cut $S_{g,b,0}$ along α to obtain $S_{g-1,b+2,0}$. Repeating this process inductively $g-1$ times, we reduce to the $g = 0$ case, which we have already resolved above.

Thus in both possible cases, φ is generated by a combination of Dehn twists and isotopies. We conclude that $\text{MCG}(S_{g,b,0})$ is generated by Dehn twists. □

We end with some remarks about the limitations of this result. In this chapter, we restricted ourselves to considering the mapping class groups of compact (i.e. non-punctured surfaces) $S_{g,b,0}$. The question remains: does Theorem 3.1 hold for the general case, where we could have any number of punctures on the finite-type surface?

In fact, this result does *not* hold in general when we have a surface $S_{g,b,p}$ that has a nontrivial number of punctures. This is because composing Dehn twists cannot permute the puncture(s) of a given surface $S_{g,b,p}$. Since any representative element of a mapping class must send the puncture(s) of $S_{g,b,p}$ to other punctures (and not to boundary components or interior points), we cannot get obtain isotopy class of orientation-preserving self-diffeomorphisms that fix the boundary point-wise via Dehn twists. Thus, we cannot generate the entire mapping class group. In order to deal with the general case, we will have to modify Theorem 3.1—which is the task of Chapter 4.

Chapter 4

Generating the Mapping Class Group: The General Case

The aim of this chapter is to extend the results from Chapter 3 to the general case, analyzing the mapping class groups of surfaces that are not necessarily compact. As discussed at the end of Chapter 3, we cannot immediately generalize Theorem 3.1 or Theorem **??** (Dehn-Lickorish) to non-compact surfaces. Rather, dealing with the general case requires us to think carefully about how self-diffeomorphisms act on the punctures of a finite-type surface.

Our aim is to prove a modified version of the statement of Theorem 3.1. Rather than analyze the generators of the mapping class group of $S_{g,b,p}$ directly, we will restrict ourselves to a special subgroup of the mapping class group called the pure mapping class group, which consists of elements of the mapping class group that fix all punctures p of $S_{g,b,p}$. We will demonstrate that the pure mapping class group of any finite-type surface is in fact generated by finitely-many Dehn twists about non-separating simple closed curves, and conclude that the special case follows as a corollary from the general case.

We develop this argument in three stages. First, we introduce additional algebraic and geometric structure that builds upon the results from Chapters 1 and 2 and provides an analogue to the concept of 'relatedness' of curves and arcs in compact surfaces from §3.1 that applies to punctured surfaces. Second, we prove the Birman Exact Sequence to induct upon the number of punctures of a finite-type surface. Finally, we prove of the general case, which will rely on a double-induction argument on the genus and number of punctures of a surface. We conclude with some remarks on the relationship between the special case and the general case and ongoing efforts to specify the generators of the mapping class group of a finite-type surface $S_{g,b,p}$ with further precision.

This chapter mostly follows [FM11, Chapter 4], which are based on results from Birman's doctoral dissertation that were later published in [Bir74]. However, while [FM11, Chapter 4] works in the topological category, we will continue to work in the smooth category. Most results can be directly extended to the smooth category due to our choice of the $\mathcal{C}^\infty$ topology for the mapping class group in §2.1.1. We also prove additional algebra and algebraic topology lemmas to elaborate on missing links in the proofs, drawing on [Hat01] and preliminary results from Chapters 1 and 2.

4.1 Additional Algebraic and Geometric Structure

We begin by providing additional algebraic and geometric structure aimed at analyzing how elements of the mapping class group act on punctures of a surface $S_{g,b,p}$. To this end, we introduce the pure mapping class group as a subgroup of the mapping class group. We then introduce the complex of curves, a simplicial complex structure whose vertices correspond to curves in the surface and edges correspond to sequences of isotopies and Dehn twists transforming one curve into another. Finally, we prove a series of lemmas regarding the connectedness of this simplicial complex that serve as an analogue to lemmas regarding the relatedness of curves and arcs in §3.1.

4.1.1 The pure mapping class group

The pure mapping class group is a special subgroup of the mapping class group that enables us to deal specifically with punctured surfaces.

Definition 4.1. *Given a surface $S_{g,b,p}$, the* ***pure mapping class group*** $\mathrm{PMCG}(S_{g,b,p})$ *is defined as the subset of the mapping class group* $\mathrm{MCG}(S_{g,b,p})$ *consisting of the isotopy classes of all orientation-preserving self-diffeomorphisms of $S_{g,b,p}$ that fix the boundary point-wise fix each puncture individually.*

Observe that $\mathrm{PMCG}(S_{g,b,p})$ clearly contains the equivalence class of the identity diffeomorphism, is closed under composition, and contains its own inverses. Thus, $\mathrm{PMCG}(S_{g,b,p})$ is indeed a subgroup of $MCG(S_{g,b,p})$.

Because any self-diffeomorphism must send punctures to punctures (and not to points in the interior or on the boundary of $S_{g,b,p}$), the mapping class group $\mathrm{MCG}(S_{g,b,p})$ acts on the set of the punctures of $S_{g,b,p}$ by permutation. This yields a useful short exact sequence linking the pure mapping class group and the mapping class group.

Lemma 4.1. *Given any finite-type surface $S_{g,b,p}$, the sequence*

$$0 \to \mathrm{PMCG}(S_{g,b,p}) \xrightarrow{i} \mathrm{MCG}(S_{g,b,p}) \xrightarrow{\psi} S_p \to 0$$

is exact, where S_p is the symmetric group of degree p containing all permutations of elements of the set $\{1, ..., p\}$, i is the inclusion, and ψ is the group action of $\mathrm{MCG}(S_{g,b,p})$ *on the set of all punctures. In other words, if f is a representative element of some mapping class in* $\mathrm{MCG}(S_{g,b,p})$*, ψ sends the equivalence class of f to the corresponding permutation of the indices in S_p. For instance, if f sends punctures $x_1 \mapsto x_{i_1}, x_2 \mapsto x_{i_2}, \ldots, x_p \mapsto x_{i_p}$, then $\psi([f]) \in S_p$ sends $1 \mapsto i_1, \ldots, p \mapsto i_p$.*

Proof of Lemma 4.1. This follows from checking definitions. Clearly $\mathrm{PMCG}(S_{g,b,p})$ is its own image under the inclusion i. To see that $\mathrm{im}(i) = \ker(\psi)$, observe that the kernel of ψ is the set of all mapping classes whose representatives are the identity on the indices of punctures, i.e., maps fixing all punctures—precisely the pure mapping class group $\mathrm{PMCG}(S_{g,b,p})$ itself. □

Remark 4.1. There are precisely two cases where the pure mapping class group in equal to the mapping class group. The first case is the trivial case where $p = 0$. The second case is where $p = 1$: since self-diffeomorphisms must send punctures to punctures, any representative of a mapping class in $\mathrm{MCG}(S_{g,b,1})$ must send the one puncture of $S_{g,b,p}$ to itself, thereby fixing the puncture. Then it is also an element of the pure mapping class group, so $\mathrm{PMCG}(S_{g,b,1}) = \mathrm{MCG}(S_{g,b,1})$.

4.1.2 The complex of curves

Having introduced additional algebraic structure with the pure mapping class group, we now introduce additional geometric structure by defining a simplicial complex on $S_{g,b,p}$.

Definition 4.2. *Given a surface $S_{g,b,p}$, the* ***complex of curves*** *$\mathcal{C}(S_{g,b,p})$ is an abstract simplicial complex associated to a surface $S_{g,b,p}$. Its 1-skeleton is given as follows:*

Vertices (0-cells): Each isotopy class of essential simple closed curves in S corresponds to exactly one vertex of $\mathcal{C}(S_{g,b,p})$.

Edges (1-cells): Given any two vertices of $\mathcal{C}(S_{g,b,p})$ corresponding to isotopy classes of curves with representative elements α and β, we say that there is an edge between these two vertices if the geometric intersection number $i(\alpha, \beta) = 0$. Recall that the geometric intersection number is unique up to isotopy, so this is well-defined.

It is possible to continue defining the n-skeleton of this complex inductively to obtain a higher-dimensional simplicial complex. However, for our purposes, we will deal exclusively with this 1-dimensional complex of curves. Introducing this simplicial complex structure proves to be immediately useful in reducing the general case to the case where $b = 0$.

Remark 4.2. Within a simplicial complex structure $\mathcal{C}(S_{g,b,p})$, a puncture and a boundary component of $S_{g,b,p}$ have the exact same effect: by definition, a simple closed curve that is homotopic to either a puncture or a boundary is not essential, and thus does not correspond to a vertex in the complex of curves. Therefore, we can without loss of generality assume that $b = 0$ and deal only with nontrivial values of p, thereby reducing the general case to surfaces of genus g with no boundary components and p punctures.

For the remainder of this chapter, we assume all surfaces have 0 boundary components, which will enable us to simplify our casework and inductive arguments for the general case.

We now move to introduce additional simplicial complexes based on the complex of curves and prove certain useful lemmas regarding the connectedness of this simplicial complex structure.

Definition 4.3. *Given a finite-type surface $S_{g,b,p}$, let $\mathcal{N}(S_{g,b,p})$ be the subcomplex of $\mathcal{C}(S_{g,b,p})$ spanned by the vertices corresponding to non-separating simple closed curves. We call this the* ***complex of non-separating curves****.*

The complex of non-separating curves is an intermediate simplicial complex, which allows us to define another simplicial complex that we will use for the majority of the proofs in this chapter called the modified complex. We define this below.

Definition 4.4. *Given a finite-type surface $S_{g,b,p}$, let $\hat{\mathcal{N}}(S_{g,b,p})$ denote the 1-dimensional simplicial complex whose vertices are isotopy classes of non-separating simple closed curves in $S_{g,b,p}$ and whose edges correspond to pairs of isotopy classes α, β with geometric intersection number $i(a, b) = 1$. We call this the* ***modified complex****.*

Note that $\hat{\mathcal{N}}(S_{g,b,p})$ is *not* a subcomplex of either $\mathcal{N}(S_{g,b,p})$ or $\mathcal{C}(S_{g,b,p})$. While its set of vertices is a subset of the vertices of $\mathcal{C}(S_{g,b,p})$, its edges are completely different: in the modified complex, we take pairs of vertices with geometric intersection number 1 rather than 0. We now investigate the conditions under which this complexes are connected.

Theorem 4.2. *If $g \geqslant 2$ and $p \geqslant 0$, then $\hat{\mathcal{N}}(S_{g,b,p})$ is connected.*

Before proceeding with the proof of Theorem 4.2, it is worth briefly discussing the motivation for thinking about the connectedness of the complex of curves and its subcomplexes. The reader might ascertain from the definition of this simplicial complex structure that there is a close relationship between this structure and the concept of 'relatedness' introduced in Chapter 3. We remark on this in more detail below.

Remark 4.3. If the complex of non-separating curves $\mathcal{N}(S_{g,b,p})$ is connected—i.e., there is a path joining any two vertices of the complex—then all non-separating simple closed curves in $S_{g,b,p}$ are related by Definition 3.1. This follows by definition of the edges and vertices of the complex. A 'path' between two vertices with representatives α and β is a sequence of adjacent edges, which corresponds to a sequence of isotopy classes $\alpha = \gamma_1, ..., \gamma_k = \beta$ such that the geometric intersection number $i(\gamma_j, \gamma_{j+1}) = 0$ for all j. But by the definition in Chapter 3, this implies that the curves α and β are related by a sequence of Dehn twists and isotopies.

We now build towards the proof of Theorem 4.2 by stating the two lemmas pertaining to the connectedness of the complex of curves and the complex of non-separating curves. The proofs for both lemmas are quite similar, relying upon casework and induction on the values of g and p. In the interest of concision, we assume the result of the first lemma and only provide a detailed proof of the second lemma.

