Notes on the complex of curves

Saul Schleimer

Author address:

Department of Mathematics, Rutgers
E-mail address: saulsch@math.rutgers.edu

CHAPTER 1

The mapping class group and the complex of curves

1. Introduction

These are the notes I prepared for talks at Caltech, January 2005. As they are doubtless riddled with typos, mathematical errors, misattributions, and other mistakes please do email me with any corrections or other improvements.

I have given references to various articles and books in the course of the notes. There are also a collection of exercises: these range in difficulty from the straight-forward to quite difficult. For the latter I have sometimes given a hint in Appendix A, if I in fact know how to solve the problem!

Much of the discussion that follows is an expository account of two papers of Masur and Minsky, [**30**] and [**31**]. Work-in-progress with Howard Masur is also present, and I thank Jason Behrstock, Jeff Brock, Ken Bromberg, Chris Leininger, Feng Luo, Dan Margalit, Yair Minsky, Hossein Namazi, Kasra Rafi, Peter Storm, and Karen Vogtmann for many interesting and instructive conversations. I further thank Behrstock for pointing out a collection of errors in a previous version of Section 7 of Chapter 2.

Of course all of the remaining errors are my own fault. I again ask you to contact me via email with those you find.

2. Basic definitions

There are many detailed discussions of the mapping class group in the literature: perhaps Birman's book [**3**] and Ivanov's [**22**] are the best known.

Before we introduce the *mapping class group* of a surface, let's recall a few basic notions. A *surface* S is a two-dimensional manifold. Unless otherwise noted, we assume that our surfaces are compact and connected. Typically we shall also require that S be orientable, with non-orientable surfaces relegated to the exercises. Recall that the disk is the surface $\{z \in \mathbb{C} \mid |z| \leq 1\}$ and the annulus is the surface $S^1 \times [0,1]$, circle cross interval.

Suppose S is a surface. A *curve* $\alpha \subset S$ is an embedded copy of the circle S^1. We say α is *separating* if $S \setminus \alpha$ (S cut along α) has two components. Otherwise α is nonseparating. We say α is *essential* if no component of $S \setminus \alpha$ is a disk. We say α is *non-peripheral* if no component of $S \setminus \alpha$ is an annulus. Virtually all curves discussed are assumed to be essential and non-peripheral.

See Figure 1.

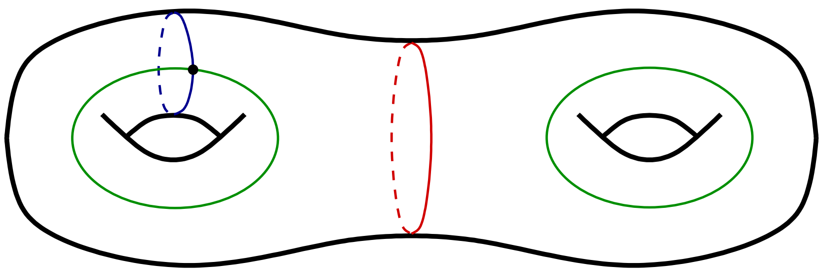

FIGURE 1. Three nonseparating curves and one separating curve in the genus two surface, S_2.

An *arc* α in S is the image of the unit interval I under a proper embedding. An arc α is *essential* if no component of $S \setminus \alpha$ has closure being a disk. (If S is an annulus, then we redefine essential arcs to be those meeting both boundary components.) Of course, a closed surface contains no essential arcs. See Figure 2.

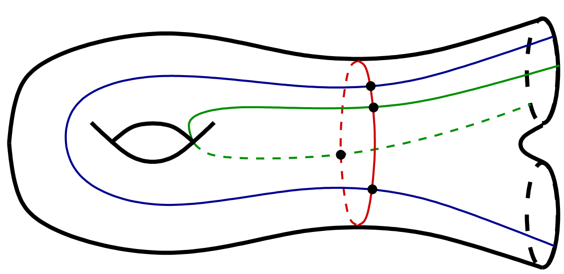

FIGURE 2. A separating curve and a few essential arcs in the twice-holed torus, $S_{1,2}$.

The *intersection number* of two curves or arcs α and β is $\iota(\alpha, \beta)$: the minimal possible number of intersections between α and β' where β' is any curve or arc properly isotopic to β. Note that the intersection number between two curves or arcs is realized exactly when $S \setminus (\alpha \cup \beta)$ contains no *bigons* and no *boundary triangles*. Here a bigon is a disk meeting α and β in exactly one subarc each while a boundary triangle is a disk meeting α, β, and ∂S in one subarc each.

EXERCISE 1.1. Suppose that α, β are any curves in a compact connected surface S. Suppose that β' is isotopic to β with $\alpha \cap \beta' = \iota(\alpha, \beta)$. Then there is an isotopy β_t between β and β' so that $\alpha \cap \beta_t \leq \alpha \cap \beta_s$ for $0 \leq t \leq s \leq 1$.

EXERCISE 1.2. Suppose that α, β are any curves in a compact connected surface S. Suppose that α and β meet once, transversely. Prove that $\iota(\alpha, \beta) = 1$.

Now suppose that $\iota(\alpha, \beta) > 0$. Prove that both α and β are essential and non-peripheral.

The *genus* of a surface S, $g(S)$, is the minimal number of disjoint essential non-peripheral curves required to cut S into a connected planar surface. So the torus $\mathbb{T}^2 = S_1 \cong S^1 \times S^1$ has genus one and the *connect sum* of g copies of the torus has genus g.

As a bit of notation we will use $S_{g,b,c}$ to denote a compact connected surface of genus g with b boundary components and c cross-caps. The classification of surfaces tells us that every surface is homeomorphic to one of these. Typically we take $S_{g,b} = S_{g,b,0}$. Also, if $b = 0$ we simply write S_g. Recall that one useful measure of complexity is $\zeta(S_{g,b}) = 3g + b - 3$.

An *essential subsurface* X in S is an embedded surface where every component of ∂X is contained in ∂S or is essential in S. We do not allow boundary parallel annuli to be essential subsurfaces.

EXERCISE 1.3. Classify all pairs (g, b) so that $S_{g,b}$ contains no essential non-peripheral curves. We will call these the *simple* surfaces.

EXERCISE 1.4. Classify all pairs (g, b) so that in $S_{g,b}$ every pair of essential non-peripheral curves, which are not isotopic, intersect. We will call these the *sporadic* surfaces.

EXERCISE 1.5. Fix a surface S. Prove that for any pair of non-separating curves in S there is a homeomorphism throwing one onto the other. How many kinds of separating curve are there in $S_{g,b}$, up to homeomorphism?

3. The mapping class group

We now spend a bit of time discussing homeomorphisms of surfaces. Before we begin, it is nice to have a few simple examples. So here is a concrete construction of a surface together with accompanying homeomorphism.

Let Γ be a finite, polygonal, connected graph in $\mathbb{R}^3$ and let $h\colon \mathbb{R}^3 \to \mathbb{R}^3$ be an isometry sending Γ to itself. Then we can obtain a closed surface S by taking the boundary of the closed ϵ neighborhood of Γ, $S = \partial N_\epsilon(\Gamma)$. The restriction of the isometry, $h|S$, is a surface homeomorphism. If Γ is not homeomorphic to an interval then h is necessarily of finite order. For example, see Figure 3.

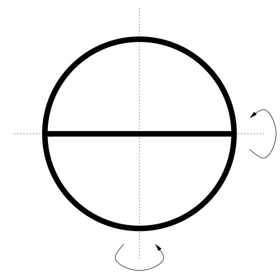

FIGURE 3. The θ-graph gives some symmetries of S_2.

EXERCISE 1.6. Find a finite order homeomorphism of S_2, the closed genus two surface, which *cannot* be obtained in this fashion.

Here is another nice construction: Recall that the torus S_1 can be obtained as the quotient of the plane $\mathbb{R}^2/\mathbb{Z}^2$, where $\mathbb{Z}^2$ acts on the plane by $(n, m) \cdot (x, y) = (x + n, y + m)$. Consider now the linear maps of the plane sending the integer lattice to itself, $\mathrm{SL}(2, \mathbb{Z})$.

