SURFACES AND THEIR SYMMETRIES: AN INTRODUCTION TO MAPPING CLASS GROUPS

TARA E. BRENDLE AND LEAH CHILDERS

1. Introduction

An overarching theme in mathematics is that one can learn a vast deal about an object by studying its group of symmetries. For example, in abstract algebra we study two fundamental objects in mathematics, a finite set, and a regular polygon, via symmetric groups and dihedral groups, respectively.

The primary goal of this chapter is to introduce the *mapping class group* $\mathrm{Mod}(S)$, that is, the group of symmetries of another fundamental object: a surface S. We will acquaint the reader with a few of its fascinating properties and give a brief glimpse of some active research related to this class of groups.

We do not assume a background in topology. Therefore in Section 2 we give an introduction to surfaces and explain the concept of a homeomorphism, our working notion of "sameness" for surfaces. If already familiar with these notions, the reader may safely skip to Section 3, where we give some examples of homeomorphisms which will play an important role in what follows. In Section 4 we define mapping class groups. Finally, we discuss some interesting facts and open problems related to mapping class groups in Section 5.

The group $\mathrm{Mod}(S)$ is connected to many areas of mathematics, including complex analysis, dynamics, algebraic geometry, algebraic topology, geometric topology (particularly in the study of 3- and 4-dimensional spaces), and group theory. Within geometric group theory, the close relationships between $\mathrm{Mod}(S)$ and groups such as braid groups, Artin groups, Coxeter groups, certain groups of matrices such as $\mathrm{SL}(n, \mathbb{Z})$, and automorphism groups of free groups, have proved to be a fascinating and rich area of study. We refer you to Farb and Margalit's excellent book *A Primer on Mapping Class Groups* [8] for further details and references on many topics mentioned in this chapter. Although their text is aimed at graduate students and researchers, large portions of it are accessible to undergraduates.

2000 *Mathematics Subject Classification.* Primary: 20F36; Secondary: 57M07.

Acknowledgements. We would like to thank Matt Clay and Dan Margalit for inviting us to be a part of this project, for their great patience, and for their many thoughtful suggestions for the improvement of this chapter. We are also grateful to Joan Birman and Saul Schleimer for helpful conversations. We would also like to thank Mante Zelvyte for comments on an earlier draft of this chapter.

2. A brief user's guide to surfaces

The word "surface" comes from the French for "on the face." Indeed, we all have an intuitive notion of a surface as the outermost layer of an object, as when we speak of resurfacing a road, or as the boundary between two substances, such as the surface of the sea. Each of these kinds of surfaces is inherently two-dimensional in nature, and mathematicians think of surfaces in similar terms.

For the sake of getting on with things, we will wholeheartedly embrace the philosophy that a picture is worth a thousand words in giving our first rough definition: a *surface* is one of the subsets of $\mathbb{R}^3$ given in Figure 1. Note that these are all "hollow;" we are imagining the outer layer rather than a solid object.

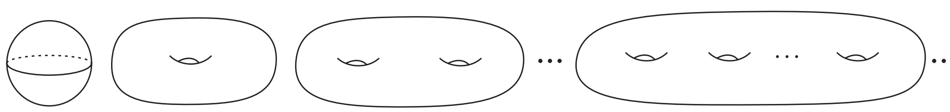

Figure 1. A list of surfaces.

The sphere and the torus. The leftmost surface is familiar to us as the *sphere* S^2. We can think of S^2 as the set of points in $\mathbb{R}^3$ which are distance 1 from the origin. The next surface is the *torus* T, which may be familiar from calculus as a surface of revolution. For example, you can obtain a torus by taking the circle of radius 1 in the xy-plane centered at the point $(2, 0) \in \mathbb{R}^2$, and revolving it around the y-axis in $\mathbb{R}^3$.

Higher genus. The torus T is often described as the frosting on a doughnut, without the doughnut. The doughnut illustration is useful for obtaining another infinite family of examples, by imagining (the frosting of) a doughnut with any number of holes, as shown in Figure 1. The number of holes is the *genus* of the surface.

The list of Figure 1 depicts surfaces increasing in genus. The sphere S^2 has genus 0. The next surface, with genus 1, is the torus T. The torus T is followed by surfaces of genus 2 and higher, which are sometimes referred to as "higher genus tori."

Three different tori? Consider the three subsets of $\mathbb{R}^3$ shown in Figure 2. According to our working definition of a surface, only the second of these is a surface, since the first and the third do not exactly match anything on our list in Figure 1. The first is much skinnier than our torus, yet we can still recognize its basic donut shape. If we "inflate" the first torus until it looks like the second torus, we get a *homeomorphism*, that is, a continuous function with continuous inverse, from the first to the second. (In this case the inverse map from the second to the first is obtained simply by "deflating.") Topologists say that two surfaces are the "same," or *homeomorphic*, if there is a homeomorphism from one to the other. So the first two subsets of $\mathbb{R}^3$ are homeomorphic; we can safely and accurately say each is a torus. But what about the third subset? We claim that it is also homeomorphic to the other two.

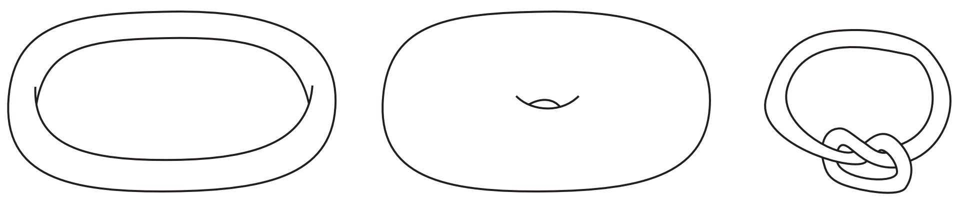

FIGURE 2. Three tori.

For a moment, we imagine the first torus as a flexible hollow tube. We cut the tube, tie it in a knot, and then reglue the tube so that every point on one side of the cut is matched up exactly as before to the points on the other side. Homeomorphisms must preserve open sets, and certainly any open sets away from the cut were not disturbed by this process. But, by careful regluing, we also have not changed any of the open sets, or *neighborhoods*, of points where we cut. In fact, this process gives a homeomorphism from a "standard" torus to a knotted torus. The main point is that the proverbial near-sighted bug of Topology 101 cannot tell the difference between the two, because all neighborhoods remain the same. We will return to this point in the next section when we talk about homeomorphisms known as "Dehn twists."

We can summarize the preceding discussion by expanding our working definition of a surface as follows: a *surface* is any subset of $\mathbb{R}^3$ which is homeomorphic to one of the subsets in Figure 1.

Exercise 1. *Determine the genus of each of the two surfaces shown in Figure 3.*

The reader who is uncomfortable with any of the topological terminology we've used so far (neighborhood, homeomorphism, near-sighted bugs...) has two options: either skip ahead to the Appendix for further details about surface topology, from Christopher Columbus to Klein bottles, or else soldier on, armed with the following take-home point: it's really the examples of homeomorphisms we give in the next section that one should focus on.