Lemma 4.3. *If $3g + p \geqslant 5$, then the complex of curves $\mathcal{C}(S_{g,b,p})$ is connected.*

Proof of Lemma 4.3. See [FM11, §4.1.1, p.92-93]. □

Lemma 4.4. *If $g \geqslant 2$, then the complex of non-separating curves $\mathcal{N}(S_{g,b,p})$ is connected.*

Proof of Lemma 4.4. Suppose $g \geqslant 2$ and consider any finite-type surface $S_{g,b,p}$. We proceed in two steps. First, we will consider the special case where $p \leqslant 1$. Second, we will induct on p to prove the general case.

Step 1: Suppose that $p \leqslant 1$. Let α, β be arbitrary isotopy classes of non-separating simple closed curves in $S_{g,b,p}$, i.e., two vertices within $\mathcal{N}(S_{g,b,p})$ (and thus within the main complex of curves). To conclude, we need to show that there is a path through the edges of $\mathcal{N}(S_{g,b,p})$ joining these two vertices, i.e., a sequence of isotopy classes of non-separating simple closed curves $\alpha = \gamma_1, ..., \gamma_n = \beta$ such that the geometric intersection number $i(\gamma_i, \gamma_{i+1}) = 0$ for all i.

There are two possible cases: either $p = 0$ or $p = 1$. In both cases, we have that $3g+p \geqslant 5$ (as $3g \geqslant 6$ since $g \geqslant 2$). Thus, by Lemma 4.3, the complex of curves $\mathcal{C}(S_{g,b,p})$ is connected. Then by definition, there exists a sequence of isotopy classes of curves $\alpha = c_1, ..., c_n = \beta$ such that the geometric intersection number $i(c_j, c_{j+1}) = 0$, giving a 'path' from edges joining the two vertices α, β within the complex of curves.

To conclude, we modify this sequence such that each $\{c_j\}$ is non-separating so that this 'path' between α and β is contained in the complex of non-separating curves. Suppose c_j is separating for some value of j. Let $d_j \in c_j$ be a representative curve, and let S and S' be the two components of $S_{g,b,p} - d_j$ (i.e., the two sub-surfaces obtained by cutting $S_{g,b,p}$ along the separating curve d_j). Since $g \geqslant 2$ and $n \leqslant 1$ by assumption, both S and S' have positive genus. Now there are precisely two cases: either c_{j-1}, c_{j+1} have representatives lying on different sub-surfaces (i.e., one lies on S and the other lies on S'), or they have representatives lying on the same sub-surface.

Case 1: Suppose c_{j-1} and c_{j+1} have representatives lying on different sub-surfaces. Then the geometric intersection number $i(c_{j-1}, c_{j+1}) = 0$, so we can delete c_j from the sequence.

Case 2: Suppose c_{j-1} and c_{j+1} have representatives that both lie on the same sub-surface. Without loss of generality, suppose they both have representatives lying on S. Because S' has positive genus, we can identify a non-separating simple closed curve in S', say ρ (as shown in Figure 4.1).

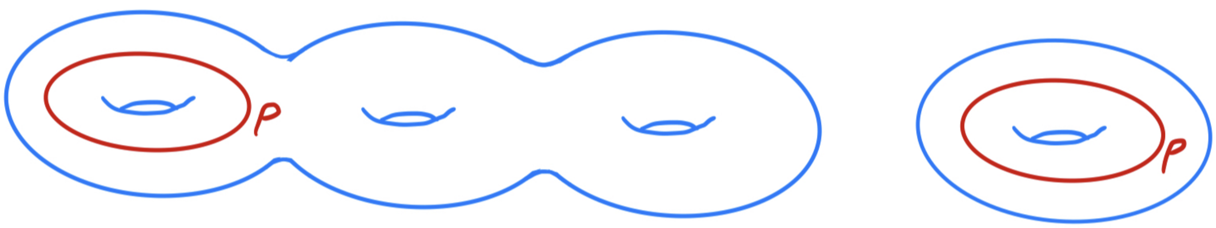


Figure 4.1: Example of obtaining a non-separating simple closed curve ρ in genus 3 and genus 1 surfaces (which can be generalized for higher genera).

Since ρ lies on a different subsurface to representatives of both c_{j-1} and c_{j+1}, it follows that the geometric intersections $i(c_{j-1}, \rho) = i(\rho, c_{j+1}) = 0$. Then, we can replace c_j with the isotopy class of ρ in the sequence.

By repeating this process inductively for each separating c_j in the sequence, we obtain a modified sequence consisting of isotopy classes of non-separating curves. This proves the $p \leqslant 1$ case.

Step 2: We will now induct on p to conclude for the general case. Suppose $p \geqslant 2$, and take any two isotopy classes of non-separating simple closed curves α and β, i.e., vertices in the subcomplex $\mathcal{N}(S_{g,b,p})$. Once again, observe that $3g + p \geqslant 5$, and thus by Lemma 4.3, we have a 'path' in the complex of curves from α to β given by $\alpha = c_1, ...c_n = \beta$ where the c_j are isotopy classes of curves in the surface such that the geometric intersection $i(c_j, c_{j+1}) = 1$ for all j.

As with Step 1, we will modify this sequence such that all the c_j are non-separating in order to conclude. If we follow the process for modifying the sequence detailed in Cases 1 and 2 of Step 1, the only issue we could encounter is if representatives of c_j and c_{j+1} lie on one of the two sub-surfaces, and the other of the two sub-surfaces has genus 0: without loss of generality, say that representatives of c_j and c_{j+1} lie on S and that S' has genus 0 (where S, S' are again given by cutting $S_{g,b,p}$ along a separating curve c_j in the sequence). In this case, we cannot follow the process from Step 1 to obtain a non-separating simple closed curve ρ in S' with which we can replace the separating c_j.

But then observe that S must have genus $g \geqslant 2$ and fewer punctures than the overall surface $S_{g,b,p}$, so $3g + p \geqslant 5$. Then we can apply Lemma 4.3 to the subsurface S to conclude that there is a 'path' in $\mathcal{C}(S)$ between the vertices that correspond to c_j and c_{j+1}. This path corresponds to a sequence of isotopy classes of non-separating simple closed curves in S (and thus, in the main surface $S_{g,b,p}$). We replace c_j with this sequence to conclude. □

We now apply these lemmas to prove Theorem 4.2.

Proof of Theorem 4.2. Suppose $g \geqslant 2$ and $p \geqslant 0$. Let α and β be two isotopy classes of non-separating simple closed curves in $S_{g,b,p}$, i.e. two vertices in the modified complex $\hat{\mathcal{N}}(S_{g,b,p})$ (and therefore also in $\mathcal{N}(S_{g,b,p})$ and $\mathcal{C}(S_{g,b,p})$).

By Lemma 4.4, we know $\mathcal{N}(S_{g,b,p})$ is connected. Thus, there exists a path in $\mathcal{N}(S_{g,b,p})$ between the vertices α and β, i.e., a sequence of isotopy classes of non-separating simple closed curves $\alpha = c_1, ..c_n = \beta$ such that the geometric intersection $i(c_i, c_{i+1}) = 0$ for all i. As with the proof for Lemma 4.4, we will modify this sequence so that this path is contained entirely within $\hat{\mathcal{N}}(S_{g,b,p})$ in order to conclude that the modified complex is connected.

But this follows directly from the change of coordinates principle from §1.1.4. For every i, there exists an isotopy class d_i of non-separating simple closed curves such that $i(c_i, d_i) = i(d_i, c_{i+1}) = 1$. Then the sequence $\alpha = c_1, d_1, c_2, ...c_{n-1}, d_{n-1}, c_n = \beta$ yields the desired path from α to β within $\hat{\mathcal{N}}(S_{g,b,p})$. □

4.2 The Birman Exact Sequence

The previous section introduced algebraic and geometric structures to study the mapping class group act on punctured surfaces. Yet we need a way to induct on p, the number of punctures, in order to develop a double-induction argument to prove the general case. This requires us to think carefully about the relationship between the pure mapping class group and mapping class group when there are a nontrivial number of punctures—and how puncturing a surface affects its mapping classes in the first place.

This section will provide the last missing link to prove the general case, based on Birman's dissertation. First, we introduce the *forgetful map*, a homomorphism that includes a punctured surface into a non-punctured surface. Next, we introduce the *push map*, a homomorphism of a punctured surface through which we can analyze the kernel of the forgetful map, and consider how the push map operates in terms of Dehn twists. Finally, we draw these concepts together by proving the Birman exact sequence, a short exact sequence that will allow us to induct on the number of punctures.

4.2.1 The forgetful map

We begin with any surface that is possibly punctured (i.e. the value of p can be either trivial or nontrivial), but has no marked points. Select any point x in the interior of $S_{g,b,p}$. Let $(S_{g,b,p}, x)$ denote the surface obtained from $S_{g,b,p}$ by marking the interior point x. An example of this is illustrated in Figure 4.2.

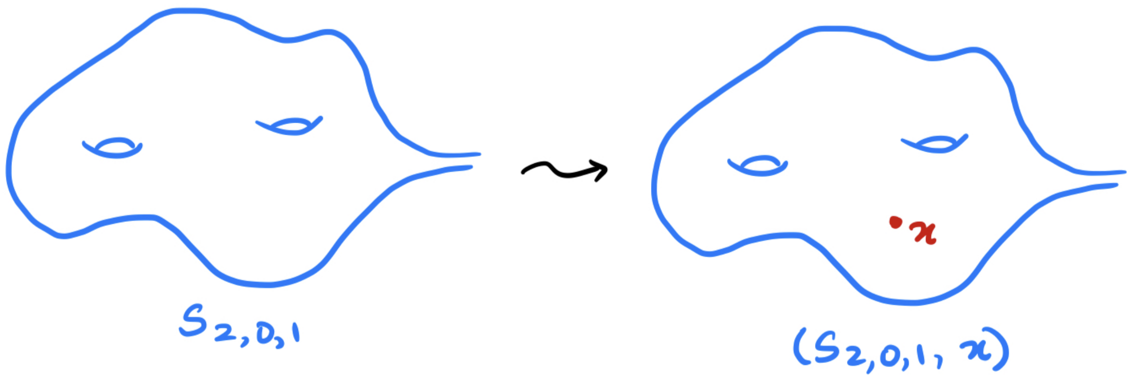


Figure 4.2: Marking a point x in the interior of the surface $S_{2,0,1}$.

Definition 4.5. *With $S_{g,b,p}$ and $(S_{g,b,p}, x)$ as defined above, there exists a natural homomorphism*

$$\mathcal{F}orget : \mathrm{MCG}((S_{g,b,p}, x)) \to \mathrm{MCG}(S_{g,b,p})$$

called the ***forgetful map****. This is the map given by 'forgetting' that the point x in the interior of $S_{g,b,p}$ is marked. More precisely, take any representative φ of a mapping class in* $\mathrm{MCG}((S_{g,b,p}, x))$*. The map φ must send each point (a, x) to some other point (b, x) within $S_{g,b,p}$. Then the forgetful map sends the isotopy class of φ to the isotopy class of the map φ' in* $\mathrm{MCG}(S_{g,b,p})$ *that sends a to b in $S_{g,b,p}$*

In fact, the forgetful map is closely related to how we study self-diffeomorphisms of a punctured surface. As discussed in §2.3.2, 'marking' a point x in the interior of a surface $S_{g,b,p}$ is roughly analagous to deleting x from $S_{g,b,p}$, i.e., to making a new puncture to $S_{g,b,p}$. We remark on this below.

Remark 4.4. The forgetful map is surjective. For any orientation-preserving self-diffeomorphism φ of $S_{g,b,p}$, we can modify φ by isotopy to a new orientation-preserving self-diffeomorphism φ' that fixes x. Then the isotopy class $[\varphi] = [\varphi'] \in \mathrm{MCG}(S_{g,b,p})$ and $\mathcal{F}orget([\varphi']) = [\varphi]$.