EXERCISE 1.7. Classify all finite order elements of $\mathrm{SL}(2, \mathbb{Z})$ up to conjugacy. Thought of as finite order homeomorphisms of S_1, which elements are induced as symmetries of a graph in $\mathbb{R}^3$?

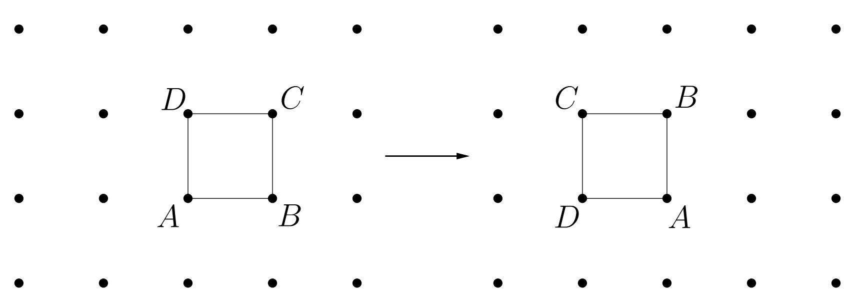

FIGURE 4. An order four map of the torus.

There are three possibilities for elements $A \in \mathrm{SL}(2, \mathbb{Z})$:

- *periodic*: the trace of A is less than 2 in absolute value,

- *reducible*: the trace equals ± 2,
- *Anosov*: the trace is greater than 2 in absolute value.

EXERCISE 1.8. You have probably already shown above that periodic elements are exactly those of finite order. Show that every reducible element leaves a parallel family of circles in S_1 invariant and fixes pointwise exactly one of these.

Of even more interest are the Anosov elements. Here the matrix A has two distinct real eigenvalues $\lambda^+ = \lambda > 1$ and $\lambda^- = 1/\lambda$. The lines parallel to the eigenspaces for A foliate the plane in two different ways. Call these foliations $\widetilde{\mathcal{F}^\pm}$. These descend to the torus to give transverse foliations $\mathcal{F}^\pm$. We may further endow each of these foliations with *transverse measures* $\mu^\pm$ as follows: for any arc α in the torus choose a lift $\widetilde{\alpha}$. Let $\mu^+(\widetilde{\alpha})$ be the distance between the two lines of $\widetilde{\mathcal{F}^+}$ which meet the endpoints of $\widetilde{\alpha}$. We could use the two transverse measures to determine a new Euclidean metric on S_1, with respect to which the Anosov element acts in a fairly standard way – it preserves the foliations $\mathcal{F}^\pm$, stretching $\mathcal{F}^+$ (shrinking $\mathcal{F}^-$) in the tangential direction. That is, A preserves the triple $(S_1, \mathcal{F}^\pm)$ and rescales the transverse measures $\mu^\pm$ by a factor of $\lambda^\pm$.

EXERCISE 1.9. Verify the claims of the above paragraph. Give an accurate picture of $\mathcal{F}^\pm$ for the map induced by $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, as shown in Figure 5.

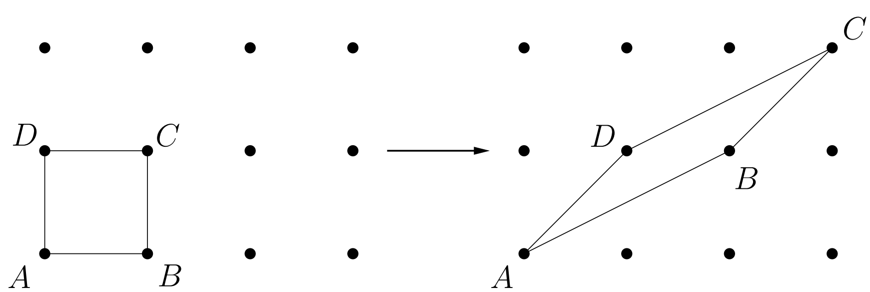

FIGURE 5. The action of A on the plane. Check that this descends to act on the torus.

Let us leave these examples and return to the more general theme. A *mapping class* of S is a proper isotopy class of homeomorphisms of S. Note that the mapping classes form a group, the *mapping class group* $\mathcal{MCG}(S)$.

EXERCISE 1.10. Show that the composition of mapping classes is well defined.

Following Thurston we call a mapping class f

- *periodic*: if f is finite order,
- *reducible*: if f permutes a collection of essential non-peripheral curves (up to isotopy),
- *pseudo-Anosov*: if there is a $f' \in f$ preserving a pair of transverse singular measured foliations, while rescaling the measures in the correct fashion.

Instead of giving a precise definition of a singular foliation, here is a collection of examples: an Anosov map of the torus S_1 gives a foliation and, lifting by a branched cover $S_g \to S_1$, gives a singular foliation in S_g. If the branched cover is chosen carefully the Anosov map will lift to a pseudo-Anosov map on S_g.

Another characterization of pseudo-Anosov maps is the following: $f \colon S \to S$ is pseudo-Anosov if for every essential non-peripheral curve α the geometric intersection number $\iota(\alpha, f^n(\alpha))$ grows without bound. As a hint of the proof: any such curve, under iteration by f, comes closer and closer to being "parallel" to the stable foliation. See Casson and Bleiler [**8**] or Fathi, Laudenbach, and Poénaru [**12**] for a detailed discussion.

EXERCISE 1.11. Convince yourself that it is somewhat ok to ignore the difference between a homeomorphism and its mapping class by proving: if f is reducible then there is a $f' \in f$ which permutes the collection of essential non-peripheral curves. (For the much more ambitious reader: prove that if f is periodic then there is a $f' \in f$ which is finite order.)

We have already shown that the reducible elements of $\mathrm{SL}(2, \mathbb{Z})$ are also reducible in this new sense. Here is a much more general set of examples: fix attention on a properly embedded curve α in an oriented surface S and let $N = N(\alpha) \cong S^1 \times [0, 1]$ be a closed annular neighborhood. (The homeomorphism is chosen so that the product orientation agrees with the given orientation on S.) Define the *positive Dehn twist* $\tau_\alpha \colon S \to S$ by setting $\tau_\alpha | S \smallsetminus N = \mathrm{Id}$ and setting $\tau_\alpha(\theta, r) = (\theta + 2\pi r, r)$.

EXERCISE 1.12. If α bounds a disk, then τ_α is isotopic to the identity.

EXERCISE 1.13. The surface S need not be orientable to define a Dehn twist. It suffices that the neighborhood $N = N(\alpha)$ be homeomorphic to an annulus. (We cannot twist about the core of a Möbius

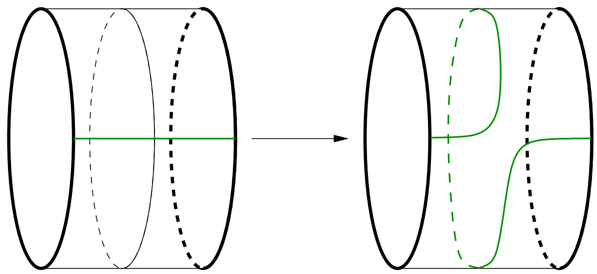

FIGURE 6. Twisting along the vertical curve transforms the horizontal one. As the twist is positive, the horizontal curve “turns left.”

band.) Such curves are called *two-sided*. If $\alpha \subset S$ bounds a Möbius band then α is certainly two-sided. Prove that if α bounds a Möbius band then τ_α is isotopic to the identity.

We now obtain simple reducible maps by performing Dehn twists on a collection of disjoint curves.

EXERCISE 1.14. Suppose that α and β are curves in S with $\iota(\alpha, \beta) = 1$. Show that $\tau_\alpha\tau_\beta\tau_\alpha = \tau_\beta\tau_\alpha\tau_\beta$ as mapping classes – this is called the *braid relation*. Once you've done this, it should be easy to show that $(\tau_\alpha\tau_\beta)^6 = \tau_\gamma$ as mapping classes, where γ is the boundary of a closed regular neighborhood of $\alpha \cup \beta$.

EXERCISE 1.15. Prove that $\mathcal{MCG}(S_1) \cong \mathrm{SL}(2, \mathbb{Z})$.