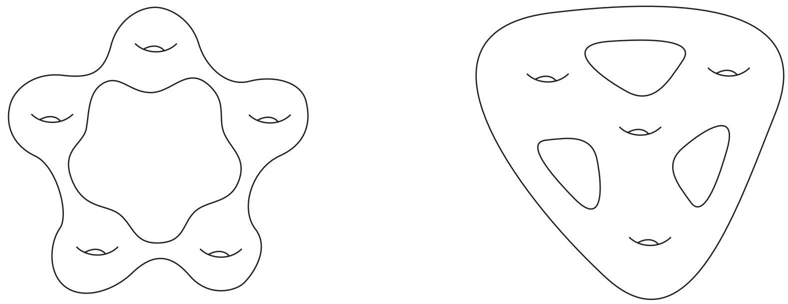

FIGURE 3. Two surfaces.

3. Examples of homeomorphisms of surfaces

So far we have been living in the world of topology, but the notion of homeomorphism of a surface immediately leads us to a group. The set Homeo(S) of all homeomorphisms from a surface S to itself is closed under the operation of function composition. Even better, Homeo(S) carries a group structure under this operation:

- **Associativity.** When defined, function composition is always associative.
- **Identity.** The identity map $\mathrm{id}_S : S \to S$ which sends every point to itself serves as the identity element in this group.
- **Inverses.** Any map which is a homeomorphism comes equipped by definition with an inverse which is also a homeomorphism.

The group Homeo(S) is still not quite the group we want, but before we say more about that, we introduce several important examples of elements in the group Homeo(S).

Examples coming from isometries of $\mathbb{R}^3$. Some of the easiest surface homeomorphisms to visualize arise from nice embeddings of surfaces in $\mathbb{R}^3$. If we arrange our genus g surface as in Figure 4, we can rotate it by $2\pi/g$, or by one "click," to obtain an element of order is g in the group Homeo(S).

Another example of a rotation is the *hyperelliptic involution* given by "skewering" the surface about the axis indicated in Figure 5 and rotating it by π.

Reflections of $\mathbb{R}^3$ can also give rise to homeomorphisms of a surface. As in Figure 6 below, we can just reflect across a plane that slices the surface in half. This homeomorphism is fundamentally different from the others we have discussed so far because the orientation of the surface has been reversed. (If you think of writing a word on the surface, then after the reflection the words will be reversed in the same way that words look backwards in a mirror.) Orientation-reversing homeomorphisms are often excluded from the discussion of symmetries of a surface, or

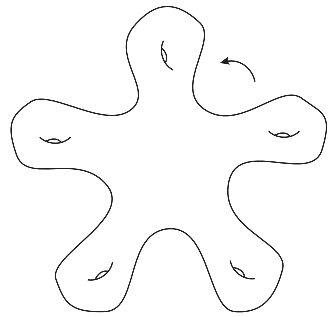

FIGURE 4. Rotation by $2\pi/g$ about the "center" of the surface pictured is a homeomorphism.

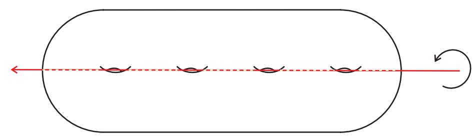

FIGURE 5. Rotation by π about the indicated axis is a hyperelliptic involution.

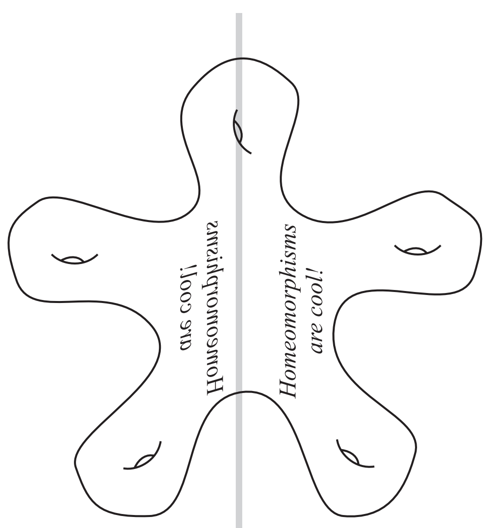

FIGURE 6. The surface is reflected across the vertical plane indicated. This homeomorphism reverses the orientation.

treated separately. Since a precise definition of orientation is beyond the scope of this chapter, we will follow the pack and say very little about this type of map.

Curves and twists. A *simple closed curve* on a surface S is the image of a circle in the surface under a continuous, injective function; three examples are shown in

Figure 7. For practical purposes we can picture a simple closed curve on a surface S as a "loop" on the surface that does not intersect itself.

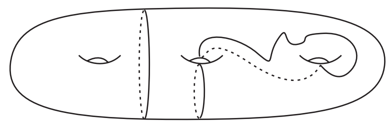

FIGURE 7. Example of three simple closed curves on a surface.

Any simple closed curves on a surface S gives rise to an important example of an element of the group Homeo(S). Imagine cutting a surface along a simple closed curve α, giving the surface a full 360-degree twist, and then carefully reglueing, as shown in Figure 8.

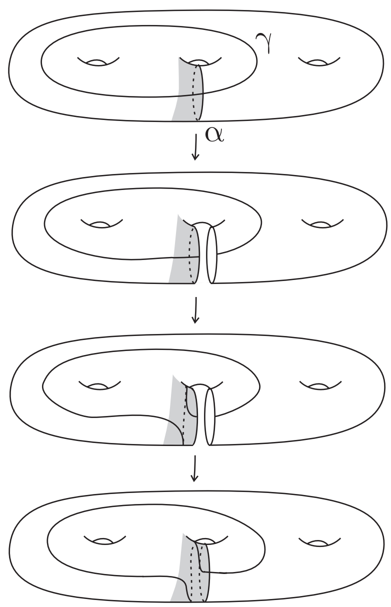

FIGURE 8. A Dehn twist seen as cut along α, twist, and reglue. The simple closed curve γ intersecting α acquires an extra twist about α.

Neighborhoods of points on S might be stretched or twisted, but are never "broken" in the process, since neighborhoods which are separated in the cutting are carefully reunited again when reglueing. Therefore the resulting map of the surface S to itself is a homeomorphism, known as a *Dehn twist about* α, denoted T_α. (We can compare this example with the homeomorphism described at the end of Section 2, which also involved cutting and reglueing.)

Dehn twists via annuli. We can make this precise as follows. Using polar coordinates (r,θ) for points in the plane $\mathbb{R}^2$, we consider the annulus A made up of those points with $1 \le r \le 2$. Then we can define a map T_A by

$$(r,\theta) \mapsto (r, \theta - 2\pi r)$$

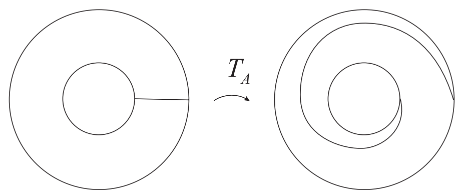

FIGURE 9. A Dehn twist on an annulus.