Remark 4.5. The forgetful map is induced by the inclusion $S_{g,b,p} - x \to S_{g,b,p}$. This is because the mapping class group $\mathrm{MCG}(S_{g,b,p}, x)$ is isomorphic to a subgroup G of the mapping class group $\mathrm{MCG}(S_{g,b,p} - x)$ that preserves the puncture that comes from removing the interior point x from the surface. Then the forgetful map can be interpreted as a map from $G \to \mathrm{MCG}(S_{g,b,p})$ that 'fills in' the puncture at x—the same as 'forgetting' that the point x is marked. But this is precisely the map induced by the inclusion.

These remarks provide an insight into the structure of the forgetful map and the geometric motivation behind its definition. In particular, Remark 4.5 suggests that the process of defining the forgetful map can be understood as puncturing a surface and then taking the inclusion of the newly-punctured surface into the original surface. This brings us one step closer to being able to induct on the number of punctures of a surface via a short exact sequence—although we first need to consider the kernel of the map to set up and prove the sequence.

4.2.2 The push map

In this section, we introduce the push map, whose image, we shall see, is precisely the kernel of the forgetful map. In order to motivate our definition of the push map, we first consider what an element of the kernel of the forgetful map would look like.

Suppose f is a mapping class that lies within $\mathrm{MCG}((S_{g,b,p}, x))$ and is an element of the kernel of the forgetful map. Let φ be a representative of f. Under the inclusion of $\mathrm{MCG}((S_{g,b,p}, x)) \hookrightarrow \mathrm{MCG}(S_{g,b,p})$, we can think of φ as an orientation-preserving self-diffeomorphism $\bar{\varphi}$ of $S_{g,b,p}$. Because $\mathcal{F}orget(f)$ is trivial, there exists an isotopy from $\bar{\varphi}$ to the identity $\mathrm{id}_{S_{g,b,p}}$. Under this isotopy, the image of x traces a loop α in $S_{g,b,p}$ that is based in x, as we must return to x since the self-diffeomorphism fixes x. We want to show that by 'pushing' x along this path in the opposite direction, i.e., along the loop α^{-1}, we can recover the original element f in $\mathrm{MCG}(S_{g,b,p}, x)$. In other words, we want to extend what Birman calls an "isotopy of points" (the loop based in x) to an isotopy of the entire surface.

Before formalizing this idea, we first visualize what this means. Suppose there is a large table with a point x marked on its surface, with a clear plastic tablecloth that lies atop it. Someone places their finger right atop the point labelled x and pushes their finger in a loop (without lifting it off the table) until they return to where they began—right above the same point x marked on the tabletop. By moving their finger in this manner, not only did the position of their finger move in a loop (which we can analogize to the 'isotopy of points') but the entire fabric of the tablecloth moved with it (an isotopy of surface). The push map, defined below, allows us to formally describe this notion of 'pushing' in order to distort the entire surface.

Definition 4.6. *Suppose α is a loop in $S_{g,b,p}$ based in some interior point x. We can view $\alpha : [0,1] \to S_{g,b,p}$ as an 'isotopy of points' from x to itself.*

We now extend this to an isotopy of the surface $S_{g,b,p}$. Let φ_α be an orientation-preserving self-diffeomorphism of $S_{g,b,p}$ obtained at the end of the isotopy. We can regard φ_α as a self-diffeomorphism of $(S_{g,b,p}, x)$ and consider its isotopy class in $\mathrm{MCG}(S_{g,b,p}, x)$, which we will denote by $[\varphi_\alpha]_i$ in order to specify that this refers to equivalence classes under isotopy. Then we define the ***push map*** *$\mathcal{P}ush$ to be the map sending α to $[\varphi_\alpha]_i$.*

We want to show that the push map as discussed in Definition 4.6 is well-defined—i.e., that it is independent of the isotopy extension and choice of the 'loop' α based in x. This draws us back into the realm of topology. Showing that the push map is independent of the choice of the loop is the same as showing that it is independent of the choice of a representative within the homotopy class of α, which we will denote as $[\alpha]_h$, considered as an element of the fundamental group $\pi_1(S_{g,b,p}, x)$. We want to show that the push map $\mathcal{P}ush : \pi_1(S_{g,b,p}, x) \to \mathrm{MCG}(S_{g,b,p}, x)$ is well-defined along the equivalence relations of homotopy classes (in the domain) and isotopy classes (in the co-domain). In other words, the following diagram should commute:

$$\begin{array}{ccccc} \alpha \in & \text{all loops based in } x & \longrightarrow & \mathrm{Diffeo}^+((S_{g,b,p}, x)) & \ni \varphi_\alpha \\ \downarrow & \downarrow \pi_h & & \downarrow \pi_i & \downarrow \\ [\alpha]_h \in & \pi_1(S_{g,b,p}, x) & \xrightarrow{\mathcal{P}ush} & \mathrm{MCG}((S_{g,b,p}, x)) & \ni [\varphi_\alpha]_i \end{array}$$

Remark 4.6. It is by no means obvious that this map is well-defined. While we can extend an isotopy of a loop to an isotopy of an entire surface, we cannot necessarily extend a homotopy of a loop to a homotopy of an entire surface. We will prove that this map is well-defined in Section 4.2.3, when we state and prove the Birman exact sequence—from which the well-definedness of this map follows a corollary.

Before this, we make an important observation about how the push map can be interpreted in terms of Dehn twists, relating our study of the kernel of the forgetful map to the broader aim of this chapter: proving that the pure mapping class group is generated by Dehn twists and isotopies.

Remark 4.7. Given a simple loop α in $S_{g,b,p}$ based in x, consider an explicit representative of $\mathcal{P}ush(\alpha)$. We can identify a neighborhood of the loop α with the annulus $S^1 \times [0,2]$. We give the annulus the standard orientation on S^1 and $[0,2]$. Now, let the marked point x (the base of the loop α) correspond to the point $(0,1)$ on the annulus. Then we obtain the following isotopy of the annulus:

$$F((\theta,r),t) = \begin{cases} (\theta + 2\pi r t, r) & 0 \leqslant r \leqslant 1 \\ (\theta + 2\pi(2-r)t, r) & 1 \leqslant r \leqslant 2 \end{cases}$$

We can extend F by the identity to get an isotopy of the entire surface $S_{g,b,p}$. Then restricting F to $x \times [0,1]$, we get

$$F((0,1),t) = (2\pi t, 1),$$

i.e., an isotopy F that pushes the point x (base point of the loop α, marked point in $S_{g,b,p}$) along the core of the annulus $S^1 \times [0,2]$.

Let φ denote the self-diffeomorphism of $S_{g,b,p}$ induced by F at time $t = 1$, which is precisely the representative element of the mapping class that α is sent to under the push map. But φ is in fact the product of two Dehn twists: we can identify the boundary curve $S^1 \times 0$ of the annulus with a simple closed curve a in $(S_{g,b,p}, x)$ and $S^1 \times 2$ with a simple closed curve b in $(S_{g,b,p}, x)$. Then φ is isotopic to $T_a {T_b}^{-1}$, the product of the Dehn twist along a and the inverse Dehn twist along b. This is illustrated in Figure 4.3 below.

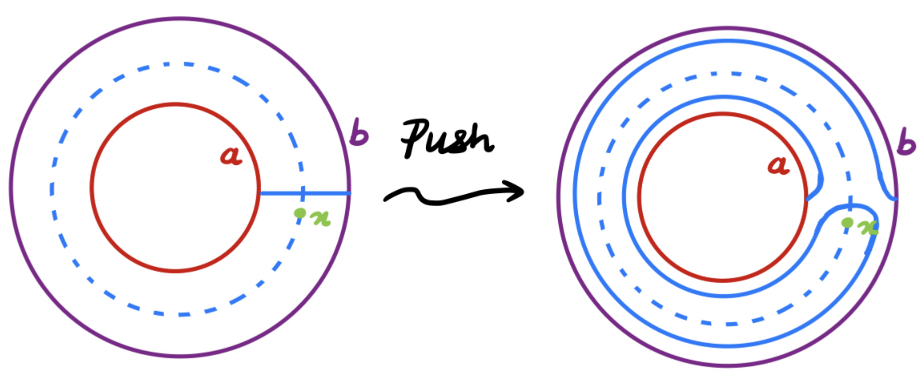


Figure 4.3: Visualizing the push map as a product of two Dehn twists

4.2.3 The short exact sequence

The previous sections defined the forgetful map and the push map and offered an intuitive understanding of the relationship between them: how the forgetful map can be understood as 'filling in' a puncture within a surface, and how the push map enables us to think about the kernel of the forgetful map and can be understood as a sequence of two Dehn twists. We now relate these notions algebraically through the Birman exact sequence.

Theorem 4.5 (Birman exact sequence)**.** *Let $S_{g,b,p}$ be a surface with negative Euler characteristic, possibly with punctures and/or boundary. Let $(S_{g,b,p}, x)$ be the surface obtained from $S_{g,b,p}$ by marking a point x in the interior of $S_{g,b,p}$. Then the sequence*

$$0 \to \pi_1(S_{g,b,p}, x) \xrightarrow{\mathcal{P}ush} \mathrm{MCG}((S_{g,b,p}, x) \xrightarrow{\mathcal{F}orget} \mathrm{MCG}(S_{g,b,p}) \to 0$$

is exact.

Note that the conditions for Theorem 4.5 to hold are slightly different from the conditions we have been operating under for the purposes of this chapter.

First, Theorem 4.5 holds for surfaces boundary as well as with punctures, i.e., finite-type surfaces $S_{g,b,p}$ with nontrivial values of b and/or p. For our purposes, following Remark 4.2, we have assumed without loss of generality that all surfaces in this chapter have no boundary components. It is nonetheless useful to know that Theorem 4.5 holds for cases where there are boundary components.

Second, Theorem 4.5 assumes that the Euler characteristic $\chi(S_{g,b,p})$ of the surface must be negative. While we have not been operating under this condition elsewhere in this thesis, it does not significantly limit the applicability of Theorem 4.5. Recall from Definition 1.1 that the Euler characteristic of a surface $S_{g,b,p}$ is given by $\chi(S_{g,b,p}) = 2-2g-b-p$. But this value will be negative for all surfaces in this chapter, as we have nontrivial values of g and p. We will see this explicitly in the proof of the $g = 0$ special case and the general case in §4.3.1.

The proof of Theorem 4.5 draws upon machinery from algebraic topology, specifically the long exact sequence for fiber bundles.

Proof of Theorem 4.5. Our proof proceeds in two steps. First, we prove that there is a fiber bundle where the total space is the space of all orientation-preserving self-diffeomorphisms of $S_{g,b,p}$, the base space is $S_{g,b,p}$ itself, and the fiber is the space of all orientation-preserving self-diffeomorphisms of $(S_{g,b,p}, x)$. Next, we apply the long exact sequence for fiber bundles to conclude that the desired result.

Step 1: We claim that there exists a fiber bundle,

$$\mathrm{Diffeo}^+(S_{g,b,p}, x) \overset{i}{\hookrightarrow} \mathrm{Diffeo}^+(S_{g,b,p}) \overset{\epsilon}{\twoheadrightarrow} S_{g,b,p}$$

where i is the inclusion and ϵ is the evaluation map at the marked point x. To see this, let U be some open neighborhood of x in $S_{g,b,p}$ that is diffeomorphic to a disc. For every point $u \in U$, we can choose a orientation-preserving self-diffeomorphism $\varphi_u \in \mathrm{Diffeo}^+(U)$ such that

$\varphi_u(x) = u$ and φ_u varies continuously as a function of u. Then, there exists a diffeomorphism from $U \times \text{Diffeo}^+(S_{g,b,p}, x) \to \epsilon^{-1}(U)$ given by

$$(u, \psi) \mapsto \varphi_u \circ \psi.$$

Its inverse map is given by $\psi \mapsto (\psi(x), \phi^{-1}_{\psi(x)} \circ \psi)$. For any y in $S_{g,b,p}$, we can choose a boundary-preserving self-diffeomorphism ξ of $S_{g,b,p}$ that takes x to y. Then, there exists a diffeomorphism $\epsilon^{-1}(U) \to \epsilon^{-1}(\xi(U))$ that sends ψ to $\xi \circ \psi$. Thus, we have shown that there exists a neighborhood U of x such that the restriction of ϵ is a projection to the first factor. In other words, the structure is 'locally a product,' which yields the desired fiber bundle.