EXERCISE 1.16. Note that it is possible for a periodic mapping to be reducible. Find one in $\mathcal{MCG}(S_2)$ which is not.

Generalizing the notion of a Dehn twist is what we will call a *partial map*: chose an essential subsurface $X \subset S$ and a mapping class $f : X \to X$. Choose a representative $f' \in f$ so that $f'|\partial X = \mathrm{Id}$ and extend f' by the identity map on $S \setminus X$. Then the extension is a *partial map*. We say X contains the *support* of the partial map.

REMARK 1.17. More formally, for any pair of disjoint essential subsurfaces $X, Y \subset S$ there is a fairly natural map $\mathcal{MCG}(X) \times \mathcal{MCG}(Y) \to \mathcal{MCG}(S)$. (We are ignoring issues regarding Dehn twists about curves parallel to boundary components of X and Y.) Other relations between mapping class groups may be found in Birman's book [**3**].

We finish this section by stating a major milestone in both the study of surface dynamics and of three-manifolds:

THEOREM 1.18 (Thurston [**40**]). *Every mapping class is either periodic, reducible, or pseudo-Anosov.*

4. The complex of curves

We now come to the main object of interest for these notes. The *complex of curves* was first defined by Harvey [**17**], who was studying the Teichmüller spaces of Riemann surfaces. More recently, Minsky began to investigate three-manifolds via the geometric structure of the curve complex. I will follow Masur and Minsky's treatment in [**30**] and [**31**].

Fix attention on a non-sporadic surface, S. Define a simplicial complex $\mathcal{C}(S)$ as follows: for vertices we take isotopy classes of essential non-peripheral curves in S. A collection of $k+1$ vertices $\{\alpha_i\}_0^k$ form a k-simplex whenever the α_i can be realized by disjoint curves in S.

EXERCISE 1.19. Show that top dimensional simplices in $\mathcal{C}(S)$ have $\zeta(S)$ many vertices. (This is one explanation of the definition of $\zeta(S)$.)

EXERCISE 1.20. Show that $\mathcal{C}(S_{0,5})$ and $\mathcal{C}(S_{1,2})$ are isomorphic as simplicial complexes. Do the same for $\mathcal{C}(S_{0,6})$ and $\mathcal{C}(S_{2,0})$. Show that $\mathcal{C}(S_{0,6})$ and $\mathcal{C}(S_{1,3})$ are *not* isomorphic.

To simplify our discussion, and to continue to follow [**30**] closely, we generally restrict attention to the one-skeleton of $\mathcal{C}(S)$.

Let $d_S(\alpha, \beta)$ denote the minimal number of edges in any edge path in $\mathcal{C}^1(S)$ starting at α and ending at β. For this to be well-defined we must show that $\mathcal{C}^1(S)$ is connected. In fact we can show quite a bit more:

LEMMA 1.21. *Fix S a compact, connected surface which is not simple. Suppose that α and β are curves with $\iota(\alpha, \beta) \neq 0$. Then*

$$d_S(\alpha, \beta) \leq 2\log_2(\iota(\alpha, \beta)) + 2.$$

This form of the inequality may be found as Lemma 2.1 in Hempel's paper [**20**].

REMARK 1.22. The proof relies on the idea of *curve surgery* – given α and β we can form many other curves by taking the union of various arcs of $\alpha \smallsetminus \beta$ with arcs of $\beta \smallsetminus \alpha$.

REMARK 1.23. The lemma implies that $\mathcal{C}(S)$ is connected. The connectedness of $\mathcal{C}(S)$ was naturally first observed by Harvey [**17**] (see his Proposition 2). Harvey cites Lickorish [**28**] for the inductive argument, which is an essential step in the proof that the mapping class group is finitely generated. Our proof of Lemma 1.21 follows Lickorish [**27**] – even the diagrams are the same.

It can be argued that the key ideas to prove connectedness first appeared in Dehn's work. However his surgery arguments appear to be much more complicated. See Stillwell's notes in [**11**].

PROOF OF LEMMA 1.21. Suppose that S is not simple. We also suppose that S is orientable and not sporadic, leaving these cases as Exercises 1.25 and 1.31.

EXERCISE 1.24. Show that if $\iota(\alpha, \beta)$ equals one or two then $d_S(\alpha, \beta) = 2$.

This exercise deals with the base case of an induction on $\iota(\alpha, \beta)$. So suppose now that $\iota(\alpha, \beta) = n > 2$. Isotope β to realize this intersection number. Orient α.

We have two cases. Suppose first that there are two intersections points $x, y \in \alpha \cap \beta$, consecutive along β, with the following property: Let γ be the subarc of β with $\partial\gamma = \{x, y\} = \gamma \cap \alpha$. Then the tangents to α at x and y agree, up to parallel translation along γ. See the left side of Figure 7.

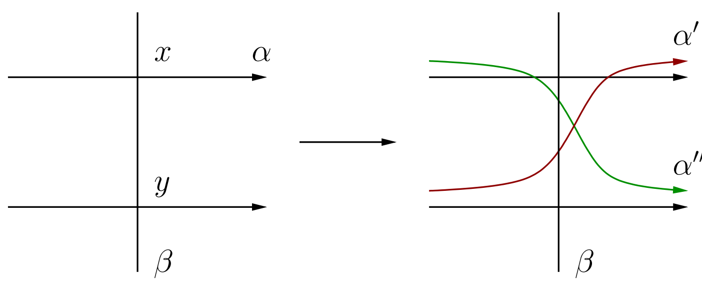

FIGURE 7. A neighborhood of the union $\alpha \cup \gamma$ is an essential once-holed torus. The curves $\alpha, \alpha', \alpha''$ span a Farey triangle in this subsurface.

Let δ' and δ'' be subarcs of α so that $\delta' \cap \delta'' = \{x, y\}$ and $\delta' \cup \delta'' = \alpha$. Form the simple closed curves α' and α'' by isotoping $\gamma \cup \delta'$ and $\gamma \cup \delta''$ slightly, into general position. See the right side of Figure 7.

Thus $\iota(\alpha', \beta) + \iota(\alpha'', \beta) \le \iota(\alpha, \beta)$. Since α' and α'' intersect once transversely by Exercise 1.2 the curves are essential and non-peripheral. One of them has intersection number with β at most half as large as α does and the induction is complete.

We now turn to the second case. Suppose now that there are three consecutive intersection points x, y, and z with the following property:

Let γ be the subarc of β containing y with endpoints $\partial\gamma = \{x, z\}$. Then the tangents to α at x and z agree, and that at y disagrees, after parallel translation along γ. See the left side of Figure 8.

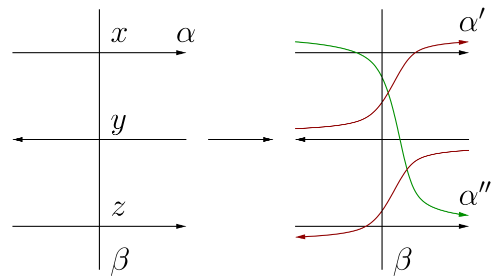

FIGURE 8. A neighborhood of the union $\alpha \cup \gamma$ is an essential four-holed sphere. The curves $\alpha, \alpha', \alpha''$ span a Farey triangle in this subsurface.

Up to relabeling, symmetry, and reorienting α we may suppose that the three arcs of $\alpha \smallsetminus \{x, y, z\}$ have endpoints $\{x, y\}$, $\{y, z\}$, $\{z, x\}$ and induced orientations pointing away from the first point in each case. Surger α as shown in the right side of Figure 8 to obtain α' and α''.

We find that, as above, $\iota(\alpha', \beta) + \iota(\alpha'', \beta) \leq \iota(\alpha, \beta)$. Now, the intersection number of any two of $\alpha, \alpha', \alpha''$ is exactly two. If the geometric intersection number of any two of them is zero (it cannot be one) then there is a bigon in the picture. This bigon can be used to reduce the intersection number of α and β, an impossibility. Thus $\iota(\alpha, \alpha') = \iota(\alpha', \alpha'') = \iota(\alpha'', \alpha) = 2$ and again, by Exercise 1.2, the curves α' and α'' are essential and non-peripheral. One of these has intersection number with β at most half as large as α does and the induction is again complete. This completes the proof of the lemma. $\square$

EXERCISE 1.25. Generalize to the case where S is non-orientable.