The important thing to notice is that *each point on the boundary of the annulus A is fixed by the map T_A*. This means that once we do our twisting on the annulus A, we can obtain an element of Homeo(S) by "extending by the identity," that is, by fixing every other point on S outside of A. The point is that our twist on the annulus A and the identity map on $S \backslash A$ agree where they meet, on the boundary of the annulus A.

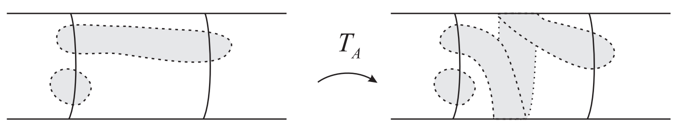

FIGURE 10. A Dehn twist preserves open sets.

But this discussion was supposed to be about simple closed curves, not annuli. The key realization is that with our working definition of a surface S, every simple closed curve α in S is the core [1] of some annulus A, as in Figure 12. (Another way to say this is that we are only considering *orientable* surfaces, that is, surfaces which do not contain a Möbius band.)

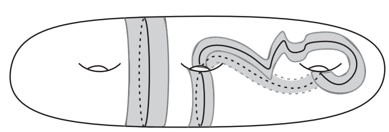

FIGURE 11. Three simple closed curves with their corresponding annuli.

So, given a simple closed curve α, we find a corresponding annulus A, and now we're in business: we can do the Dehn twist T_A which is an element of Homeo(S). In fact, the map T_A we have just defined is really just T_α, a Dehn twist about the curve α as defined above. To see this, look again at Figure 8. We can understand this map

[1] In our previous discussion, the *core* of A in the plane $\mathbb{R}^2$ is the set of points with $r = \frac{3}{2}$.

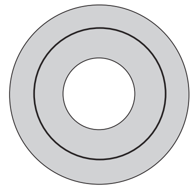

FIGURE 12. The core of an annulus.

by seeing what happens to a simple closed curve γ which crosses α: away from A, nothing happens to γ, but as it nears α, the simple closed curve γ suddenly turns and "traces" α before continuing on its way.

Well-definedness of twists. If you were paying close attention in the previous paragraph, you will have spotted that we were very careful to talk about "a" twist about a simple closed curve α. This is because given α, we have to choose a corresponding annulus A, and each choice of A results in a slightly different way of performing the twist. And if you were paying really close attention, you will also have noted that our definition of T_α also depends on a *parametrization* of A, that is, a homeomorphism which tells us how to go between the annulus A which actually lies in the surface S and the annulus which lives in the plane $\mathbb{R}^2$ which we used to define our twist in the first place. For each step, we actually have an (uncountably) infinite number of choices, each giving rise to a slightly different element of Homeo(S). Ugh!

Isotopy to the rescue. The idea of defining a homeomorphism by twisting a surface about a simple closed curve α is a simple and intuitive one. Yet the whole point of the preceding discussion is that there is no single element of Homeo(S) which we can sensibly refer to as "the" Dehn twist about α because of all the choices involved. In trying to write it down precisely, our definition suddenly became very technical with lots of details to worry about – it would be nice to be able to talk about a Dehn twist without having to make reference to some parametrized annulus every time! This suggests that we have not yet arrived at the correct notion of symmetries of a surface S. The group Homeo(S) is simply too large.

The concept of *isotopy*, explained in the next section, allows us to deal with these difficulties in an efficient and elegant manner.[2] Isotopy also leads us to our main goal: the mapping class group.

[2]A more cynical person might say that isotopy allows us to be lazy and avoid these annoying details!

4. Mapping class groups

In our quest to define the appropriate notion of symmetries on a surface, we have seen, through the example of Dehn twists, that the group Homeo(S) is somehow much too large. In order to address this, we would like to group together homeomorphisms that are in some sense the same, and declare them to *be* the same. In other words, we are going to introduce an equivalence relation, called isotopy, on the set Homeo(S). The goal is to distill Homeo(S) into a more manageable group that still incorporates all the essential features of Homeo(S).

Isotopy. We like to think of isotopy as the technical tool which allows us to get away with not being very good artists when drawing simple closed curves and surfaces – a bump here or a wiggle there does not matter; drawing objects to scale is unimportant. Informally, we say two simple closed curves on a surface are *isotopic* if one can be "deformed" to the other; see Figure 13 for examples and non-examples. One way to think of this is to imagine that the simple closed curve on the surface

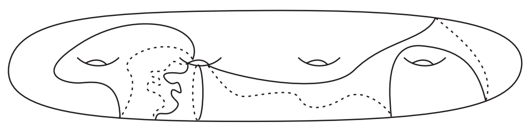

Figure 13. The left simple closed curve is not isotopic to the other three curves, which are all pairwise isotopic.

is made of a rubber band. If you stretch the rubber band and move it around you will get a new curve isotopic to the original.

More precisely, isotopy is a continuous deformation of one simple closed curve to another in such a way that at each stage, we still have a simple closed curve, as opposed to, say, letting the curve intersect itself at some point along the way as in the example of Figure 14.

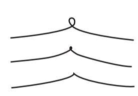

Figure 14. A continuous deformation that is not an isotopy.

Two elements $f, g \in \text{Homeo}(S)$ are *isotopic* if the simple closed curve $f(\alpha)$ is isotopic to $g(\alpha)$ for all simple closed curves α on the surface S. (*A priori*, there are infinitely many such curves to check, but it turns out you can get away with only checking finitely many.)

An equivalent definition is that two homeomorphisms $f, g : S \to S$ are isotopic if there is a continuous function $I : S \times [0, 1] \to S$ where $I(x, 0) = f(x), I(x, 1) = g(x)$, and for each parameter value $t \in [0,1]$, the map $I(x,t)$ is a homeomorphism. In words, two homeomorphisms are isotopic if you can "deform" one to the other, through a continuous family of homeomorphisms.

Exercise 2. *Show that isotopy is an equivalence relation on the set* $\mathrm{Homeo}(S)$ *of all homeomorphisms from the surface* S *to itself.*

Note that we use t as our parameter because we often think of an isotopy as playing a movie where at time $t = 0$, we see the first map f, and then we watch f slowly being deformed so that by the time $t = 1$, we have arrived at the map g. Note that this second definition of isotopy can be generalized easily to other maps, for example from the circle S^1 to a surface S; this allows us to write down a similarly precise definition for isotopy of simple closed curves.

Mapping class groups. A *mapping class* of a surface S is an isotopy class of homeomorphisms from the surface S to itself. If $h \in \mathrm{Homeo}(S)$, we let $[h]$ denote the set of all homeomorphisms from S to S that are isotopic to h, and we say that $[h]$ is the mapping class of the homeomorphism h. Alternatively we say that the homeomorphism h *represents* the mapping class $[h]$.