Step 2: We can now apply the long exact sequence of homotopy groups of fiber bundles, given in [Hat01, Thm 4.41]:

$$\cdots \to \pi_1(\text{Diffeo}^+(S_{g,b,p})) \to \pi_1(S_{g,b,p}) \to \pi_0(\text{Diffeo}^+(S_{g,b,p})) \to \pi_0(S_{g,b,p}) \to \dots$$

Note that $\pi_0(S_{g,b,p})$ is trivial. Moreover, by Theorem 1.1 from Chapter 1, $\text{Diffeo}^+(S_{g,b,p})$ is simply connected, so its fundamental group $\pi_1(\text{Diffeo}^+(S_{g,b,p}))$ is also trivial. Thus, this long exact sequence reduces to a short exact sequence,

$$0 \to \pi_1(S_{g,b,p}) \xrightarrow{\partial} \pi_0(\text{Diffeo}^+(S_{g,b,p}, x)) \xrightarrow{i*} \pi_0(\text{Diffeo}^+(S_{g,b,p})) \to 0,$$

where i_* and ϵ_* are the maps induced by the inclusion i and evaluation map ϵ, respectively, on the homotopy groups, and where ∂ is the boundary map as defined in [Hat01, Thm 4.3].

But the terms of the sequence can be simplified further:

1. By Remark 2.1, $\text{MCG}(S_{g,b,p}, x) = \pi_0(\text{Diffeo}^+(S_{g,b,p}, x))$ as any two orientation-preserving self-diffeomorphisms of $S_{g,b,p}$ are isotopic if and only if they are homotopic, or in the same path component.

2. Applying Remark 2.1 again, we have that $\pi_0(\text{Diffeo}^+(S_{g,b,p})) = \text{MCG}(S_{g,b,p})$.

3. The space $S_{g,b,p}$ is path-connected, so the fundamental group of $S_{g,b,p}$ is independent of choice of basepoint (in this case, the point x) up to isomorphism. Thus, we have that $\pi_1(S_{g,b,p}) \cong \pi_1(S_{g,b,p}, x)$.

It remains to show that the maps ∂ and i_* are indeed the push and forgetful maps, respectively. We established in Remark 4.5 that the forgetful map is the map induced by the inclusion of $S_{g,b,p} - x \cong (S_{g,b,p}, x) \to S_{g,b,p}$ on their mapping class groups, so we already have that $\mathcal{F}orget = i_*$. We now verify that $\partial = \mathcal{P}ush$ via their respective definitions:

Given an element $[\alpha]_h$ in $\pi_1((S_{g,b,p}, x))$, take any representative $\alpha : (D^1, \partial D^1) \to (S_{g,b,p}, x)$, i.e. a loop α based in x. Because we have shown in Step 1 that $\epsilon : \text{Diffeo}^+(S_{g,b,p}) \twoheadrightarrow S_{g,b,p}$ is a fiber bundle over $\text{Diffeo}^+(S_{g,b,p}, x)$, it follows from [Hat01, Prop 4.48] that the map $\epsilon : \text{Diffeo}^+(S_{g,b,p}) \to S_{g,b,p}$ has the homotopy lifting property with respect to D^k for any $k \geqslant 0$. Applying the homotopy lifting property with respect to D^1, we observe that there exists a map

$$\tilde{\alpha} : (D^1, \partial D^1) \to (\text{Diffeo}^+(S_{g,b,p}), \epsilon^{-1}(x) = \text{Diffeo}^+((S_{g,b,p}, x)))$$

such that $\epsilon \circ \tilde{\alpha} = \alpha$. Then the boundary map ∂ is defined to be the map sending $[\alpha]_h \in \pi_1(S_{g,b,p}, x)$ to $[\tilde{\alpha}\big|_{\partial D^1}]_h \in \pi_0(\text{Diffeo}^+(S_{g,b,p}, x))$, which, as we have established above, is equivalent to $[\tilde{\alpha}\big|_{\partial D^1}]_i \in \text{MCG}(\text{Diffeo}^+(S_{g,b,p}, x))$.

But $\tilde{\alpha}\big|_{\partial D^1}$ is precisely the same as φ_α as given in Definition 4.6: namely, it is the self-homeomorphism of $S_{g,b,p}$ obtained at the end of the 'isotopy of points' given by α. Thus, both the push map and the boundary map send the equivalence class of α to the equivalence class of $\tilde{\alpha}\big|_{\partial D^1} = \varphi_\alpha$. We conclude that $\mathcal{P}ush = \partial$, completing the proof.

□

We now remark on the implications of Theorem 4.5.

Remark 4.8. Theorem 4.5 immediately demonstrates that the push map $\mathcal{P}ush : \pi_1(S_{g,b,p}, x) \to \text{MCG}((S_{g,b,p}, x))$ is well-defined, resolving the issue raised in Remark 4.6. Moreover, the exactness of the sequence confirms our intuition from §4.2.2 that the kernel of the forgetful map is precisely the image of the push map.

Remark 4.9. We can restrict the Birman exact sequence to any subgroup of the mapping class group $\text{MCG}((S_{g,b,p}, x))$, in particular to the pure mapping class group $\text{PMCG}((S_{g,b,p}, x))$. Recall that by Remark 4.5, marking an interior point x in the surface $S_{g,b,p}$ to obtain a marked surface $(S_{g,b,p}, x)$ has the same effect on the mapping class group as *deleting* an interior point x from the surface, i.e., of *puncturing* the surface once (and thus increasing the number of punctures from p to $p+1$). Then, replacing $\text{MCG}(S_{g,b,p})$ with $\text{PMCG}(S_{g,b,p})$, we have that the sequence

$$0 \to \pi_1(S_{g,b,p}) \to \text{PMCG}(S_{g,b,p+1}) \to \text{PMCG}(S_{g,b,p}) \to 0$$

is exact.

As indicated in Remark 4.9, we can apply the Birman exact sequence iteratively to induct on the number of punctures of a given finite-type surface. This will be essential to our attempts to prove the general case, as we shall see in the following section.

4.3 Proving the general case

We now have established the algebraic and geometric machinery to study the generators of the punctured finite-type surface. Our aim in this section is to prove that the pure mapping class group of a finite-type surface, possibly with punctures, is generated by finitely many Dehn twists and isotopies.

We will build towards this proof in two parts. First, we deal with another kind of special case: the genus 0 case, i.e., the n-punctured 2-sphere, which serves as the base case for our double-induction argument. Next, we prove the general case, first by establishing a useful algebra lemma and then by inducting on both the genus g and the number of punctures p of the surface. We also discuss its relationship to the special case proved in Chapter 3 of compact surfaces.

4.3.1 Another special case: the n-punctured sphere

We first consider the genus $g = 0$ case, i.e., surfaces of the form $S_{0,0,n}$, or the n-punctured sphere. This will serve as the base case for our inductive proof of the general case in the following section.

Theorem 4.6. *For $p \geqslant 0$, the pure mapping class group $\mathrm{PMCG}(S_{0,0,p})$ of the sphere with p punctures $S_{0,0,p}$ is generated by finitely many Dehn twists and isotopies.*

To develop this proof, we will use Remark 4.1 to induct on the number of punctures. Our computations of the mapping class groups of the twice- and thrice-punctured 2-sphere in §2.3.2 will serve as the base cases.

Proof of Theorem 4.6. Consider any surface $S_{0,0,p}$ of genus 0 with p punctures. We will induct on the number of punctures p to conclude.

Base case: First we consider the case where $p \leqslant 0$. There are precisely four sub-cases: $p = 0, 1, 2, 3$. We will show that in all of these cases, the pure mapping class group is trivial, and is therefore generated by finitely many Dehn twists and isotopies.

Sub-case (a): For $p = 0$, our surface is the standard 2-sphere $S_{0,0,0}$. By Remark 4.1, we have that $\mathrm{PMCG}(S_{0,0,0}) = \mathrm{MCG}(S_{0,0,0})$ for $p = 0$. But we have already computed $\mathrm{MCG}(S_{0,0,0})$ to be trivial in Example 2.5. Thus, $\mathrm{PMCG}(S_{0,0,0})$ is trivial, generated by the isotopy class of the identity.

Sub-case (b): For $p = 0$, our surface is the punctured 2-sphere $S_{0,0,1}$, homeomorphic to the 2-disc D^2 under the stereographic projection. Again, for $p = 1$, we apply Remark 4.1 to observe that $\mathrm{PMCG}(S_{0,0,1}) = \mathrm{MCG}(S_{0,0,1})$, and apply Example 2.4 from Chapter 2 to observe that $\mathrm{MCG}(S_{0,0,1})$ is also trivial. Thus, $\mathrm{PMCG}(S_{0,0,0})$ is trivial, generated by the isotopy class of the identity.

Sub-case (c): For $p = 2$, our surface is the twice-punctured sphere $S_{0,0,2}$. By Example 2.9, we have that $\mathrm{MCG}(S_{0,0,2}) \cong S_2 \cong \mathbb{Z}/2\mathbb{Z}$, i.e., that the mapping class group is isotopic to the group of symmetries of the two-element set that permute the two punctures. Then any orientation-preserving self-diffeomorphism of $S_{0,0,2}$ that *fixes* the two punctures of the surface must be isotopic to the identity. Since $\mathrm{PMCG}(S_{0,0,2})$ consists of the elements of $\mathrm{MCG}(S_{0,0,2})$ that fix the two punctures, it follows that $\mathrm{PMCG}(S_{0,0,2})$ must be trivial, consisting only of the isotopy class of the identity.

Sub-case (d): For $p = 3$, our surface is the thrice-punctured sphere $S_{0,0,3}$. We follow a similar process to subcase (c) above. By Example 2.8, we have that $\mathrm{MCG}(S_{0,0,2}) \cong S_3$, the symmetric group of degree 3 consisting of all symmetries of the three-element set that permute the three punctures. Again, this means that any orientation-preserving self-diffeomorphism of $S_{0,0,3}$ that fixes the three punctures of the surface must be isotopic to the identity, so then $\mathrm{PMCG}(S_{0,0,3})$ is also trivial, consisting only of the isotopy class of the identity.

This establishes the base case.

Inductive step: We now assume that $p \geqslant 4$. Then the Euler characteristic of $S_{0,0,p}$ is negative: $\chi(S_{0,0,p}) = 2 - 2 \cdot 0 - 0 - p = 2 - p \leqslant 2 - 4 = -2 < 0$. Then we can apply the restriction of the Birman exact sequence from Remark 4.1, from which it follows that the sequence

$$0 \to \pi_1(S_{0,0,3}) \to \mathrm{PMCG}(S_{0,0,4}) \to \mathrm{PMCG}(S_{0,0,3}) \to 0$$

is exact. By sub-case (d) above, $\mathrm{PMCG}(S_{0,0,3})$ is trivial. Moreover, observe that the thrice-punctured 2-sphere $S_{0,0,3}$ is homotopy equivalent to $\bigvee_{i\leqslant 2} S_i^1$, the wedge of two copies of the 1-sphere, as shown in Figure 4.4.