Note that Lemma 1.21 is *not* sharp: there are curves α and β with $d_S(\alpha, \beta) = 2$ but with $\iota(\alpha, \beta)$ as large as desired.

EXERCISE 1.26. Find such pairs.

However, it is also true that Lemma 1.21 cannot be improved in an obvious way. For example, if α is any essential non-peripheral curve and $f \colon S \to S$ is any pseudo-Anosov mapping then $d_S(\alpha, f^n(\alpha))$ grows linearly with n while $\iota(\alpha, f^n(\alpha))$ grows exponentially. As a reference for the first fact, see Theorem 2.24 below. As for the second, see Casson and Bleiler [**8**].

Here is another standard interpretation of curves at high distance in $\mathcal{C}(S)$. Suppose that $d_S(\alpha, \beta) \geq 3$. Then the curves α and β *fill* the surface S: any other curve γ meets either α or β. It follows that $S \smallsetminus (\alpha \cup \beta)$ is a collection of disks and annuli, with exactly one annulus containing each component of ∂S.

Exercise 1.27. Give explicit examples of a pair of curves at distance exactly 3 in $\mathcal{C}(S_2)$. Can you find an explicit pair of curves with distance exactly 4?

Another result which we should mention, on the global topology of $\mathcal{C}(S)$, is Harer's theorem: the curve complex is homotopy equivalent to a wedge of infinitely many spheres, all of the same dimension. See Ivanov's article (Theorem 3.3.A of [**23**]) for a more complete statement. The relevant paper of Harer's is [**16**].

Exercise 1.28. Find spheres in $\mathcal{C}(S_{0,5})$ and in $\mathcal{C}(S_2)$. Can you prove that these are not contractible?

5. A trip to the zoo: the Farey graph

Here we fill the gap left by the fact that $\mathcal{C}(S)$ has no edges when S is a sporadic, non-simple surface. That is, when S is either S_1, $S_{1,1}$, or $S_{0,4}$.

The *Farey graph* $\mathcal{F}$ has vertex set the set of isotopy classes of essential non-peripheral curves in S_1. A collection of $k+1$ vertices span a k-simplex if all of the curves pairwise meet exactly once.

Exercise 1.29. Prove that the top dimensional simplex in $\mathcal{F}$ is a triangle.

Note that the Farey graph is also the "curve complex" for the surfaces $S_{1,1}$ and $S_{0,4}$. For the former no change in the definition is necessary. For the latter we take vertices to be essential non-peripheral curves in $S_{0,4}$ and take an edge between any pair that meets exactly twice.

Exercise 1.30. Prove this. The fact that $S_{0,4}$ and $S_{1,1}$ both cover the orbifold $S^2(2, 2, 2, \infty)$ may be useful.

We denote distance in the Farey graph by $d_{\mathcal{F}}(\cdot, \cdot)$.

Exercise 1.31. Prove that $d_{\mathcal{F}}(\alpha, \beta) \leq \log_2(\iota(\alpha, \beta)) + 1$ for curves in S_1 or $S_{1,1}$ while $d_{\mathcal{F}}(\alpha, \beta) \leq \log_2(\iota(\alpha, \beta))$ for curves in $S_{0,4}$.

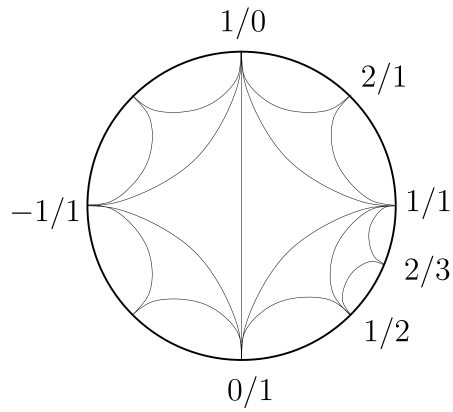

FIGURE 9. A terrible picture of a small part of the Farey graph. Some of the vertices are labeled via the corresponding elements of $\widehat{\mathbb{Q}}$.

6. A trip to the zoo: the annulus and the pants

When $S \cong S_{0,2}$ the standard definition of $\mathcal{C}(S)$ has no edges *or* vertices. It is important to fill this gap — we will need the "curve" complex of the annulus to help keep track of Dehn twists. We note that we do *not* need to invent a curve complex for the pants $S_{0,3}$: any mapping class on $S_{0,3}$ is a product of twists on the boundary, and these are recorded by the curve complex of the corresponding annuli.

Fix $\mathbb{A} = S^1 \times I$. Let $\mathcal{C}(\mathbb{A})$ be the complex where vertices are isotopy classes of spanning arcs, via isotopies fixing the boundary pointwise. Two vertices are connected by an edge if the isotopy classes have disjoint representatives.

EXERCISE 1.32. Check that $\mathcal{C}(\mathbb{A})$ is quasi-isometric to $\mathbb{R}$. See Section 3 of Chapter 2 for definitions.

CHAPTER 2

Coarse geometry

1. Basic definitions

In this part of the notes we will recall a number of definitions from metric and coarse metric geometry. A wonderful but difficult introductory paper in the topic is Gromov's article on hyperbolic groups [**14**]. A more readable introduction is the book by Coornaert, Delzant, and Papadopoulos [**10**].

Let $(\mathcal{X}, d_{\mathcal{X}})$ be a metric space. Recall that a *path* in $\mathcal{X}$ from x to y is a continuous map $P\colon [a,b] \to \mathcal{X}$ from the interval $[a,b] \subset \mathbb{R}$ so that $P(a) = x$ and $P(b) = y$. Also, P is a *geodesic* if $d_{\mathcal{X}}(P(a'), P(b')) = |b' - a'|$ for all $a', b' \in [a,b]$. We say $\mathcal{X}$ is a *geodesic metric space* if every pair of points of $\mathcal{X}$ is connected by a geodesic. We denote a geodesic connecting x to y by $[x,y]$ when the parameterization and image may safely be forgotten.

We now have an important definition:

Definition 2.1. The metric space $\mathcal{X}$ is *Gromov hyperbolic* (or *δ-hyperbolic* or simply *hyperbolic*) with constant $\delta \geq 0$ if, for every geodesic triangle xyz the closed δ neighborhood of the two sides $[x,y]$ and $[y,z]$ contains the third side $[z,x]$.

Exercise 2.2. Prove that there are points $a \in [x,y], b \in [y,z], c \in [z,x]$ so that $d_{\mathcal{X}}(a,b)$ and $d_{\mathcal{X}}(a,c)$ are at most δ. (Note the asymmetry.)

Loosely speaking, Gromov hyperbolicity says that geodesic triangles in the space $\mathcal{X}$ are "slim." This (plus great deal of work) has many applications to the study of infinite groups. For example, if the Cayley graph of a finitely presented group is hyperbolic then Dehn's *word problem* is solvable in that group.

Exercise 2.3. Show that a tree (for example, a graph without loops) is hyperbolic with constant $\delta = 0$. Show that any metric space with bounded diameter is hyperbolic for *some* constant δ. Show that $\mathbb{R}^2$ is not hyperbolic for any constant. It follows that $\mathbb{R}^n$ and hence $\mathbb{Z}^n$ (for $n > 1$) are not Gromov hyperbolic.

We may now state:

THEOREM 2.4 (Masur and Minsky [30]). *The curve complex $C(S)$ is hyperbolic.*

We cannot even sketch the proof here. Instead we refer the reader to their original paper, or to Bowditch's version [4]. As just a hint we note that their proof finds, for any pair of vertices α and β, a "good" path between them. They then go on to show that closest point projection of $\mathcal{C}(S)$ to one of these paths greatly contracts distance. This in turn implies hyperbolicity.