The set of all mapping classes of a surface S is denoted $\mathrm{Mod}(S)$. Since the elements of $\mathrm{Mod}(S)$ are classes of homeomorphisms, we will use composition of homeomorphisms to help define a group operation on the set $\mathrm{Mod}(S)$. If $f, g \in \mathrm{Homeo}(S)$, and if $[f], [g] \in \mathrm{Mod}(S)$ are their respective mapping classes, then we can define an operation on $\mathrm{Mod}(S)$ as follows:

$$[f] \cdot [g] = [f \circ g].$$

This operation is usually referred to as the composition of mapping classes, and it is not too hard to show that it is well defined and associative, that the mapping class of the identity function on S is the identity in $\mathrm{Mod}(S)$, and that $[f]^{-1} = [f^{-1}]$ for any mapping class $[f] \in \mathrm{Mod}(S)$. Thus $\mathrm{Mod}(S)$ together with this operation is known as the *mapping class group of the surface* S. The notation $\mathrm{Mod}(S)$ is short for *Teichmüller modular group*, an alternative name sometimes used for this group.[3]

Dehn twists as mapping classes. Returning to our example of Dehn twists, let us consider a simple closed curve α in a surface S. Recall that in defining a Dehn twist corresponding to α, we had to make choices: an annulus A with core α, and a parametrization of the annulus A, and the resulting homeomorphism depends heavily on these choices. We seemed to have a serious problem: in the context of homeomorphisms, it made no sense to talk about "the" Dehn twist about the simple closed curve α. Rather, we obtained an uncountably infinite number of different Dehn twists about α!

[3]What we have defined here is often referred to as the *extended mapping class group*. In much of the literature, the term *mapping class group* refers only to isotopy classes of orientation-preserving homeomorphisms. In this chapter, we will use the notation $\mathrm{Mod}^+(S)$ if we wish to restrict to the orientation-preserving case.

However, *the isotopy class of the resulting homeomorphism is independent of both choices*, although it is a somewhat tedious exercise to prove this carefully. In other words, while it does not make sense to talk about "the" Dehn twist about α in the context of Homeo(S), it *does* make sense in the context of Mod(S). Even better, it turns out that if α' is another simple closed curve in the surface S which is isotopic to α, then their corresponding Dehn twists are also isotopic! So not only can we choose whatever annulus and whatever parametrization we like, we are also free to choose any simple closed curve that is isotopic to α. Thus given any simple closed curve α on a surface S, we can safely write T_α as a well defined element of Mod(S) (although we probably should write $T_{[\alpha]}$ to emphasize that it is only the isotopy class of α we care about).

We have finally arrived at the correct notion of the group of symmetries of a surface S: it is the mapping class group Mod(S).

Before we give examples of mapping class groups we would like to make sure you really understand Dehn twists (they really are that important!). So for practice see if you can do the following exercise.

Exercise 3. *Take the simple closed curve α to β using Dehn twists about the simple closed curves α, β, and γ (or their inverses) shown in Figure 15.*

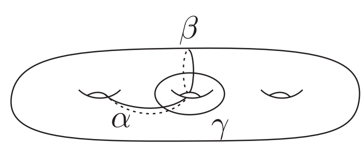

FIGURE 15. The simple closed curves α, β, and γ.

While this exercise sounds simple enough it does take some practice. You may need to twist about some curves more than once, or not at all. There is a really fun computer program called Teruaki, written by Kazushi Ahara of Meiji University, which allows you to play around with Dehn twists (at the time of this writing, it is available for free from his website [1]).

5. Fun facts and open questions.

As noted in the introduction, the study of Mod(S) is a vast area of current research intersecting many branches of mathematics. We will end here with a somewhat random collection of interesting facts about Mod(S), leading up to a brief sample of open questions related to its structure. We begin with some remarks about certain low-genus cases. Details and proofs and/or references to proofs of all facts mentioned here can be found in [8].

The Sphere. It is intuitively clear that any simple closed curve on the sphere S^2 is isotopic to any other simple closed curve on S^2 – sketching and staring at a few pictures should convince you, although writing down a careful proof is a nontrivial exercise. The deep and powerful Jordan-Schönflies Theorem tells us that any simple closed curve on S^2 separates it into two disks. A short lemma known as the *Alexander trick* tells us that all orientation-preserving homeomorphisms of a disk are isotopic (as a technical point, homeomorphisms of a disk and isotopies should all fix the boundary of the disk). In other words, $\mathrm{Mod}^+(S^2)$ is just the trivial group!

The Torus. We next consider the example of the torus T. Consider the two simple closed curves in the torus of Figure 16.

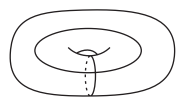

FIGURE 16. Two simple closed curves on a torus.

The Dehn twist about each of these simple closed curves is nontrivial and in fact has infinite order in $\mathrm{Mod}(T)$. Dehn twists are orientation-preserving and so can never generate all of $\mathrm{Mod}(S)$, for any surface S. But it turns out that these two Dehn twists actually generate $\mathrm{Mod}^+(T)$. In other words, any homeomorphism of the torus is *isotopic* to some finite product of these two Dehn twists (or their inverses – this is just achieved by twisting the annuli in the other direction). This may seem like a surprising result, but we turn to linear algebra for some guiding intuition.

The torus is the only surface under our consideration that admits a Euclidean metric; the torus can be realized for example as the quotient of the plane $\mathbb{R}^2$ by the integer lattice $\mathbb{Z}^2$. The general linear group $\mathrm{GL}(2, \mathbb{Z})$ consisting of all invertible 2×2 integer matrices is the group of automorphisms of $\mathbb{Z}^2$.

In fact, $\mathrm{Mod}(T)$ is isomorphic to $\mathrm{GL}(2, \mathbb{Z})$, and $\mathrm{Mod}^+(T)$ is isomorphic to $\mathrm{SL}(2, \mathbb{Z})$, the index-two subgroup of $\mathrm{GL}(2, \mathbb{Z})$ consisting of those matrices with determinant 1. It is classically known that $\mathrm{SL}(2, \mathbb{Z})$ is generated by the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

The isomorphism between the two groups can be defined by sending the Dehn twists about the two simple closed curves in Figure 16 to these matrices. For the complete proof that this defines an isomorphism, see ([8], Section 2.2).

This example illustrates that understanding the simple closed curves on a surface S turns out to be extremely helpful in understanding both S itself and its symmetries. There is a powerful analogy between aspects of surface topology and the vector space $\mathbb{R}^2$. This should not come as a great surprise, since surfaces are inherently planar. In many important ways, simple closed curves are to surfaces as vectors are to vector spaces.