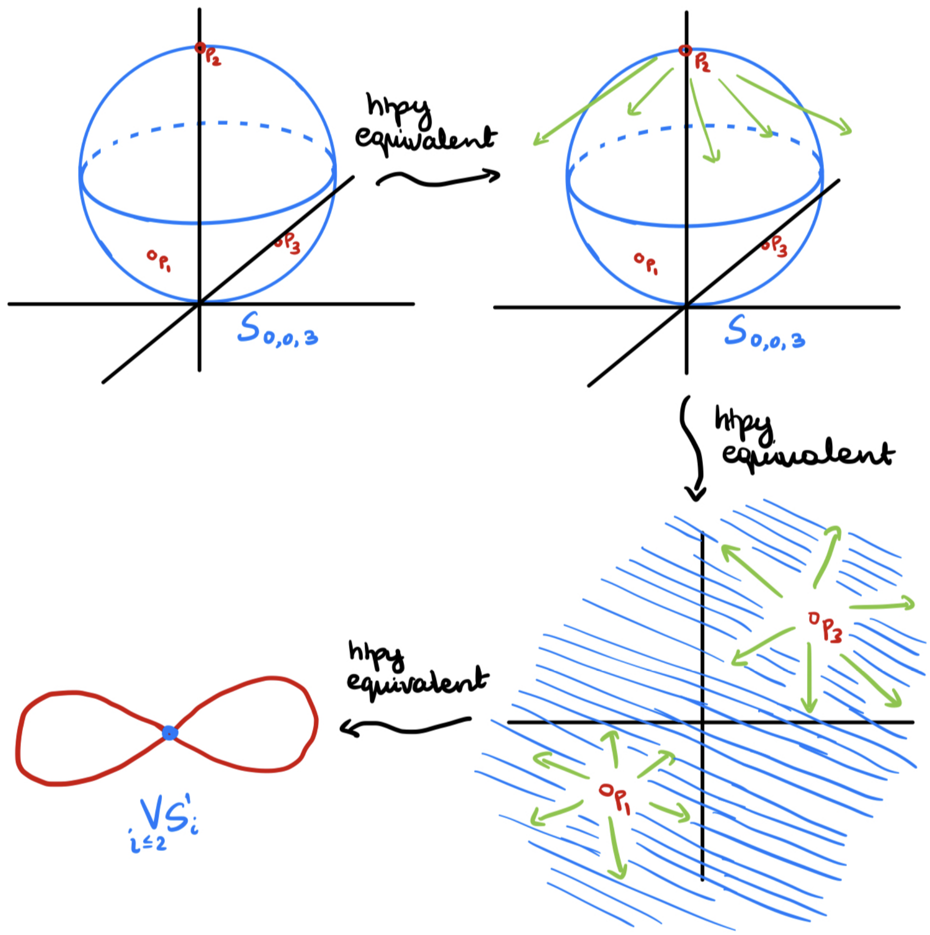


Figure 4.4: $S_{0,0,3}$ is homotopy equivalent to $\bigvee_{i\leqslant 2} S_i^1$, which we see in two steps. Because the once-punctured 2-sphere is homotopy equivalent to $\mathbb{R}^2$ under the stereographic projection (top right), we can homotope $S_{0,0,3}$ to $\mathbb{R}^2$ with two remaining punctures (bottom right). This is equivalent to the wedge of two 1-spheres as desired (bottom left).

Then applying Van Kampen's theorem, given in [Hat01, Thm 1.20, Ex 1.21], the fundamental group $\pi_1(S_{0,0,3}) \cong \pi_1(\bigvee_2 S_i^1) \cong F_2$, the free abelian group with two generators. Thus, the above short exact sequence simplifies to the sequence

$$0 \to F_2 \to \mathrm{PMCG}(S_{0,0,4}) \to 0.$$

By exactness, the map from F_2 to $\mathrm{PMCG}(S_{0,0,4})$ is an isomorphism, implying that $\mathrm{PMCG}(S_{0,0,4})$ is finitely generated.

Having established the algebraic structure of $\mathrm{PMCG}(S_{0,0,4})$, we can now give a geometric interpretation. By Remark 4.7, elements of $\pi_1(S_{0,0,3})$ are represented by simple loops that map to products of the Dehn twists in $\mathrm{PMCG}(S_{0,0,4})$. Thus, the standard generators of $\pi_1(S_{0,0,3})$ give the generating set for $\mathrm{PMCG}(S_{0,0,4})$: two Dehn twists along simple closed curves with a geometric intersection number of 2.

The remainder of the proof follows by repeating this process, iteratively applying the modified version of the Birman exact sequence to induct on p. For instance, at $p = 5$, we have that the sequence

$$0 \to \pi_1(S_{0,0,4}) \to \mathrm{PMCG}(S_{0,0,5}) \to \mathrm{PMCG}(S_{0,0,4}) \to 0$$

is exact. As with the $p = 4$ case, we can once again simplify this. We already have that $\mathrm{PMCG}(S_{0,0,4}) \cong F_2$, and we also observe that the sphere with four punctures $S_{0,0,4}$ is homotopy equivalent to $\bigvee_{i \leqslant 3} S_i^1$, the wedge of three copies of the 1-sphere, via a similar homotopy as was indicated in Figure 4.4. Applying Van Kampen's theorem again, this implies that $\pi_1(S_{0,0,4}) \cong \pi_1(\bigvee_{i \leqslant 3} S_i^1) \cong F_3$, the free abelian group with three generators. Then by exactness of the sequence, $\mathrm{PMCG}(S_{0,0,5})$ must be finitely generated as both F_2 and F_3 are finitely many generators, and the generators of $\mathrm{PMCG}(S_{0,0,5})$ consist of the lifts of generators of F_3 and the images of generators of F_2—which is a finite set of Dehn twists. We again apply Remark 4.7 and the fact that $\pi_1(S_{0,0,n})$ is generated by simple loops to conclude. □

4.3.2 Proof of the general case

We are now ready to prove the general case—that the pure mapping class group of a finite-type surface $S_{g,b,p}$ with genus g and with p punctures is generated by finitely many Dehn twists and isotopies. Our proof uses double-induction, inducting on the values of g and p, and will proceed in three stages. First, we prove an important lemma from geometric group theory that allows us to use results about the connectedness of the modified complex from §4.1.2. Second, we use the forgetful map from §4.2.1 to compute the mapping class group of the once-punctured torus. Together with the results from Theorem 4.6, this will allow us to establish the base case for double-induction. Finally, we prove the general case.

We first return to connectedness of the modified complex from §4.1.2, which serves as a more intricate and combinatorial analogue to the notion of 'relatedness' of curves and arcs from §3.1. Observe that for any surface $S_{g,b,p}$, its mapping class group $\mathrm{MCG}(S_{g,b,p})$ acts on the complex $\hat{\mathcal{N}}(S)$ in the following manner: orientation-preserving self-diffeomorphisms in the mapping class group take non-separating simple closed curves to other non-separating simple closed curves, while preserving the geometric intersection number. In other words, representatives of mapping classes take vertices of $\hat{\mathcal{N}}(S)$ to other vertices.

We want to use the fact that the group $\mathrm{MCG}(S_{g,b,p})$ acts on the modified complex $\hat{\mathcal{N}}(S)$ to say something about the generators of the group—specifically its subgroup, $\mathrm{PMCG}(S_{g,b,p})$. To do this, we need a more general fact from geometric group theory: that if a group G acts cellularly on a connected cell complex X, and D is a subcomplex of X whose G-translates (i.e., orbits) cover G, then G is generated by the set $\{g \in G | gD \cap D \neq \phi\}$. Below, we prove a specialized version of this fact as a lemma.

Lemma 4.7. *Let X be a connected, 1-dimensional simplicial complex and let G be a group that acts on X by simplicial automorphisms. Suppose G acts transitively on the vertices of X and also on the pairs of vertices of X that are connected by an edge. Let v and w be two vertices of X connected by an edge, and choose $h \in G$ such that $h(w) = v$. Then the group G is generated by h together with the stabilizer of v, i.e., of* $\mathrm{Stab}(v) \subset G$.

Proof of Lemma 4.7. Let $g \in G$ and let H be the subgroup of G that is generated by h and the stabilier $\mathrm{Stab}(v)$ of v. We want to show that g lies in H in order to conclude. Observe that g must send vertices to other vertices, and thus, $g(v)$ is also a vertex of the simplicial complex X. Because X is connected, there exists a sequence of vertices, $v = v_0, ..., v_k = g(v)$ that connect the vertices v and $g(v)$, where adjacent vertices in the sequence are connected by an edge (i.e., there exists an edge between each v_i, v_{i+1}).

Because G acts transitively on pairs of vertices of X containing an edge, we can choose elements $g_i \in G$ such that $g_i(v) = v_i$. We take g_0 to be the identity and g_k to be g, so that $g_0(v) = v_0 = v$ and $g_k(v) = v_k = g(v)$. To conclude, it suffices to show that g_i lies in H for all i. We will prove this by induction.

Base case: Since H is a subgroup of G, it must contain the identity, and thus $g_0 = \mathrm{id} \in H$. This establishes the base case.

Inductive step: Now suppose that $g_i \in H$ for some $i \leqslant k$. We want to show that $g_{i+1} \in H$ to conclude.

We first apply the inverse g_i^{-1} to the edge of X that is between the vertices v and $g_i^{-1}g_{i+1}(v)$. Since G acts transitively on ordered pairs of vertices containing an edge, there exists an element r in G taking the pair $(v, g_i^{-1}g_{i+1}(v))$ to the pair (v, w).

Observe then that r lies in the stabilizer $\mathrm{Stab}(v)$, as it fixes v, and that $rg_i^{-1}g_{i+1}(v) = w$. Since $h(w) = v$, we have that $h^{-1}rg_i^{-1}g_{i+1}(v) = v$, from which it follows that $hrg_i^{-1}g_{i+1}$ lies in Stab(v). Then in particular, r lies in Stab(v) and therefore also in H. Then h and r are both contained in H and g_i lies in H by the inductive hypothesis. We can compose with inverses to obtain that g_{i+1} also lies in H. Then by induction, $g_i \in H$ for all $i \leqslant k$. In particular, $g_k = g \in H$, so we conclude. □

Our final preparation is to compute the mapping class group of the once-punctured 2-torus. We did not compute this as a motivating example in §2.3 because the complication is greatly simplified by introducing the forgetful map. As we shall see, the forgetful map reduces this computation to the case of the standard 2-torus that we already computed in Example 2.7.

Example 4.10 (Mapping class group of the once-punctured 2-torus). Consider $S_{1,0,1}$, the once-punctured 2-torus. We claim that its mapping class group is isomorphic to that of the standard torus $T^2 = S_{1,0,0}$, i.e., that $\mathrm{MCG}(S_{1,0,1}) \cong \mathrm{MCG}(S_{1,0,0}) \cong \mathrm{SL}_2(\mathbb{Z})$ from Example 2.7.

Recall from Remark 4.5 that puncturing a surface has the same effect on its mapping class group as marking a point x in its interior, so $\mathrm{MCG}((S_{1,0,0}, x)) \cong \mathrm{MCG}(S_{1,0,1})$. Then by Definition 4.5, we have a natural homomorphism $\mathrm{MCG}((S_{1,0,0}, x)) \cong \mathrm{MCG}(S_{1,0,1}) \to \mathrm{MCG}(S_{1,0,0})$, namely the forgetful map $\mathcal{F}orget$ given by 'forgetting' the marked point x, or equivalently, by 'filling in' the puncture.