2. Boundaries

Here we recall the definition of the *Gromov boundary* of a hyperbolic metric space. Fix $(\mathcal{X}, d_{\mathcal{X}})$ a δ-hyperbolic metric space. Pick $\omega \in \mathcal{X}$ a basepoint. Define the *Gromov product* of x and y in $\mathcal{X}$ to be:

$$(x \cdot y) = (x \cdot y)_\omega = \frac{1}{2}(d_{\mathcal{X}}(x, \omega) + d_{\mathcal{X}}(y, \omega) - d_{\mathcal{X}}(x, y)).$$

Roughly speaking, this is the distance between the basepoint ω and a geodesic connecting x to y. Put another way, fix geodesics P and Q connecting ω to x and y. The Gromov product measures how long P and Q fellow-travel.

EXERCISE 2.5. Justify the first remark by proving, if $x, y, \omega \in \mathcal{X}$ and $R = [x, y]$ is a geodesic from x to y, then $d_{\mathcal{X}}(\omega, R) - 4\delta \leq (x \cdot y) \leq d_{\mathcal{X}}(\omega, R)$.

EXERCISE 2.6. Here is a sort of triangle inequality: For all $x, y, z, \omega \in \mathcal{X}$ we have:

$$(x \cdot y) \geq \min\{(x \cdot z), (z \cdot y)\} - 5\delta.$$

(In fact, this condition is equivalent to hyperbolicity. For the details see [6], page 411.) It is trivial to deduce:

$$(x \cdot y) \geq \min\{(x \cdot z), (z \cdot w), (w \cdot y)\} - 10\delta.$$

We say that a sequence $\{x_i\}_{i \in \mathbb{N}}$ *converges at infinity* if $\lim_{i,j \to \infty}(x_i \cdot x_j) = \infty$. Sequences $\{x_i\}$ and $\{y_i\}$ converging at infinity are *equivalent* if $\lim_{i,j \to \infty}(x_i \cdot y_j) = \infty$. Finally, define $\partial_\infty \mathcal{X}$, the *Gromov boundary* of $\mathcal{X}$ to be the set of these equivalence classes.

EXERCISE 2.7. Check that these notions are independent of the choice of basepoint ω. Check that equivalence *is* an equivalence relation.

EXERCISE 2.8. Prove that if $\mathcal{X}$ has bounded diameter then $\partial_\infty \mathcal{X}$ is empty.

It is also possible to give a metric on $\partial_\infty \mathcal{X}$, but the construction is delicate and is also basepoint-dependent. See the book of Ghys and de la Harpe [**13**] or that of Bridson and Haefliger ([**6**], pages 429-437). We will content ourselves with giving a topology on $\partial_\infty \mathcal{X}$.

Fix points X and Y in $\partial_\infty \mathcal{X}$. Extend the Gromov product to $\partial_\infty \mathcal{X}$ by taking:

$$(X \cdot Y) = \inf\{\liminf_{i,j\to\infty} (x_i \cdot y_j) \mid \{x_i\} \in X, \{y_j\} \in Y\}.$$

We say that a sequence $\{X_n\} \subset \partial\mathcal{X}$ *converges* to X if $\lim_{n\to\infty}(X_n \cdot X) = \infty$.

Exercise 2.9. Check that convergence is independent of our choice of basepoint, ω.

In using $\inf\{\liminf\}$ in the definition we are following several authors: for example [**10**] and [**1**]. Others, such as [**6**] and [**13**], use $\sup\{\liminf\}$ instead.

Exercise 2.10. Does it make any difference? What if we used $\inf\{\limsup\}$ or $\sup\{\limsup\}$ instead? Prove, for example, if $\{x_i\}$ and $\{y_j\}$ converge to distinct points on the boundary then there is an $N \in \mathbb{N}$ so that the set $\{(x_i \cdot y_j) \mid i, j \geq N\}$ has diameter less than 10δ. (Hint: use the "rectangle inequality" of Exercise 2.6.)

The next exercise is taken from Remark 3.17 (page 432) of [**6**]:

Exercise 2.11. The extended Gromov product is similar to a metric:

- $(X \cdot Y) = \infty$ if and only if $X = Y$.
- $(X \cdot Y) = (Y \cdot X)$.
- $(X \cdot Y) \geq \min\{(X \cdot Z), (Z \cdot Y)\} - 10\delta$.

Here is a useful bit of analysis:

Exercise 2.12. Prove the *Cauchy criterion*: $\{X_n\} \subset \partial_\infty \mathcal{X}$ converges if and only if for all $M \in \mathbb{N}$ there is an $N \in \mathbb{N}$ where $m, n \geq N$ implies that $(X_m \cdot X_n) \geq M$.

Before attempting the curve complex, perhaps it would be wise to consider a few examples:

Exercise 2.13. Show that the boundary of the three-valent tree, $\partial_\infty T_3$, is a Cantor set and so is compact. Prove that $\partial_\infty T_\infty$ is not compact. (This last can be done directly or by using the Cauchy criterion.)

Exercise 2.14. Can you prove that $\partial_\infty \mathcal{C}(S)$ is not compact?

We may now state an important theorem of Klarreich [**25**]:

THEOREM 2.15. *The boundary at infinity of $\mathcal{C}(S)$ is homeomorphic to $\mathcal{EF}(S)$, the space of ending foliations.*

We only sketch the definition of $\mathcal{EF}(S)$ – we refer the reader to Kapovich's book [**24**] for details about foliations. Let $\mathcal{PMF}(S)$ be the space of *projectively measured foliations* on S. Let Δ be the set of all foliations containing a closed leaf and let $\bar{\Delta}$ be the closure. Then $\mathcal{EF}(S)$ is the quotient of $\mathcal{PMF}(S) \smallsetminus \bar{\Delta}$ which forgets the measures, remembering only the topological equivalence class.

As a companion to Theorem 2.15 there is an outstanding question of Peter Storm:

QUESTION 2.16. Is the boundary of the curve complex, $\partial_\infty \mathcal{C}(S)$, connected?

Of course, for $\mathcal{F} = \mathcal{C}(S_{1,1}) = \mathcal{C}(S_{0,4})$ the answer is "no". However, as the complexity $\zeta(S)$ grows perhaps $\partial_\infty \mathcal{C}(S)$ becomes highly connected... See Chapter 5 for a related discussion.

3. Quasi-isometric embeddings

We now turn to another branch of Gromov's program of "coarse geometry."

A function $f\colon \mathcal{Y} \to \mathcal{X}$ is a (K, E) *quasi-isometric embedding* for $K \geq 1, E \geq 0$ if, for every $x, y \in \mathcal{Y}$, we have

$$\frac{1}{K}(d_{\mathcal{Y}}(x, y) - E) \leq d_{\mathcal{X}}(x', y') \leq K \cdot d_{\mathcal{Y}}(x, y) + E$$

where $x' = f(x)$ and $y' = f(y)$. If, in addition, f is *E-dense* (an E neighborhood of $f(\mathcal{Y})$ equals all of $\mathcal{X}$) then we say that f is a *quasi-isometry* and that $\mathcal{X}$ is *quasi-isometric* to $\mathcal{Y}$. We note that a quasi-isometry need not be continuous.

(As a bit of notation, if $r, s \in \mathbb{R}$ and $\frac{1}{K}(r - E) \leq s \leq K \cdot r + E$ then we write $r \overset{K,E}{=} s$. We also call r and s *quasi-equal* with constants (K, E) if this occurs.)

By now you have the hang of things: coarsen your favorite metric concept by replacing exact measurements by measurements with bounded multiplicative and additive error (and by sticking a "quasi" in front of the word).

Given a space $\mathcal{X}$ we can ask:

- Which more familiar space is $\mathcal{X}$ quasi-isometric to?
- What is the group of self quasi-isometries of $\mathcal{X}$ (up to the equivalence relation $f \sim g$ if $g^{-1}f$ moves all points at most a bounded amount)?

- What other invariants of $\mathcal{X}$ can we compute? (Such as the Gromov boundary, size and growth of metric balls, metrically interesting subspaces...)