Dehn twists in higher genus. Perhaps more surprising is the fact that $\mathrm{Mod}^+(S)$ is generated by Dehn twists for any surface S. In other words, we can view Dehn twists as the building blocks of all mapping classes: every element of $\mathrm{Mod}^+(S)$ can be written as a *finite* product $T_{\alpha_1}^{e_1} T_{\alpha_2}^{e_2} \cdots T_{\alpha_n}^{e_n}$ for some simple closed curves $\alpha_1, \ldots, \alpha_n$ and $e_i = \{\pm 1\}$. Even better, it turns out we can choose the α_i from some finite list of simple closed curves. Humphries has shown that every mapping class can be generated by Dehn twists about the $2g + 1$ simple closed curves in Figure 17 [11]. To get some appreciation for the beauty of this fact, try to find a product of these

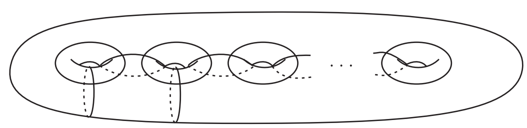

FIGURE 17. Dehn twists about these $2g + 1$ simple closed curves generate the mapping class group.

Dehn twists which achieves the rotation of Figure 5 – it is not at all obvious how to do this (you are allowed to cheat and look it up in [8] after 10 minutes)! It is often said that the idea that any mapping class in $\mathrm{Mod}^+(S)$ can be obtained by a finite sequence of Dehn twists about various curves is analogous to the idea of being able to solve a Rubik's cube in a finite number of moves.

One way to prove that Dehn twists generate $\mathrm{Mod}^+(S)$ is to study its action on nice sets built out of simple closed curves, which we describe next.

The Complex of Curves. For any surface S, we can define an infinite graph, which we denote $\mathcal{C}^1(S)$. Here a *graph* means a collection of points, or *vertices*, with some *edges* connecting them. The graph $\mathcal{C}^1(S)$ has one vertex for each isotopy class of simple closed curves on S, and we join two vertices by an edge if (and only if) we can find a pair of simple closed curves, one representing each vertex, that are disjoint.

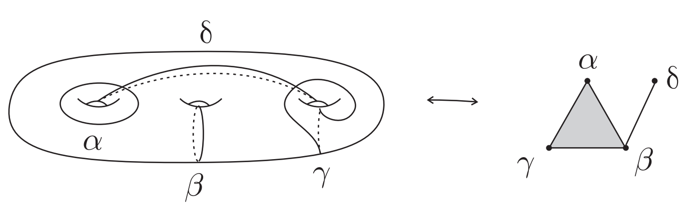

FIGURE 18. Four simple closed curves on a surface and their corresponding span in the complex of curves.

When you are new to surface topology, it might seem like the graph $\mathcal{C}^1(S)$ could be finite. In fact, it has infinitely many vertices and edges. Even more, the graph $\mathcal{C}^1(S)$ is *locally infinite*: each vertex is adjacent to infinitely many edges!

Exercise 4. *Sketch infinitely many distinct, non-isotopic simple closed curves on a surface. Then find a curve disjoint from infinitely many distinct, non-isotopic simple closed curves on a surface. (If you've done the first part cleverly, you can reuse it in the second part.)*

Our graph $\mathcal{C}^1(S)$ is a 1-dimensional *simplicial complex*. We can add dimensions by turning $\mathcal{C}^1(S)$ into a special kind of simplicial complex called a *flag complex*. This just means that we glue in a "filled-in" triangle each time we see the outline of a triangle in our graph, and then we fill in a solid tetrahedron every time we see the boundary of one, and so on in higher dimensions. Figure 18 shows the subcomplex of the curve complex of a genus 3 surface spanned by 4 curves.

The resulting structure, known as the *complex of curves* $\mathcal{C}(S)$, records all the combinatorial structure of intersection patterns of simple closed curves on a surface S. Formally, $\mathcal{C}^1(S)$ is the 1-*skeleton* of $\mathcal{C}(S)$, as the objects which make up $\mathcal{C}^1(S)$ are at most 1-dimensional.

A major theme in geometric group theory is that one can learn much about a group's structure, for example its generating sets and relations, by studying its various "nice" actions on "nice" sets. Since homeomorphisms take curves to curves, and preserve disjointness, the group $\mathrm{Mod}(S)$ acts nicely on $\mathcal{C}(S)$, in the sense that it takes vertices of $\mathcal{C}(S)$ to vertices, edges to edges, etc. Hence we can use this action to write down, for example, a complete proof of the generation of the mapping class group by Dehn twists via the complex of curves ([8], Chapter 4).

The importance of the curve complex is also revealed by an incredibly useful theorem of Ivanov, that the *simplicial automorphism group* $\mathrm{Aut}(\mathcal{C}(S))$ is isomorphic to $\mathrm{Mod}(S)$ [12]. In other words, the combinatorial data of curves and their intersections encodes the entire algebraic structure of $\mathrm{Mod}(S)$.

Relations among Dehn twists. Since Dehn twists generate $\mathrm{Mod}^+(S)$, it is important to understand their algebraic properties. It turns out that we can completely characterize how Dehn twists interact with each other algebraically in terms of combinatorial properties of the underlying simple closed curves. In what follows, everything we say is "up to isotopy." For example, we will apply the term "disjoint" to any two curves α and β which can be isotoped to be disjoint from each other.

Now, if two curves are disjoint, then twisting about one has no effect on the other, so you can do your twisting in either order. In other words, Dehn twists about disjoint curves commute.

What if the two curves intersect? We will consider the easiest case, where two curves intersect once. We will need the following useful fact:

Fact 5.1. *For any $f \in \mathrm{Mod}(S)$ and any simple closed curve α in S we have*

$$T_{f(\alpha)} = f T_\alpha f^{-1}.$$

A bit of thought will convince you that this equation does not really require proof: following a homeomorphism to another copy of S, doing the Dehn twist there, and then going back again, is the same as if you just Dehn twist about the image of your curve under the very same homeomorphism .

Now we can formally state and prove the *braid relation* for Dehn twists.

Theorem 5.2 (Braid relation)**.** *If α and β are simple closed curves in S that intersect exactly once, then $T_\alpha T_\beta T_\alpha = T_\beta T_\alpha T_\beta$.*

Proof. The relation $T_\alpha T_\beta T_\alpha = T_\beta T_\alpha T_\beta$ is the same as $T_\beta^{-1} T_\alpha T_\beta = T_\alpha T_\beta T_\alpha^{-1}$. Using Fact 5.1, this is the same as $T_{T_\beta^{-1}(\alpha)} = T_{T_\alpha(\beta)}$. So to complete the proof we just need to verify that $T_\beta^{-1}(\alpha)$ and $T_\alpha(\beta)$ are the same curves. To do this we simply pick two curves that intersect once and verify that the resulting curves are the same. In Figure 19 we show one such choice for the simple closed curves α and β; a quick calculation shows the simple closed curve $T_\beta^{-1}(\alpha)$ is indeed the same as $T_\alpha(\beta)$, as in the second part of the same Figure.