We claim that in this case, the forgetful map is actually an isomorphism. We already have by Remark 4.4 that the map is surjective, so it remains to show that the map is also injective. To see this, we identify the punctured torus $S_{1,0,1}$ with the standard torus with a marked interior point, i.e., with $(S_{1,0,0}, x)$ where x is some point in the interior of $S_{1,0,0}$ (recalling that the mapping class groups of the once-punctured torus and of the standard torus with a marked interior point are isomorphic). Let φ_t be an isotopy between two orientation-preserving self-diffeomorphisms φ and φ' that both preserve the marked point x, i.e., two representatives of the same mapping class in $\mathrm{MCG}((S_{1,0,0}, x)) \cong \mathrm{MCG}(S_{1,0,1})$. Let $\alpha(t) = \varphi_t(x)$, which is a loop based in x. Then we can obtain an isotopy F_t from φ to φ' that fixes x at every time t by setting F_t to be the inverse of the loop $\alpha(t)$ composed with φ_t. In other words, we set $F_t(y) = (\alpha(t))^{-1}\varphi_t(x)$. This concludes the proof that the forgetful map is injective.

It follows that the mapping class group of the once-punctured torus $S_{1,0,1}$ is isomorphic to the mapping class group of the standard torus $S_{1,0,0}$. Then by Example 2.7, $\mathrm{MCG}(S_{1,0,1})$ is also generated by finitely many Dehn twists and isotopies.

We now have all the necessary results to induct on the genus g and number of punctures p of a surface $S_{g,b,p}$.

Theorem 4.8. *Given any finite-type surface $S_{g,b,p}$, the pure mapping class group $\mathrm{PMCG}(S_{g,b,p})$ is generated by finitely many Dehn twists around non-separating simple closed curves in $S_{g,b,p}$.*

Proof of Theorem 4.8. Take any surface $S_{g,b,p}$. By Remark 4.2, we can assume without loss of generality that $b = 0$ and induct on g and p.

Base Case: We have already proved the $g = 0$ case in Theorem 4.6 from Section 4.3.1. Thus, suppose $g \geqslant 1$. Our base cases are then at $p = 0$ and $p = 1$, i.e., the standard 2-torus $S_{1,0,0}$ and the once-punctured 2-torus $S_{1,0,1}$. We have already shown that their mapping class groups are generated by finitely many Dehn twists along non-separating simple closed curves in Example 2.7 and Example 4.10, respectively. As the pure mapping class group is a subgroup of the mapping class group, both $\mathrm{PMCG}(S_{1,0,0})$ and $\mathrm{PMCG}(S_{1,0,1})$ are also generated by finitely many Dehn twists along non-separating simple closed curves. This establishes the base case.

Inductive step on p: Now we induct on p. Let $g \geqslant 1$ and $p \geqslant 0$. For the inductive hypothesis, suppose $\mathrm{PMCG}(S_{g,0,p})$ is generated by finitely many Dehn twists about non-separating simple closed curves $\{\alpha_i\}$ in $S_{g,0,p}$. We want to show that $\mathrm{PMCG}(S_{g,0,p+1})$ is generated by finitely many Dehn twists about non-separating simple closed curves in $S_{g,0,p+1}$. As we have already addressed the case where $g = 1$ and $p = 0$ in the base case for the standard 2-torus, we assume $(g, p) \neq (1, 0)$, i.e., if $g = 1$, then $p \geqslant 1$. Observe that then the Euler characteristic $\chi(S_{g,b,p})$ is always negative: if $g = 1$, $\chi(S_{g,b,p}) = 2 - 2g - 0 - p \leqslant 2 - 2 - 1 = -1 < 0$, and if $g > 1$, then $\chi(S_{g,b,p}) < 2 - 2 - 0 - p = 0$.

Then we can apply the modified version of the Birman exact sequence from Remark 4.1 to observe that the sequence

$$0 \to \pi_1(S_{g,0,p}) \to \mathrm{PMCG}(S_{g,0,p+1}) \to \mathrm{PMCG}(S_{g,0,p}) \to 0$$

is exact. As $g \geqslant 1$, the fundamental group $\pi_1(S_{g,0,p})$ is generated by classes of finitely many simple loops. By Remark 4.7, the image of each loop is the product of two Dehn twists about non-separating simple closed curves.

Our next step is building a generating set for the pure mapping class group $\mathrm{PMCG}(S_{g,0,p+1})$. Taking these Dehn twists individually, we need to choose a lift of each Dehn twist generator T_{α_i} of $\mathrm{PMCG}(S_{g,0,p})$ to $\mathrm{PMCG}(S_{g,0,p+1})$. We will do this by using the forgetful map from §4.2.1. Recall by Remark 4.4 that the forgetful map is surjective. Then, given any non-separating simple closed curve $\alpha_i \in S_{g,0,p}$, there exists a non-separating simple closed curve in $S_{g,0,p+1}$ that maps to α_i under the forgetful map.

This implies that the corresponding Dehn twist generator T_{α_i} has a preimage in $\mathrm{PMCG}(S_{g,0,p+1})$ that is a Dehn twist about a non-separating simple closed curve in $S_{g,0,p+1}$. This concludes our induction on p. By induction, the pure mapping class group $\mathrm{PMCG}(S_{g,0,p})$ is generated by finitely many Dehn twists around non-separating simple closed curves for all $p \geqslant 0$ for a fixed $g \geqslant 1$.

Inductive step on g: Our next step is to induct on g. Take any surface $S_{g,0,p}$. As we have already inducted on p, we can assume without loss of generality that $p = 0$. Then our $g = 0$ and $g = 1$ reduce to the base case, so we can also without loss of generality assume $g \geqslant 2$. Thus we are dealing with surfaces of the form $S_{g,0,0}$, where by Remark 4.1, $\mathrm{PMCG}(S_{g,0,0}) = \mathrm{MCG}(S_{g,0,0})$.

For the inductive hypothesis, suppose $\mathrm{PMCG}(S_{g-1,0,p})$ is finitely generated by isotopies and Dehn twists about non-separating simple closed curves for all $p \geqslant 0$. We want to show that $\mathrm{PMCG}(S_{g,0,0}) = \mathrm{MCG}(S_{g,0,0})$ is also finitely generated by isotopies and Dehn twists about non-separating simple closed curves.

We now apply Theorem 4.2 and Lemma 4.6 to the modified complex of the surface $S_{g,0,0}$. Since $g \geqslant 2$, Theorem 4.2 implies that the modified complex $\hat{\mathcal{N}}(S_{g,0,0})$ is connected. Moreover, observe that $\mathrm{MCG}(S_{g,0,0})$ acts transitively on ordered pairs of isotopy classes of simple closed curves with a geometric intersection number of 1—i.e., pairs of vertices of the modified complex $\hat{\mathcal{N}}(S_{g,0,0})$ that are joined by edges. Then we can apply Lemma 4.7 to the group action of $\mathrm{MCG}(S_{g,0,0})$ on the simplicial complex $\hat{\mathcal{N}}(S_{g,0,0})$. Let a be any isotopy class of non-separating simple closed curves in $S_{g,0,0}$, and let b be an isotopy class of non-separating simple closed curves in $S_{g,0,0}$ such that the geometric intersection number $i(a, b) = 1$ (a and b are joined by an edge in the modified complex $\hat{\mathcal{N}}(S_{g,0,0})$). Let $\mathrm{MCG}(S_{g,0,0}, a)$ denote the stabilizer in $\mathrm{MCG}(S_{g,0,0})$ of a. By Theorem 2.2, $T_bT_a(b) = a$. Then Lemma 4.7 implies that $\mathrm{MCG}(S_{g,0,0})$ is generated by $\mathrm{MCG}(S_g, a)$, T_a, and T_b.

Thus to conclude, it suffices to show that $\mathrm{MCG}(S_g, a)$ is generated by finitely many Dehn twists about non-separating simple closed curves. Let $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ be the subgroup of $\mathrm{MCG}(S_g, a)$ of elements preserving the orientation of a. Observe that the sequence

$$0 \to \mathrm{MCG}(S_{g,0,0}, \vec{a}) \to \mathrm{MCG}(S_{g,0,0}, a) \to \mathbb{Z}/2\mathbb{Z} \to 0$$

is exact, where the map from $\mathrm{MCG}(S_{g,0,0}, \vec{a}) \to \mathrm{MCG}(S_{g,0,0}, a)$ is the inclusion and the map from $\mathrm{MCG}(S_{g,0,0}, a) \to \mathbb{Z}/2\mathbb{Z}$ is the map that sends elements of the mapping class group $\mathrm{MCG}(S_{g,0,0}, a)$ to $0 \in \mathbb{Z}/2\mathbb{Z}$ if they preserve the orientation of a and to $1 \in \mathbb{Z}/2\mathbb{Z}$ if they reverse the orientation of a. Using the change of coordinates principle from §1.1.4, we see that as the combination of Dehn twists $T_b T_a^2 T_b$ switches the orientation of a, its equivalence class must represent the nontrivial coset of $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ in $\mathrm{MCG}(S_{g,0,0}, a)$ (where $\mathrm{MCG}(S_{g,0,0}, a)$ is the same as modding out by orientation).

We now apply the cutting homomorphism from Theorem 2.3. Taking some representative α of a, we have a homomorphism from $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ to $\mathrm{MCG}(S_{g,0,0} - \alpha)$ with kernel $< T_\alpha >$, where $S_{g,0,0} - \alpha$ is the surface obtained from $S_{g,0,0}$ by deleting α from the surface. Since the surface $S_{g,0,0}$ (and therefore $S_{g,0,0} - \alpha$) has no punctures, Remark 4.1 implies that $\mathrm{MCG}(S_{g,0,0} - \alpha) = \mathrm{PMCG}(S_{g,0,0} - \alpha)$. Then by the isomorphism theorems for group homomorphisms, the sequence

$$0 \to < T_a > \to \mathrm{MCG}(S_{g,0,0}, \vec{a}) \to \mathrm{PMCG}(S_{g,0,0} - \alpha) \to 0$$

is exact.

But $S_{g,0,0} - \alpha$ is equivalent to $S_{g-1,0,2}$. In other words, cutting the surface along a non-separating simple closed curve α is equivalent to decreasing its genus by one and adding two punctures, as is illustrated in Figure 4.5.

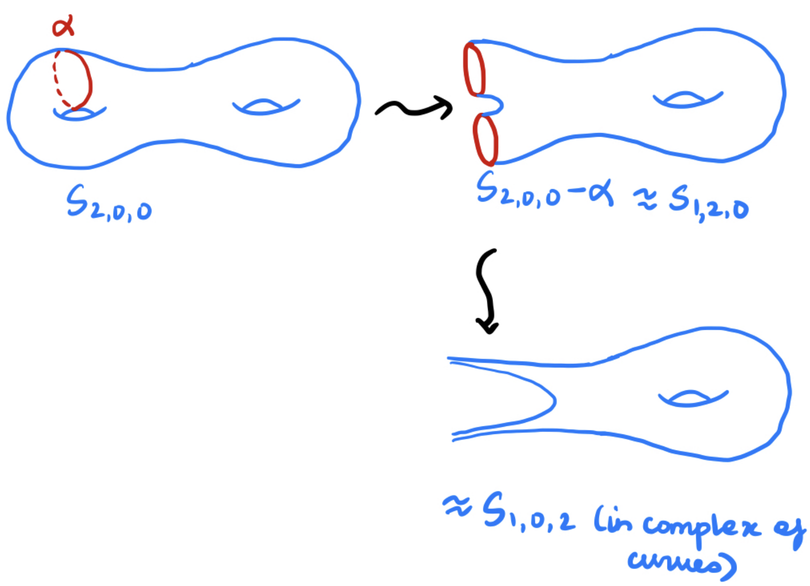


Figure 4.5: Cutting the genus 2 surface $S_{2,0,0}$ by a non-separating simple closed curve α yields a surface that is equivalent to $S_{1,0,2}$ in the complex of curves.