EXERCISE 2.17. Show that $\mathbb{Z}$ with the standard metric is quasi-isometric to $\mathbb{R}$. Recall that T_k is regular k-valent tree. So T_2 is isometric to $\mathbb{R}$. Show that T_3 is quasi-isometric to T_4. (In fact T_3 is quasi-isometric to T_k for any $k > 2$.) However, T_3 is not quasi-isometric T_∞, the regular tree with countably infinite valence. (Some care must be taken here: the natural CW structure on T_∞ is not metrizable.)

EXERCISE 2.18. Prove that T_∞ is quasi-isometric to $\mathcal{F}$, the Farey graph.

EXERCISE 2.19. Show that $\mathbb{R}^2$ with the metric $d_1(x,y) = |x_1 - y_1| + |x_2 - y_2|$ is quasi-isometric to $\mathbb{R}^2$ with the metric $d_2(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}$. In fact the additive error E may be taken to be zero and so the two metrics are *bi-Lipschitz*.

EXERCISE 2.20. Show that Gromov hyperbolicity is a *quasi-isometry invariant*: if $f\colon \mathcal{Y} \to \mathcal{X}$ is a quasi-isometry, and $\mathcal{Y}$ is δ hyperbolic, then $\mathcal{X}$ is δ' hyperbolic. Nonetheless, one can show that T_3 is not quasi-isometric to $\mathbb{H}^n$, the hyperbolic space of dimension $n > 0$. A model for $\mathbb{H}^n$ is the upper-half space $\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid x_n > 0\}$ with metric $\frac{ds}{x_n}$ where ds is the standard L^2 line element for $\mathbb{R}^n$. (Hint: the complement of a large ball in $\mathbb{H}^n$ is connected.) Find instead an explicit quasi-isometric embedding $f\colon T_3 \to \mathbb{H}^2$.

EXERCISE 2.21. Show that if $n \neq m$ then $\mathbb{R}^n$ is not quasi-isometric to $\mathbb{R}^m$. (Hint: show that $\mathbb{Z}^n$ is not quasi-isometric to $\mathbb{Z}^m$ by counting the number of points in a ball of radius R.)

We define a (K,E) *quasi-geodesic* L in X to be a (K,E) quasi-isometric embedding of a closed connected subset of $\mathbb{R}$.

EXERCISE 2.22. Given (K,E) classify all possible images of (K,E) quasi-geodesics in T_3, the three-valent tree.

4. Action of the mapping class group

Suppose now that $f\colon \mathcal{X} \to \mathcal{X}$ is an *isometry*: f is onto and, for any $x, y \in \mathcal{X}$, we have $d_{\mathcal{X}}(x,y) = d_{\mathcal{X}}(f(x), f(y))$. An orbit $\mathcal{O}(x)$ of f is the set $\{f^n(x) \mid n \in \mathbb{Z}\}$.

We say f is:

- *elliptic* if every orbit of f is bounded,

- *hyperbolic* if every orbit of f is a quasi-isometric embedding of $\mathbb{Z}$ (and thus is a quasi-geodesic), or
- *parabolic* if f is neither elliptic nor hyperbolic.

Note that quasi-isometries fall into the same classification.

EXERCISE 2.23. Show that any isometry of T_3 is elliptic or hyperbolic. This is not the case for $\mathbb{H}^2$.

We now have:

THEOREM 2.24 (Masur and Minsky [**30**]). *Periodic and reducible elements of $\mathcal{MCG}(S)$ act elliptically on $\mathcal{C}(S)$. Pseudo-Anosov elements act hyperbolically. (Also, Dehn twists act hyperbolically on $\mathcal{C}(S_{0,2})$.)*

As a consequence we have:

COROLLARY 2.25. *The curve complex $\mathcal{C}(S)$ has infinite diameter.*

REMARK 2.26. Masur and Minsky [**30**] also present a more direct proof of Corollary 2.25 which is due to Luo, adapting an argument of Kobayashi [**26**]. Here is a brief sketch – we refer the reader to Kapovich's book [**24**] for details about foliations: Choose $\mathcal{F}$ a uniquely ergodic minimal foliation and let α_i be a sequence of curves which converge to $\mathcal{F}$ as projectively measured foliations. If all of these curves are distance at most M from some curve β we can choose sequences $\{\beta_i^j\}_{j=0}^M$ so that

- $\beta_i^0 = \beta$,
- $\beta_i^M = \alpha_i$, and
- $\beta_i^j \cap \beta_i^{j+1} = \emptyset$.

We now take subsequences with respect to the j index to find a collection $\{\beta_i^j\}_{j=0}^M$ with all of the above properties and in addition the sequence $\{\beta_i^j\}_{i=0}^\infty$ converges to a projective measured foliation $\mathcal{F}_i$. Using the fact that $\mathcal{F}$ is minimal and uniquely ergodic we can deduce that $\mathcal{F} = \mathcal{F}_{M-1} = \mathcal{F}_{M-2} = \ldots = \mathcal{F}_0 = \beta$, a contradiction. $\square$

5. Subsurface projections

We now turn to one of the most important definitions used by Masur and Minsky in their study of the curve complex. There are two definitions commonly given. The first is more concrete, but the second generalizes correctly to annuli. We will give both.

5.1. Cutting up curves. Suppose that X is an essential non-simple subsurface of S. Define the *subsurface projection* $\pi_X \colon \mathcal{C}(S) \to \mathcal{C}(X)$, as follows: fix attention on $\alpha \in \mathcal{C}(S)$ and isotope α to minimize $|\alpha \cap X|$. Now,

- if $\alpha \subset X$ then set $\pi_X(\alpha) = \alpha$,
- if $\iota(\alpha, \partial X) > 0$ then pick any arc $\alpha' \subset \alpha \cap X$, set N equal to a closed neighborhood of $\alpha' \cup \partial X$, and set $\pi_X(\alpha)$ equal to any component α'' of ∂N which is essential and non-peripheral in X, or
- if $\alpha \subset S \setminus X$ set $\pi_X(\alpha) = \emptyset$.

In the first two cases we say that α *cuts* X. In the last case we say α *misses* X. So, properly speaking, π_X is defined only on those vertices of $\mathcal{C}(S)$ which cut X. We extend π_X to a disjoint collection of curves in the obvious way.

5.2. Lifting curves. We now turn to the second definition. Fix a hyperbolic metric on S. Suppose that X is an essential non-pants subsurface of S. Redefine the subsurface projection $\pi_X : \mathcal{C}(S) \to \mathcal{C}(X)$, as follows: fix attention on a curve α. Straighten α to be a geodesic. Let $\widetilde{\alpha}$ be the collection of lifts of α to $\widetilde{X}$, the cover of S corresponding to the subgroup $\pi_1(X)$. Compactify $\widetilde{X}$ by adding its Gromov boundary and take the closure of $\widetilde{\alpha}$ in the resulting surface, which we identify with X.

- if $\widetilde{\alpha}$ is a single curve, not peripheral in X, then take $\pi_X(\alpha) = \alpha$, or
- if X is not an annulus and if $\widetilde{\alpha}$ contains arcs essential in X then pick any one of them, say $\alpha' \subset \widetilde{\alpha}$, set N equal to a closed neighborhood of $\alpha' \cup \partial X$, and set $\pi_X(\alpha)$ equal to any component α'' of ∂N which is essential and non-peripheral in X, or
- if X is an annulus and if $\widetilde{\alpha}$ contains spanning arcs essential in X then pick any one of them, say $\alpha' \subset \widetilde{\alpha}$, to be $\pi_X(\alpha)$, or
- if all components of $\widetilde{\alpha}$ are parallel into the boundary set $\pi_X(\alpha) = \emptyset$.

This definition is a bit more technical, but is required to deal with the annulus case.

For either definition we could take π_X to be set-valued to avoid losing information. As it turns out, such care will not be necessary. Also, as a bit of notation we write $d_X(\alpha, \beta)$ for $d_X(\pi_X(\alpha), \pi_X(\beta))$ when $\pi_X(\alpha)$ and $\pi_X(\beta))$ are not the empty set.