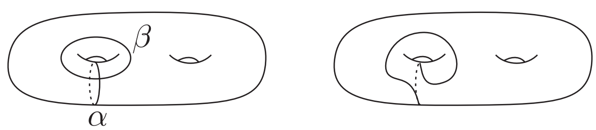

FIGURE 19. The simple closed curves α, β, and $T_\beta^{-1}(\alpha) = T_\alpha(\beta)$.

It might seem like cheating that we have only checked what happens for just one pair of curves, while in fact there are infinitely many pairs of curves on a surface that intersect exactly once. However, for every such pair of curves, there is a homeomorphism of the surface taking them to “our” pair. This is known as the *change of coordinates principle*; it is similar to the principle of changing bases in linear algebra. □

As an aside, we note that this relation is named for *braid groups*, which are mapping class groups of disks with n marked points, if we insist that all maps involved fix the disk’s boundary pointwise and fix the n marked points setwise. In this way, a traditional braid, thought of as intertwined strands, is a bit like watching a movie of n points in a disk moving around the disk without bumping into each other, and ultimately returning (as a set) to their points of origin. Braid groups have been studied classically in their own right and have a rich interaction with the types of mapping class groups we have focused on here.

We are left with the case of curves which intersect at least twice. It turns out that in this case, we get no relations whatsoever! One proof of this fact relies on the so-called Ping-Pong Lemma [8].

Is that discussed in another chapter in this volume? If so, give reference? Johanna is doing it. Don't know a chapter number

We summarize this discussion, as follows.

(1) Two curves α and β are disjoint if and only if $T_\alpha T_\beta = T_\beta T_\alpha$.
(2) Two curves α and β intersect exactly once if and only if $T_\alpha T_\beta T_\alpha = T_\beta T_\alpha T_\beta$.
(3) Two curves α and β intersect more than once if and only if T_α and T_α satisfy no non-trivial equations. (Another way to say this is that T_α and T_β generate a free group.)

See ([8], Chapter 3) for further details about the above facts and for more information about Dehn twists.

A surprising fact about Dehn twists. We have perhaps created the false impression that everything about $\mathrm{Mod}(S)$, or at least everything about Dehn twists, is well understood and somewhat intuitive. This is far from true! $\mathrm{Mod}(S)$ exhibits some strange phenomena. Mathematicians have studied Dehn twists for nearly a century, but it was only in 2009 that a strange phenomenon was discovered: Dehn twists have roots [15]! In other words, for any Dehn twist T_α, there is some element $f \in \mathrm{Mod}(S)$ such that $T_\alpha = (f)^k$ for some $k \in \mathbb{Z}$.

In fact, a vast number of basic questions related to mapping class groups remain unknown. In what follows we give a brief sample.

Open questions. We end this chapter with a brief mention of a few easy to state (and not at all original) open problems, but one only needs to open the book *Problems on Mapping Class Groups and Related Topics* [7] to find literally dozens upon dozens of unanswered questions about this fascinating class of groups. This book will also lead you to the connections between mapping class groups and fields ranging from dynamics to algebraic geometry.

Linearity. When encountering a new group, the first question we should ask is: is this group familiar? Have we ever encountered an isomorphic copy of it in some other totally different context? Earlier in this section, we saw that in the case of genus 1, the mapping class group $\mathrm{Mod}(T)$ is an example of a *linear group*, that is, it is isomorphic to a multiplicative group of matrices $\mathrm{GL}(n, \mathbb{C})$, or one of its subgroups, for some natural number n. This fact was known classically, but it was only relatively recently that Bigelow and Budney [3] gave a proof that when S is a surface of genus 2, $\mathrm{Mod}(S)$ is also a linear group, although a much more complicated one – their proof requires n to be at least 64.

There is a great deal of mathematical literature dedicated to demonstrating that mapping class groups share just about every conceivable property possessed by finitely generated linear groups, but yet the following question is currently unsolved.

Open Question 1. *Is* $\mathrm{Mod}(S)$ *a linear group for a surface* S *of any genus?*

Other generating sets. Recall that the mapping class group can be generated by $2g+1$ Dehn twists. A reasonable question is whether we can do any better: can we generate $\mathrm{Mod}^+(S)$ by a smaller set of elements? It turns out we can't do better with Dehn twists, but one can generate all of $\mathrm{Mod}^+(S)$ with just two elements if we allow other types of elements [19].

A discussion of various notions of "small" generating sets can be found in the introduction to [4]. As a sample, we can consider generating sets consisting of *involutions*, or elements of order 2, such as the hyperelliptic involution shown in Figure 5. Various mathematicians have given generating sets for $\mathrm{Mod}^+(S)$ consisting of involutions; for instance, Kassabov [13] and Monden [17] have shown that for certain surfaces S, just 4 involutions suffice to generate $\mathrm{Mod}^+(S)$.

Open Question 2. *Can* $\mathrm{Mod}^+(S)$ *be generated by fewer than four involutions?*

Relations between higher order Dehn twists. Recall that we discussed many relations involving Dehn twists of two simple closed curves. Similar questions can be asked about the k-th powers of Dehn twists.

Open Question 3. *Is there a nontrivial relation between* k*-th powers of Dehn twists for any* k*?*

Homomorphisms onto the integers. For genus at least 3, one can use the famous *lantern relation* in mapping class groups to prove easily that these mapping class groups are *perfect*, that is, their abelianizations are trivial [8]. By the "universal property of abelianization," this fact tells us that there are no homomorphisms from these mapping class groups onto any nontrivial abelian groups, such as the integers $\mathbb{Z}$. However, this does not rule out the possibility that some very large (finite index) subgroup of a mapping class may admit such a homomorphism.

Perhaps surprisingly, the existence of homomorphisms of finite index subgroups of a group onto $\mathbb{Z}$ can have deep implications for the geometry and topology of related structures: this question is related to Kazhdan's Property (T) and to a conjecture of Thurston about aspherical 3-manifolds. Hence the existence of such homomorphisms has been studied for many classes of groups. For example, Gonciulea [9] and Cooper-Long-Reid [5] have proven that infinite Coxeter groups have this property (these two papers are just two of many excellent starting points for references on these topics). Intriguingly, the fact that mapping class groups are generated by involutions (see discussion above) implies that mapping class groups are quotients of infinite

Coxeter groups, but we do not currently know if mapping class groups admit such homomorphisms.

Open Question 4. *Do finite index subgroups of mapping class groups admit a surjective homomorphism onto the integers?*

Distance in the complex of curves. Returning to the complex of curves, there is a simple notion of distance we can define between any two vertices: assign each edge to have length 1, and then define the distance between two vertices, or between curves representing them, to be the length of the shortest path connecting them. (It is not too hard to show that you can always find such paths [8]). By definition, two vertices have distance 1 in the complex of curves if and only they correspond to curves that are disjoint, or at least can be isotoped to be disjoint. It is also not too hard to find examples of curves which have distance 2 in the curve complex: you just need to find two curves which are not disjoint, but which are both disjoint from a third curve.