Then $\mathrm{PMCG}(S_{g,0,0} - \alpha) \cong \mathrm{PMCG}(S_{g-1,0,2})$. But $\mathrm{PMCG}(S_{g-1,0,2})$ is generated by finitely many Dehn twists along non-separating simple closed curves by the inductive hypothesis. Each of those Dehn twists has a preimage in $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ that is also a Dehn twist about a non-separating simple closed curve. It follows that $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ is generated by finitely many Dehn twists about non-separating simple closed curves. By exactness, as both $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ and $\mathbb{Z}/2\mathbb{Z}$ are finitely generated, it follows that $\mathrm{MCG}(S_{g,0,0})$ is finitely generated as well, and that its generators are the finite set of images of generators of $\mathrm{MCG}(S_{g,0,0}, \vec{a})$ and lifts of generators of $\mathbb{Z}/2\mathbb{Z}$, which are Dehn twists about non-separating simple closed curves in $S_{g,0,0}$. Thus we conclude. □

We close with some observations regarding the relationship between Theorem 4.8 and the special case where $p = 0$ dealt with in Chapter 3.

Corollary 4.9. *Theorem 3.1.*

Proof of Corollary 4.9. Taking $p = 0$, it follows from Theorem 4.8 that the pure mapping class group of the surface $S_{g,b,0}$ is generated by finitely many Dehn twists about non-separating simple closed curves. Since $p = 0$, it follows from Remark 4.1 that $\mathrm{PMCG}(S_{g,b,0}) = \mathrm{MCG}(S_{g,b,0})$. Thus we conclude that $\mathrm{MCG}(S_{g,b,0})$ is generated by finitely many Dehn twists about non-separating simple closed curves. Note that this is in fact a *stronger* version of the statement of Theorem 3.1 proved in Chapter 3, since it demonstrates not only that these Dehn twists generate but that they *finitely* generate, which was not part of the original statement of Theorem 3.1. □

Corollary 4.10. *Theorem* **??** *(Dehn-Lickorish).*

Proof of Corollary 4.10. Taking $b = 0$ and $p = 0$, it follows from Theorem 4.8 that the pure mapping class group of the surface $S_{g,0,0}$ is generated by finitely many Dehn twists about non-separating simple closed curves. Once again, since $p = 0$, we apply Remark 4.1 to observe that $\mathrm{PMCG}(S_{g,0,0}) = \mathrm{MCG}(S_{g,0,0})$, and thus we conclude. □

As a final note on this subject, we discuss ongoing efforts to specify the generators of the mapping class group of punctured surfaces. Chapter 3 discusses how both Dehn and Lickorish not only individually demonstrated that Dehn twists finitely generate the mapping class group of the surface $S_{g,0,0}$ in Theorem **??**, but that they also computed precisely how many Dehn twists were needed to generate in terms of the genus g. This raises the question of whether there is more we can say to specify the number of generators of the mapping class group or the pure mapping class group of a finite-type surface $S_{g,b,p}$—or the generators themselves.

Example 2.7 gave us a flavor for this work, as we were able to specify the generating set of the mapping class group of the 2-torus to be Dehn twists around the meridinal and longitudinal curves m and l. In 1934, Wilhelm Magnus (1907-90), who was Dehn's doctoral student at the University of Frankfurt, identified a presentation for the mapping class group of the twice-punctured 2-torus in [Mag34, §4, p.634-39]. Birman, who was Magnus's doctoral student at the New York University Courant Institute, took up the problem of finding a representation of the mapping group of the thrice-punctured 2-torus, working in higher generality to link mapping class groups and braid groups (the group consisting of isotopy classes of n-braids) in [Bir74].

Subsequently, Lickorish defined a collection of non-separating simple closed curves in a compact surface that would determine the Dehn twists that generated the mapping class group in [Lic08]. These curves are known as the Lickorish generators. During the same period, Stephen Humphries specified a different set of curves known as the Humphries generators in [Hum06] that generate the mapping class group of a compact surface. This question is still being investigated in the general case of the mapping class group of a surface with boundaries and/or punctures. While a fuller discussion of this subject casts beyond the scope of this thesis, the reader can refer to [FM11, §4.4, p.108-15] for a survey of the Lickorish and Humphries generators.

Epilogue

"I have a feeling that every mathematician thinks in terms of pictures. And the pictures may not be pictures like these, but they're pictures that work in a different way."
— Joan Birman in [Bir24]

Birman's comments—referencing Figure 4.3 that interprets the push map geometrically as the product of two Dehn twists—point to a larger question at the heart of this thesis. *What is the relationship between algebra and geometry, between abstraction and visualization?* These map onto what Steingart refers to as "two modes of mathematical engagement": one that is "abstract" and "theoretical," the other that is "concrete" and "tangible" [Ste15, p.46]. Our investigation into the generators of the mapping class group reflects an intricate relationship between these two modes. To cite just a few examples we have encountered in preceding chapters:

1. The complex of curves in §4.1.2 uses the geometric intersection number—defined in §1.1.4 using hyperbolic geometry—to think combinatorially about the curves on a surface, as in Theorem 4.2;

2. The two group homomorphisms introduced in §4.2—the forgetful map and the push map—both have rich geometric interpretations: the forgetful map is like 'filling in a puncture,' and the push map is like 'pushing one's finger in a loop based at the marked point' to distort the surface;

3. The generation of the n-punctured sphere in §4.3.1, where the exactness of the sequence in Remark 4.9 gave us that the pure mapping class group was finitely generated, while the geometric interpretation of the push map in Remark 4.7 reveals that its generators were Dehn twists.

In turn, the epilogue to this thesis proposes to situate the mathematical investigation pursued in preceding chapters—providing geometric and algebraic preliminaries, defining and computing examples of the mapping class group, proving generation theorems for the mapping class groups of compact and non-compact surfaces—within the history of these two modes of mathematical thought. We draw primarily on an interview the author conducted with Professor Birman, where she generously shared her reflections on her mathematical formation, doctoral work on the Birman exact sequence, and the broader role of visual intuition in the process of mathematical discovery. When contextualized within the history of twentieth-century pure mathematics, Birman's reflections suggest that the generation of the mapping class group has an interesting role to play in the merging of these two modes of engagement during the 1960s and 1970s.

The relationship between geometry and algebra has a much longer history, spanning regions and periods as disparate as ancient Greece, early modern Europe, and ancient and medieval India. As Reviel Netz argues in [Net99, Chapter 1], Ancient Greek mathematical treatises such as Euclid's *Elements of Geometry* (300 BCE) developed a close relationship between diagrams and text to develop early proof methods through deductive reasoning. In early modern Europe, Mary Domski writes in [Domcs] that texts such as René Descartes's *La Géometrie* (1637) introducing the coordinate system on the Euclidian plane developed a renewed emphasis on "the blending of algebra and geometry." Nor is this story uniquely restricted to Greco-Roman antiquity and European modernity. T.A. Sarasvati Amma argues that "the basis and inspiration for the whole of Indian mathematics is geometry," although there was a "shift in emphasis" between the ancient and medieval periods. Where earlier Sanskrit mathematical texts such as the *Śulbasūtra* (1000 BCE) were "primarily interested in geometric constructions, the applied algebraic truths coming in by the side door," for the "later mathematicians the algebraic results are the most important, the geometric figures being merely an aid to make the algebraical results more immediately convincing" [Amm79, Introduction, §9.2]. While a fuller account casts well beyond the scope of this thesis, we highlight a few interesting episodes—if only to suggest that there is a precedent for the manner in which twentieth-century mathematicians negotiated between 'concrete' and 'abstract' modes.

Mathematical developments since the late nineteenth century marked another reconfiguration of the relationship between algebra and geometry. Historians of mathematics such as Jeremy Gray refer to the late nineteenth and early twentieth century as an era of "mathematical modernism," where mathematicians eschewed pictures and diagrams for abstraction and formalism [Gra08]. This was the result of a number of shifts in late-nineteenth- and early-twentieth-century mathematics: Emmy Noether's work on axiomatizing the field of algebra (discussed in [Ste23, p.37], the growing acceptance of hyperbolic geometry (discussed in [Gra08, §2.1.2] and [Mil82]), which broke from Euclidian geometry and thus also with the direct correspondence to sensory experience, and the emergence of the field of topology.

As Henri Poincaré wrote in *Analysis Situs* (1895)—a text that [Ste23, p.31] notes marked the advent of topology—"geometry, in effect, has a unique *raison d'être* as the description of the objects that underlie our senses," while algebraic abstraction has the benefit of being "more concise" and suggesting "useful generalizations" [Poi95, p.1-2]. Topology proposed to reconfigure the relationship between geometry and algebra altogether—and the role of figures and diagrams in mathematics.

This move towards abstraction intensified in the mid-twentieth century as pure mathematics emerged as a separate field from applied mathematics. Tracing the development of algebriac topology in American universities in the 1940s-50s, Steingart observes how mathematics entered into a period of "high modernism" [Ste23, p.28]. They drew influence from the Bourbaki school that had emerged in 1930s France, which eschewed pictures and other forms of concrete visualizations to work towards axiomatization (discussed further [Cor01, 170-85]).

Steingart's analysis of Samuel Eilenberg and Norman Steenrod's *Foundations of Algebraic Topology* (1952) offers a particularly revealing example of how 'high modernism' related geometry to algebra. As Steingart notes in [Ste23, p.44], Eilenberg and Steenrod developed a four-part schema for the development of algebraic topology, which moved from the "geometrical questions" about spaces to "algebraic theory" regarding homology groups in [ES52, p.viii]. Their efforts to axiomatize the field of algebraic topology "began at the end," as Steingart puts it: the geometric roots of the field were effaced, and all that would be presented to students and to researchers was the abstract perspective based on an axiomatized algebra. Within the framework of this kind of "modern" mathematics, geometry might be the beginning, but algebra—and pure abstraction—were the telos, a sign of progress in mathematical research.

Yet the relationship between geometry and algebra proved to be more complex, as the 1960s-70s would reveal. Steingart observes a growing disillusionment amongst mathematicians towards the total dominance of the "abstract mode" in [Ste23, p.29, 50-54], which she surveys elsewhere in [Ste15] by tracking how mathematicians such as Charles Pugh and William Thurston approached the problem of sphere eversion, seeking pictorial, tactile, and even animated video "manifestations" to "understand" Smale's abstract proof. Birman's research during this very period—the 1960s onwards—would suggest that the problem of generating the mapping class group was another important episode in this history, reflecting a sustained attempt to draw together these "two modes of engagement."

"The things I was doing where I was drawing pictures were very much out of fashion, and not very highly regarded," Birman explained, when asked about her doctoral work at the New York University Courant Institute of Mathematical Sciences. There was a "snobbery.. from the Bourbaki school" towards drawing pictures, although she recalled being aware of two mathematicians who worked more visually. One was the Danish mathematician Jakob Nielsen (1890-1959), whose work on automorphisms of surfaces is surveyed in [CB88]. The other was the Austrian mathematician Emil Artin (1898-1962), whose paper "Theorie der Zöpfe" [Theory of Braids] (1925) Birman had to translate to English for her graduate school language exam. It was Artin's paper that initially sparked her interest in braid groups and topology, Birman recalled, although both Nielsen and Artin "were really doing things that were not so fashionable at the time."

Birman's own mathematical formation was highly unusual, by her account. As she has spoken and written elsewhere about her childhood and later career, our conversation focused primarily on the period between 1948, Birman's graduation from Barnard College, and 1974, her return to the Barnard and Columbia Mathematics Departments as a faculty member—precisely the period when she began her pioneering research on mapping class groups and proved the Birman exact sequence. She completed her undergraduate degree in mathematics at Barnard College in 1948, and it was during her final year at Barnard that she began to envision pursuing graduate work in mathematics. She was in a graduate-level probability theory class at the time, and spent "three of four days in the library" right before the exam, immersing herself in the material. "I enjoyed it," Birman recalled, and thought she would "have to choose either between making a mess of graduate school or really working like [she] did during those few days in the library." She chose to take up an industry job and marry, a decision she has spoken of elsewhere (see [JT06]), and returned to pursue graduate study in mathematics in 1961.