5.3. Consequences.

LEMMA 2.27. *Suppose that X is an essential non-pants subsurface of S, α and β both cut X, and $d_S(\alpha, \beta) = 1$. Then $d_X(\alpha, \beta) \leq 6$.*

PROOF. This follows from the fact that $\iota(\pi_X(\alpha), \pi_X(\beta)) \leq 4$ and from Lemma 1.21. (When X is sporadic we will need the version of

Lemma 1.21 provided by Exercise 1.31. When X is an annulus the bound improves from 6 to 1.) □

LEMMA 2.28. *Suppose that* $g = \{\alpha_0, \alpha_1, \ldots, \alpha_n\}$ *is a path in* $\mathcal{C}(S)$. *Suppose that every* α_i *cuts* X, *a non-pants essential surface in* S. *Then* $d_X(\alpha_0, \alpha_n) \leq 6n$. □

This has a useful generalization:

COROLLARY 2.29. *For every* $a \in \mathbb{N}$ *there is a number* $b \in \mathbb{N}$ *with the following property: Suppose that* $g = (\alpha_0, \alpha_1, \ldots, \alpha_n)$ *is a sequence of vertices in* $\mathcal{C}(S)$ *each of which cuts* X *and where* $\iota(\alpha_i, \alpha_{i+1}) \leq a$. *Then* $d_X(\alpha_0, \alpha_n) \leq b \cdot n$.

6. An example of curves at distance four

We are now equipped to give a fairly explicit example of a pair of curves at distance four.

Let S be the five-holed sphere $S_{0,5}$. Choose disjoint arcs $\delta_0, \delta_1, \delta_2$ in S connecting pairs of distinct boundary components so that δ_0 and δ_2 share exactly one boundary component, and δ_1 shares none. Let α_i be the essential non-peripheral boundary component of a closed regular neighborhood of $\delta_i \cup \partial S$. See Figure 1. Note that $\{\alpha_0, \alpha_1, \alpha_2\}$ is a geodesic of length two in $\mathcal{C}(S)$.

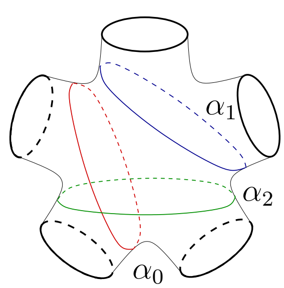

FIGURE 1. The curve α_2 separates the bottom two boundary components from the top three.

Note that α_2 cuts S into P, a pair of pants, and X_2, a four-holed sphere. Choose f a partial pseudo-Anosov supported in X_2 and raise it to a large enough power, say f^n, so that $d_X(\alpha_0, f^n(\alpha_0)) \geq 25$. This is possible by Theorem 2.24. (As $\mathcal{C}(X_2) = \mathcal{F}$ is the Farey graph the map f and the power n can be made explicit.) Let $\alpha_3 = f^n(\alpha_1)$ and let $\alpha_4 = f^n(\alpha_0)$.

CLAIM. The curves α_0 and α_4 have distance exactly four in $\mathcal{C}(S)$.

Note that $d_S(\alpha_0, \alpha_4) \leq 4$ as the α_i gives a path of length exactly four. It is also clear that any path from α_0 to α_4 passing through α_2 has length at least four. Suppose then that we have a path $g = \{\beta_j\}_0^m$ in $\mathcal{C}(S)$ from $\alpha_0 = \beta_0$ to $\alpha_4 = \beta_m$ which does not pass through α_2 (no β_j equals α_2). It follows that every β_j cuts the four-holed sphere X_2. We apply Lemma 2.28 to find that $25 \leq d_X(\alpha_0, \alpha_4) = d_X(\beta_0, \beta_m) =\leq 6m$. Thus $m > 4$ and we have proved that $d_S(\alpha_0, \alpha_4) = 4$. □

Exercise 2.30. Make the vague parts of this construction concrete. You will need to investigate how $\mathcal{MCG}(S_{0,4})$ acts on the Farey graph $\mathcal{F}$. Can you find the smallest possible intersection number between two curves in $S_{0,5}$ which are at distance four in $\mathcal{C}(S_{0,5})$? Lemma 1.21 does not give a very good lower bound... (Perhaps you need at least sixteen intersection points?)

Remark 2.31. The construction above underlines the similarity between the action of a partial map on $\mathcal{C}(S)$ and the action of a rotation on $\mathbb{R}^2$. The partial map f "rotates" all of $\mathcal{C}(S)$ about the non-peripheral boundary of X_2, the support of f.

Remark 2.32. It is also interesting to note the combinatorial nature of $\mathcal{C}(S)$. Despite the fact that $\mathcal{C}(S)$ is locally infinite, in the above construction any not-too-long path from α_0 to α_4 must go *through* α_2, and not merely travel close to α_2.

Remark 2.33. The above techniques also give another way to think about Corollary 2.25: Let $X_4 \subset S$ be the four-holed sphere component of $S \setminus \alpha_4$. We can find a conjugate of f supported in X_4, take a large power of it (twice as large as before), and apply it to α_0. This produces α_8 which is at distance eight from α_0. Proceeding in similar fashion constructs a geodesic segment of any desired length.

Again, this chain of thought can be summarized by:

Lemma 2.28. *Suppose that $g = \{\alpha_0, \alpha_1, \ldots, \alpha_n\}$ is a path in $\mathcal{C}(S)$. Suppose that every α_i cuts X, a non-pants essential surface in S. Then $d_X(\alpha_0, \alpha_n) \leq 6n$.* □

Lemma 2.28 works for any path in $\mathcal{C}$. If we instead have a geodesic, the conclusion becomes much stronger:

Theorem 2.34 (Theorem 3.1 of Masur and Minsky [**31**]). *Suppose that $g = \{\alpha_0, \alpha_1, \ldots, \alpha_n\}$ is a geodesic in $\mathcal{C}(S)$. Suppose that every α_i cuts X, a non-pants essential surface in S. Then $d_X(\alpha_0, \alpha_n) \leq M_1$, with M_1 a constant depending only on $\zeta(S)$.*

The point here is that, if we restrict our attention to geodesics, the size of the projection of the endpoints need only be sufficiently large (bigger than M). The projection does *not* need to be as long as the geodesic itself to ensure that the path meets the link of the boundary of X.

Again, we only hint at the proof: it relies on finding a sequence of singular Euclidean metrics on S (based on the geodesic g) and examining how these restrict to the subsurface X.

Equipped with Theorem 2.34 or Lemma 2.28 we may now deduce

PROPOSITION 2.35. *Suppose that S is not simple. Then $\partial_\infty \mathcal{C}(S)$ is not sequentially compact.* □

EXERCISE 2.36. Prove the proposition for sporadic S (that is, prove that $\partial_\infty \mathcal{F}$ is not compact).

EXERCISE 2.37. Prove the proposition when $\zeta(S) \geq 2$.

We end this section with an example of Hempel's (Summer 2005, Technion), shown in Figure 2.

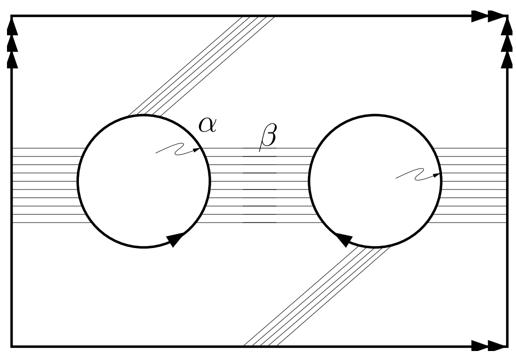

FIGURE 2. The sides of the rectangle are glued as shown. The two circles are glued by a reflection followed by 4/25 of a rotation to make the marked points agree. The result is a closed genus two surface. The circles become a simple closed curve α and the light lines close up to give β. Hempel claims that $d(\alpha, \beta) = 4$.

7. A trip to the zoo: separating and nonseparating curves

7.1. Definitions. Define $\mathrm{Nonsep}(S)$ to be the subcomplex of $\mathcal{C}(S)$ spanned by the vertices which are nonseparating curves. In general $\mathrm{Nonsep}(S)$ is non-empty and connected.

Exercise 2.38. Give the list of orientable compact surfaces so that Nonsep(S) is either empty or is not connected. (Caution: the connectedness proof is less trivial when ∂S is non-empty.)