Exercise 5. *Draw an example of a pair of curves of distance 2 in the curve complex. Hint: if you use a closed surface, you will need its genus to be at least two.*

However, it gets harder when you get to higher distances. The following exercise is much more difficult than the previous one.

Exercise 6. *Draw an example of a pair of curves of distance 3 in the curve complex.*

Need to add references for Open Question 5 and the exercise preceding it...

Open Question 5. *Construct an explicit example of distance n curves.*

Analogs of the complex of curves. The complex of curves $\mathcal{C}(S)$ is a useful tool in studying mapping class groups because of various nice properties as a space on which $\mathrm{Mod}(S)$ acts. For example, $\mathcal{C}(S)$ is analogous to the notion of *buildings* for arithmetic groups, a class of groups which includes many familiar matrix groups such as $\mathrm{SL}(n, \mathbb{Z})$. Another family of groups closely related to mapping class groups are the *automorphism groups of free groups*, and various analogs of the complex of curves exist for these groups. The next problem is more philosophical and open-ended.

Open Question 6. *What is the best analog of the complex of curves for an automorphism group of a free group?*

If you are interested in this question, we refer you to recent work of Bestvina-Feighn [2] and Handel-Mosher [10] as a starting point for discussion of this topic.

6. Appendix: The topology of surfaces

In this section we will make more precise the topological terminology we introduced in Section 2 relating to surfaces. Consider two-dimensional space itself: the Euclidean plane $\mathbb{R}^2$. It has two key properties: first, it comes equipped with a notion of distance, and second, at every point in $\mathbb{R}^2$, the surface clearly resembles a plane – after all, it *is* a plane – in both a purely intuitive sense as well as in a precise mathematical sense, which we will soon describe.

Following this guiding example, we will define a *surface* as a *metrizable* space which is *locally planar*. [4] The term *metrizable* simply means that a space admits some meaningful notion of distance. It is slightly more difficult to explain what we mean by "locally planar," but this concept is the key to understanding not only surfaces but also their symmetries.

Columbus and manifolds. When mathematicians try to explain what we mean by a "local" property of a space, we often end up talking about a small bug who lives in the space and can't see very far. It is helpful to think about early explorers, such as Christopher Columbus. We have all heard the stories that Columbus thought he could sail west from Europe to Asia, because he believed the world was round. At the time, this was thought by many to be nonsense, even heretical, since everyone knew perfectly well that the earth was flat. With the benefit of a few centuries of hindsight, modern man often indulges in a condescending chuckle at Columbus's contemporaries for their ignorance.

But those disbelievers were on to something: the surface of the earth, roughly a sphere, is a surface in the mathematical sense as well. A human being standing on the earth is exactly like a bug who can't see very far – the immediate vicinity, or *neighborhood*, of a human appears to be roughly planar, apart from the odd mountain or canyon. Indeed, early man had intuitively grasped the deep mathematical concept of a manifold. An *n-manifold* is a space that resembles, in this "local" sense, n-dimensional Euclidean space for some natural number n, together with some technical conditions that rule out pathological examples (in our case we avoid these by insisting our surfaces be metrizable). In other words, the term *surface* and *2-manifold* are mathematically interchangeable.

Orientability. We have already mentioned the example of the plane $\mathbb{R}^2$. We will give some further examples for the sake of developing our intuition before making the term "locally planar" precise. For simplicity, we will limit our discussion to orientable surfaces, in which it is impossible to reverse one's sense of right-handedness versus left-handedness. An *orientable* surface is one which does not contain a Möbius

[4] Topologists usually use a slightly more general definition, but this definition is useful for our purposes.

band, that is, a band with a half-twist, as in Figure 20. A *Klein bottle*, shown in Figure 21 is a non-orientable surface. (We're fans of the Acme Klein Bottle company; check them out at www.kleinbottle.com.)

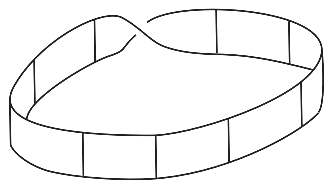

FIGURE 20. A Möbius band.

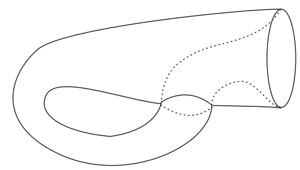

FIGURE 21. A Klein bottle.

Even though the mathematical definition does not require a surface to be situated in any ambient 3-dimensional space, it is often useful to describe them as such. For example,

The Classification of Compact Surfaces. Some surfaces display certain kinds of "infinite" behavior. The Euclidean plane $\mathbb{R}^2$ is infinite in an obvious sense. It literally goes on forever; distances are unbounded. But infinity can rear its ugly head in other ways. For example, we could have a bounded surface with infinite genus as in Figure 22.

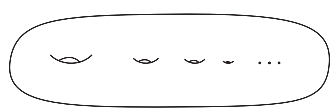

FIGURE 22. A surface with infinite genus.

The open unit disk $\{(x, y) \mid x^2 + y^2 < 1\}$ is also bounded in terms of distance, but is infinite in a more subtle sense: a bug residing in this space can never get to any boundary or end of the space. The same is true of any surface with a point removed. We often draw such a "missing point" as a cusp, as in Figure 23.

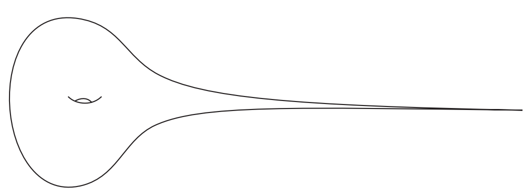

FIGURE 23. A torus with a cusp.

We will rule out these kinds of examples by insisting that our surfaces satisfy a condition called *compactness*. Observe that in each of our examples, the space in question contains at least one sequence without containing its limit point. In the case of $\mathbb{R}^2$, we can use the sequence $\{(n, 0)\}$; for the open unit disk we can use $\{(\frac{1}{n}, 0)\}$. Worse yet, there is no subsequence we can pass to in order to fix this problem. Therefore we say that a space is *compact* if every sequence in the space either converges to a point in the space, or at least contains a subsequence which converges to a point also in the space (if your space has two distinct points x and y, you can always bounce back and forth between them, giving a sequence that never converges).

Open sets in a surface. If you've encountered open sets in the plane $\mathbb{R}^2$ before, then your intuition for open sets in surfaces is probably more or less correct. Picture a surface sitting inside $\mathbb{R}^3$, and draw a disc-shaped subset of S bounded by dashed lines, indicating that the boundary of the disk is not included in the subset. This gives a pretty good notion of what open sets in S look like.

In any space with a distance function d, we can define the *open ball* $B(x, r)$ of radius r about the point x is the set of all points y in the space such that $d(x, y) < r$. In a surface (inspired by $\mathbb{R}^2$), we often speak of *open disks* rather than open balls. We say that a subset U of the space is *open* if for any $x \in U$, we can find an $r > 0$ such that $B(x, r) \subset U$. A *neighborhood* of a point x in a surface S is just an open set containing x. Most of the time, we can just picture an open ball $B(x, r)$.