In the context of this epilogue, Birman's decision to work in the industry—and then to return to graduate school at the Courant Institute, which primarily focused on applied mathematics—meant that she had far more exposure to the 'concrete, geometric' mode of engagement with mathematics, than she might have had she remained strictly within the realm of 1950s 'high modernist mathematics.' Her first job after college was at the Polytechnic Research Group, where she was tasked with solving an engineering problem of how to produce microwave frequency meters. During our conversation, Birman analogized this problem to the 'sliding ladder' problem in mathematics and physics (the setup is roughly: the top of a ladder is sliding down a wall at a particular speed; how fast is the bottom of the ladder sliding on the floor from the base of the wall?). Once she solved this problem, they ran out of mathematical problems for her and instead tasked her with taking various measurements. In search of a "a more interesting job," Birman re-enrolled for a master's degree in physics at Columbia, where she found she was most comfortable with running experiments in the lab setting.

Interestingly, Birman recalled of this period that she did not have a deep "intuition about physics"—suggesting an interesting distinction between how she experienced visual intuition in physics versus in her subsequent work in pure mathematics. The physics program convinced her that should she return to graduate study, it would instead be in mathematics, and in 1960 she began taking night classes at the Courant Institute; she became a full-time doctoral student soon afterwards when the program awarded her a scholarship after her performance in qualifying exams. It was at Courant that she took her first course in topology after translating the Artin paper on braid theory. The course was not taught by a topologist—Courant itself was known for its emphasis on applied mathematics, and Birman's first topology instructor was Jacob Schwartz (1930-2009), who worked primarily on linear operators and "was known as somebody on the faculty who liked to teach things he didn't know." There were a few faculty who did work in pure mathematics, among whom was Magnus, who came to the United States in 1948 after training under Dehn in Germany.

Magnus was an algebraicist, whose mathematical engagement tended more to the side of the 'abstract' mode of engagement. As discussed at the end of Chapter 4, he had worked on a presentation of the mapping class group of the twice-punctured 2-torus in 1934, published in [Mag34]. Observing Birman's interest in topology and the braid group, he encouraged her to focus her doctoral work on generalizing these results to consider the mapping class groups of surfaces with more punctures and higher genera. It was in this context that she developed the Birman exact sequence and her later research linking the braid group to the mapping class group that was eventually published in [Bir74].

"I was thinking about it visually," Birman recalled, when asked about how she understood the push map and pieced together the fact that it was the kernel of the forgetful map. Elsewhere, she has written of how even Magnus had not anticipated the direction in which her research would develop when he first encouraged her to study the mapping class group [Bir16, p.2]. Where he had approached the problem of finding a presentation for the mapping class group algebraically in his early work in the 1930s, Birman's work in the 1960s with the Birman exact sequence and later on braid groups made connections to geometry and topology that had yet to be fully pursued.

Birman's engagement with applied mathematics and engineering continued even beyond her industry work and doctoral training: her first academic appointment was as a faculty member at the Stevens Institute of Technology, where she taught engineering students, while also collaborating with researchers from Princeton. While drawing pictures might have been "out of fashion" in pure mathematics—with the Princeton Mathematics Department featuring prominently in Steingart's account of 1950s "high modernist mathematics" in [Ste23, Ch. 1]—Birman recalled that this was a central part of teaching and research in engineering. During this period, she was invited by Ralph Fox to give a seminar on her research on braid groups and mapping class groups at Princeton. ("There was somebody who was named John" in the audience, Birman remembered. Afterwards, she asked Fox, "Who is John? Who is asking all these questions?". The individual in question was in fact John Milnor.) She was subsequently invited to spend a semester as a visiting assistant professor at Princeton, and then hired by Eilenberg to join the mathematics faculty at Columbia. There, she taught courses ranging from the undergraduate honors mathematics sequence to graduate seminars in topology.

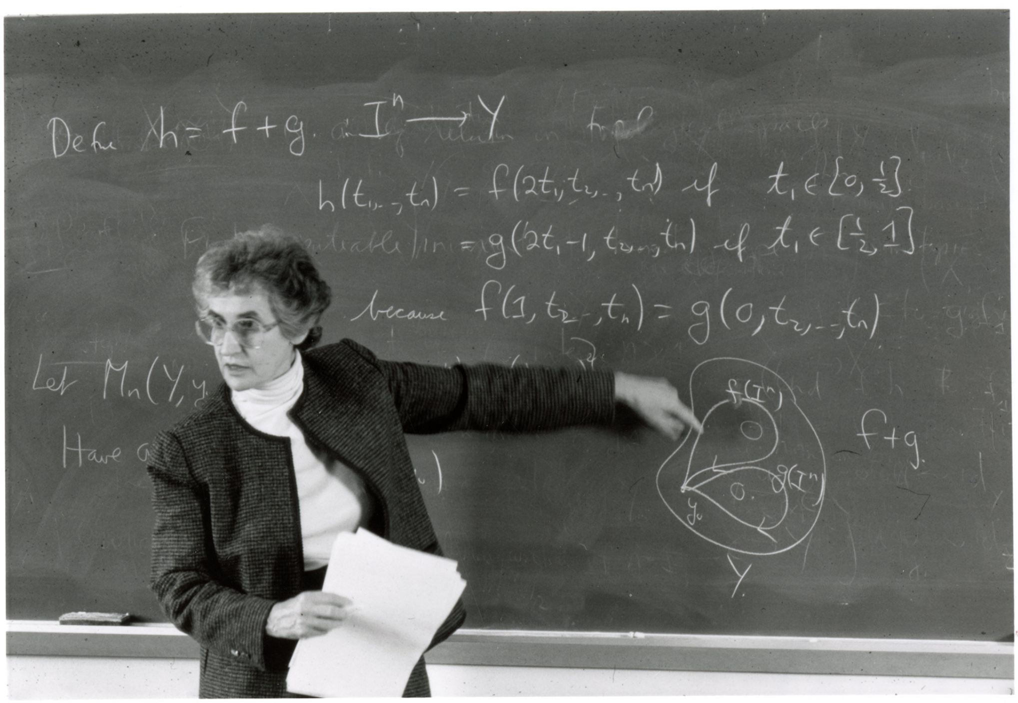


Figure 4.6: Birman teaching at Columbia in 1985.

When asked about her trajectory, Birman initially offered the following reflection: "I had a very non-traditional education... So when I got a really good job, I never felt like I knew enough. Maybe that's true of everybody, but I'm not sure of that." Yet as our conversation developed further, she noted that her unconventional exposure to applied mathematics and engineering alongside pure mathematics during the 1950s might have shaped her approach to research as she led the field in new directions. On a higher level, she maintained that all mathematicians thought visually in one way or another, giving the example of a number theorist whom she studied with in graduate school who used to use colored chalk in his own work and in lecture. "I was wondering whether he was in some way visualizing it," she said—suggesting that even the most abstract, algebraic work in mathematics might also draw on the visual mode of engagement.

Birman's contributions are foundational for low-dimensional topology, specifically for the study of mapping class groups and braid groups, as discussed in [GML98, p.ix]. Dan Margalit, the co-author of [FM11], describes how her work has "has often been ahead of its time" [Mar19, p.341]. His observation might be true in more than one sense. Birman's exact sequence and other early results not only spurred a new interest in braid groups and the mapping class group at a time when they were "out of fashion," but they also appear to have been part of this broader shift in how mathematicians drew on concrete versus abstract modes of engagement in their research. If the 1960s were a kind of 'meridian' in pure mathematics—as a generation of mathematicians shifted away from high modernism's exclusive emphasis on abstraction towards an approach that would draw together algebra and geometry—then Birman's early research suggests she was *in* this meridian, and that the problem of generating the mapping class group had a signficant role to play in this period of transition.

Bibliography

[Amm79] T.A. Sarasvati Amma. *Geometry in Ancient and Medieval India.* Motilal Banarsidass, 1979.

[Bir74] Joan Birman. *Braids, Links, and Mapping Class Groups.* Princeton University Press, 1974.

[Bir16] Joan Birman. My favorite paper. 2016.

[Bir24] Joan Birman. Interview by author, February 28, 2024.

[CB88] A.J. Casson and Steven A. Bleiler. *Automorphisms of Surfaces After Nielson and Thurston.* University of Cambridge Press, 1988.

[Cor01] Leo Corry. Mathematical structures from hilbert to bourbaki: The evolution of an image of mathematics. *Changing Images in Mathematics: From the French Revolution to the New Millennium*, edited by Umberto Bottazzini and Amy Dahan-Dalmedico, 167-85. Routledge, 2001.

[Deh87] Max Dehn. *Papers on Group Theory and Topology.* Translated by John Stillwell. Springer-Verlag, 1987.

[Domcs] Mary Domski. Descartes' mathematics. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta and Uri Nodelman (eds.), 2022. https://plato.stanford.edu/archives/fall2022/entries/descartes-mathematics/.

[ES52] Samuel Eilenberg and Norman Steenrod. *Foundations of Algebraic Topology.* Princeton University Press, 1952.

[FHV21] Federica Fanoni, Sebastian Handel, and Nicholas G. Vlamis. Big mapping class groups acting on homology. *Indiana University Math Journal*, 2021.

[FM11] Benson Farb and Dan Margalit. *A Primer on Mapping Class Groups.* Princeton University Press, 2011.

[GML98] Jane Gilman, William W. Menasco, and Xiao-Song Lin. *Knots, Braids, and Mapping Class Groups—Papers Dedicated to Joan S. Birman.* American Mathematical Society, 1998.

[GP78] Victor Guillemin and Alan Pollack. *Differential Topology.* AMS Chelsea Publishing, 1978.

[Gra73] André Gramain. Le type d'homotopie du groupe des difféomorphismes d'une surface compacte. *Annales scientifiques de l'École Normale Supérieure*, Vol. 6, No. 4, 1973.

[Gra08] Jeremy Gray. *Plato's Ghost: The Modernist Transformation of Mathematics.* Princeton University Press, 2008.

[Hat01] Allen Hatcher. *Algebraic topology.* Cambridge University Press, 2001.

[Hir76] Morris W. Hirsch. *Differential Topology.* Springer, 1976.

[Hum06] Stephen P. Humphries. Generators for the mapping class group. *Topology of Low Dimensional Manifolds*, 2006.

[JT06] Allyn Jackson and Lisa Traynor. Interview with joan birman. *Notices of the American Mathematical Society*, 2006.

[Lic08] William Lickorish. A finite set of generators for the homeotopy group of a 2-manifold. *Mathematical Proceedings of the Cambridge Philosophical Society*, 2008.

[Mag34] Wilhelm Magnus. Über Automorphismen von Fundamentalgruppen berandeter Flächen. *Mathematische Annalen*, Vol. 109, 1934.

[Mar19] Dan Margalit. The Mathematics of Joan Birman. *Notices of the American Mathematical Society*, Vol. 66, No. 3, 2019.

[Mar22] Bruno Martelli. *An Introduction to Geometric Topology.* CreateSpace Independent Publishing Platform, 2022.

[Mil82] John Milnor. Hyperbolic geometry: The first 150 years. *Bulletin of the American Mathematical Soceity*, Vol. 6, No. 1, 1982.

[Mor98] Shigeyuki Morika. *Geometry of Differential Forms.* American Mathematical Society, 1998.

[Net99] Reviel Netz. *The Shaping of Deduction in Greek Mathematics: A Cognitive History.* Cambridge University Press, 1999.

[Pat67] C.W. Patty. Isotopy classes of imbeddings. *American Mathematical Society*, Vol. 128, No. 2, 1967.

[Poi95] Henri Poincaré. Analysis situs. *Journal de l'École Polytechnique*, Vol. 1, No. 1, 1895.

[Ste15] Alma Steingart. Inside: Out. *Grey Room*, 2015.

[Ste23] Alma Steingart. *Axiomatics: Mathematical Thought and High Modernism.* University of Chicago Press, 2023.

[Tu07] Loring W. Tu. *An Introduction to Manifolds.* Springer, 2007.