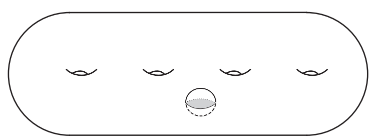

FIGURE 24. The shaded region is an open set in a surface.

Since we've insisted that our spaces are metrizable, we can always take the approach we just described. However, if a surface S happens to be embedded in $\mathbb{R}^3$ (and indeed, the Classification Theorem tells us this can always be achieved), then S inherits the usual Euclidean distance function. In this case, the open sets that result are intersections of open sets in $\mathbb{R}^3$ with S, as in Figure 24. The notion of distance inherited from an embedding may or may not be the same as the distance function we started with. In fact, one can prove that any reasonable choice of distance function yields the same collection of open sets. In other words, the notion of distance is only a means to an end, the "end" being open sets!

We are now in a position to define the first word in our key phrase "locally planar." In mathematics, a space is *"locally X"* if every point in the space has a neighborhood with property X. In other words, the defining structure of a surface, and hence of its symmetries, depends entirely on its open sets.

6.1. **Symmetries, continuous functions, and homeomorphisms.** In general, a symmetry of an object is a bijective function from the object to itself, which preserves the essential features, or defining properties of that object. In geometry, for example, we can rotate a regular n-gon in the obvious way by $2\pi/n$. We can perform a similar operation on the genus g surface as shown in Figure 4, rotating it by $2\pi/g$, or by one "click."

It seems clear that we would want to include the example of Figure 4 in our list of symmetries of a surface. But the underlying structure of a surface does not come from the way it is drawn in 3-space, its structure comes from its open sets. The example above certainly takes an open set to an open set, and so respects the open-set structure in some way.

But the example of Figure 4 does a lot more than simply preserve open sets. If we think about the usual metric on $\mathbb{R}^3$, we see that the example above is actually an *isometry*, that is, it preserves distances. In the previous section, however, we hopefully made it clear that the distance function is not an essential part of a surface's structure. We are only interested in the open sets it produces for us. The example of Figure 4 meets some extra unnecessary criteria; our definition of a "symmetry" of a surface need not be this restrictive.

Continuity is a condition commonly placed on functions to ensure a healthy respect for open sets. A function is *continuous* if the inverse image of an open set is also an open set. However, continuity is not quite enough to ensure that open sets always correspond to open sets. We also need to require that the function is invertible and that its inverse is continuous.

We are led to make the following definition: a *homeomorphism* is a continuous function with continuous inverse (and must therefore be bijective). When topologists say that two spaces are the "same," they mean there exists a homeomorphism from one to the other, as in the canonical joke that a topologist thinks a coffee cup is the same as a doughnut.

We can finally define our key phrase: a space is *locally planar* if every point in the space has a neighborhood which is homeomorphic to an open disk. The choice of terminology makes a bit more sense in light of the following exercise.

Exercise 7. *Show the open disk is homeomorphic to the plane* $\mathbb{R}^2$.

The Classification Theorem. The following incredibly deep theorem, a crowning achievement of early topology, essentially tells us that surfaces are completely determined by their genus.

Theorem 6.1 (Classification of Surfaces)**.** *Every compact orientable surface is homeomorphic to one of the surfaces on the list given in Figure 1.*

The proof of the Classification of Surfaces is nontrivial to say the least. The first complete proof took several decades to complete, with contributions by Möbius [16], Dehn and Heegaard [6], Kerekjarto [14], and Rado [18].

References

[1] Kazushi Ahara. Teruaki. *http://www51.atwiki.jp/kazushiahara/pages/32.html*.

[2] Mladen Bestvina and Mark Feign. Hyperbolicity of the complex of free factors. *ArXiv Mathematics e-prints*. arXiv:math/1107.3308.

[3] Stephen Bigelow and Ryan Budney. The mapping class group of a genus two surface is linear. *Algebr. Geom. Topol*, 1(34):699–708, 2001.

[4] Tara E. Brendle and Benson Farb. Every mapping class group is generated by 6 involutions. *Journal of Algebra*, 278(1):187–198, 2004.

[5] D. Cooper, D. D. Long, and A. W. Reid. Infinite Coxeter groups are virtually indicable. *Proc. Edinburgh Math. Soc. (2)*, 41(2):303–313, 1998.

[6] Max Dehn and Poul Heegaard. Analysis situs. *Encykl. Math. Wiss., III AB3 Leipzig*, pages 153–220, 1907.

[7] Benson Farb. *Problems on mapping class groups and related topics*, volume 74 of *Proc. Sympos. Pure Math.* Amer. Math. Soc., Providence, RI, 2006.

[8] Benson Farb and Dan Margalit. *A primer on mapping class groups.* Princeton University Press, 2011.

[9] Constantin Gonciulea. Infinite Coxeter groups virtually surject onto **Z**. *Comment. Math. Helv.*, 72(2):257–265, 1997.

[10] Michael Handel and Lee Mosher. The free splitting complex of a free group i: Hyperbolicity. *ArXiv Mathematics e-prints.* arXiv:math/1111.1994.

[11] Stephen P. Humphries. Generators for the mapping class group. In *Topology of low-dimensional manifolds (Proc. Second Sussex Conf., Chelwood Gate, 1977)*, volume 722 of *Lecture Notes in Math.*, pages 44–47. Springer, Berlin, 1979.

[12] Nikolai V. Ivanov. Automorphism of complexes of curves and of Teichmüller spaces. *Internat. Math. Res. Notices*, (14):651–666, 1997.

[13] Martin Kassabov. Generating Mapping Class Groups by Involutions. *ArXiv Mathematics e-prints*, November 2003. arXiv:math/0311455.

[14] Béla Kerékjártó. *Vorlesungen über Topologie*, volume 1. J. Springer, 1923.

[15] Dan Margalit and Saul Schleimer. Dehn twists have roots. *Geom. Topol*, 13(3):1495–1497, 2009.

[16] A.F. Möbius. Zur theorie der polyëder und der elementarverwandtschaft. *Gesammelte Werke*, 2:513–560, 1886.

[17] Naoyuki Monden. Generating the mapping class group of a punctured surface by involutions. *ArXiv e-prints*, July 2008. 0807.0916.

[18] T. Radó. Über den begriff der riemannschen fläche. *Acta Litt. Sci. Szeged*, 2:101–121, 1925.

[19] Bronislaw Wajnryb. Mapping class group of a surface is generated by two elements. *Topology*, 35(2):377–383, 1996.

Tara E. Brendle, School of Mathematics & Statistics, University Gardens, University of Glasgow, G12 8QW, tara.brendle@glasgow.ac.uk

Leah Childers, Mathematics Department, Benedictine College, 1201 N 2nd St, Atchison, KS 66002, lchilders@benedictine.edu