To simplify our discussion, for the moment, we restrict our attention to closed surfaces.

Exercise 2.39. Prove, as long as $g \geq 2$, that the natural inclusion of Nonsep(S_g) into $\mathcal{C}(S_g)$ is an isometric embedding and induces a quasi-isometry of the two spaces.

Thus, in these cases, we deduce that Nonsep(S) is not especially interesting – or at least is nothing new. Define now Sep(S) to be the subcomplex of $\mathcal{C}(S)$ spanned by the vertices which are separating curves. In general Sep(S) is non-empty and connected.

Exercise 2.40. Prove that when $g \geq 3$ the complex Sep(S_g) is in fact connected. (As a hint: straight-forward curve surgery can't work, as it may produce non-separating curves. Think instead about doing a pair of curve surgeries. For a solution see [**32**].)

7.2. Sep(S) **is not quasi-isometrically embedded.** Again fix $g \geq 3$ and $S = S_g$. Let $\nu \colon \mathrm{Sep}(S) \to \mathcal{C}(S)$ be the natural inclusion.

Claim 2.41. The map $\nu \colon \mathrm{Sep}(S) \to \mathcal{C}(S)$ is not a quasi-isometric embedding.

Proof. Fix attention on a nonseparating curve $\alpha \subset S$. Let $X = S \smallsetminus N(\alpha)$ be the complement of a neighborhood of α. Note that every separating curve in S cuts X.

Now, for any $n \in \mathbb{N}$ we may choose β and β' in X which are separating in S and so that $d_X(\beta, \beta')$ is greater than $6n$ (use a partial pseudo-Anosov supported in X). Note that $d_S(\beta, \beta') = 2$.

On the other hand, suppose that $g = \{\beta_i\}_0^m$ is a path from β to β' in Sep(S). Then we may apply Lemma 2.28 to find that $m \geq n$. That is, $d_{\mathrm{Sep}}(\beta, \beta') \geq n$.

So vertices at arbitrarily large distance in Sep(S) are reduced to distance two by the map ν. So ν is not a quasi-isometric embedding. □

Exercise 2.42. Suppose that $g \geq 3$. Show that Sep(S_g) is not Gromov hyperbolic by finding a quasi-isometric embedding of $\mathbb{Z}^2$.

7.3. Bizzarre, irregular behavior. As we perhaps now expect, when S has low genus or boundary the complexes Nonsep(S) and Sep(S) display exceptional behavior. Please note that this section is not referenced by the rest of the text and may safely be ignored. Note also that this material was written to correct an significant error, kindly brought to my attention by Jason Behrstock, in a previous version of these notes.

We have the following exercise:

EXERCISE 2.43. If S is planar then Nonsep(S) is empty. If $S = S_{1,b}$ then Nonsep(S) is nonempty, but is not connected. If $S = S_{g,b}$ where $g \geq 2$ and b is arbitrary, then Nonsep(S) is connected. However, if $b \geq 2$ then Nonsep(S) is not quasi-isometrically embedded in $\mathcal{C}(S)$.

Fix $S = S_{1,b}$ with $b \geq 2$. Then Nonsep(S) is the first example we have seen where the complex does contain edges, but is not connected. As with $\mathcal{C}(S_1)$ and with $\mathcal{C}(S_{1,1})$, it is tempting to add edges between curves which intersect exactly once. After doing so there is again a fairly natural inclusion of Nonsep(S) into $\mathcal{C}(S)$ (send the center of an edge between intersecting curves to the boundary of the regular neighborhood of the union of the curves). Again, this is not a quasi-isometric embedding: for example, all $S_{1,1}$ subsurfaces are obstructions.

The complex of separating curves also has exceptional behavior in the bounded case.

EXERCISE 2.44. As above, if $S = S_{0,4}$ then Sep(S) is disconnected. If $S = S_1$ or $S_{1,1}$ then Nonsep(S) is empty. If S is any one of the surfaces $S_{1,2}$, S_2, or $S_{2,1}$ then Sep(S) is nonempty, but is not connected. If $S = S_{g,b}$ where $g \geq 2$ and $b \geq 2$, then Sep(S) is connected. The argument of Claim 2.41 shows that $\nu\colon \mathrm{Sep}(S) \to \mathcal{C}(S)$ is not a quasi-isometric embedding.

As above, we can add edges to Sep($S_{0,4}$) between curves that meet exactly twice, obtaining the Farey graph. This naturally suggests how to add edges to the complex of separating curves for other surfaces. That is, for S being any one of the surfaces $S_{1,2}$, S_2, or $S_{2,1}$ add edges between curves that meet exactly four times.

EXERCISE 2.45. Suppose S is one of $S_{1,2}$, S_2, or $S_{2,1}$. With this new definition of Sep(S), is Sep(S) connected?

EXERCISE 2.46. Determine the dimension of a maximal simplex in Sep(S_2). It is at least five.

EXERCISE 2.47. Even with this new definition of Sep(S_2) there is still a fairly natural map of Sep(S_2) into $\mathcal{C}(S_2)$ induced by sending a curve to itself. Following Claim 2.41 show that this map is not a quasi-isometric embedding. (Hint: You will need to use Corollary 2.29 instead of Lemma 2.28.)

Nonetheless we have:

CONJECTURE 2.48. $\mathrm{Sep}(S_2)$ *is Gromov hyperbolic.*

7.4. Curing the bizzare, irregular behavior. Fix $S = S_{g,b}$ with $b > 1$. We say that a separating essential non-peripheral curve α is a *pants curve* if $S \smallsetminus \alpha$ has a component which is a pair of pants. Let $\mathrm{Nonsep}'(S)$ be the subcomplex of $\mathcal{C}(S)$ containing all nonseparating curves and all pants curves.

EXERCISE 2.49. Prove that the natural inclusion of $\mathrm{Nonsep}'(S)$ into $\mathcal{C}(S)$ is an isometric embedding and induces a quasi-isometry of the two spaces.

CHAPTER 3

Estimating distance and hierarchies

1. A few simple examples

Let us leave the realm of the curve complex for a moment and discuss how to estimate distance in a few simple metric spaces.

As our first example consider $\mathbb{R}^2$ with the standard L^1 metric. Let $X, Y \subset \mathbb{R}^2$ be the x and y axes. Let $\pi_X\colon \mathbb{R}^2 \to X$ be the closest point projection map. Define π_Y similarly. As usual, for points $x, y \in \mathbb{R}^2$ define $d_X(x, y) = d_X(\pi_X(x), \pi_X(y))$ and define d_Y similarly. We have the expected formula:

$$d_{\mathbb{R}^2}(x, y) = d_X(x, y) + d_Y(x, y).$$

Here is a less trivial example. Let $F_2 = \langle a, b\rangle$ be the free group on two generators. Let Γ be the Cayley graph of F_2 with respect to the generators a and b: vertices are group elements and $g, g' \in F_2$ are connected by an edge if we can multiply g on the right by a, b, a^{-1}, or b^{-1} to obtain g'. Note that Γ is a copy of the four-valent tree. For any reduced word w which does not end in a or a^{-1} we define the *a-line* $L(w) \subset \Gamma$ to be the geodesic with vertex set $\{wa^n \mid n \in \mathbb{Z}\}$. Define the b-lines similarly.

For any a or b-line L we again have a closest point projection $\pi_L\colon \Gamma \to L$ and we can, for any $x, y \in F_2$, define

$$d_L(x, y) = d_L(\pi_L(x), \pi_L(y)).$$

Exercise 3.1. Show that $d_\Gamma(x, y) = \sum d_L(x, y)$ where the sum is taken over all a and b-lines. Also, if we are given the pairs $(w, d_{L(w)}(x, y))$, and if there are at least two such pairs, we can recover the points x and y.

Here is a final example: fix attention on $\mathbb{H}^2$ and tile it with copies of P, the regular right angled pentagon. Note that the tiling forms an infinite collection of geodesics. For any one of these, say L, we take $\pi_L\colon \mathbb{H}^2 \to L$ to be the closest point projection map. For points $x, y \in \mathbb{H}^2$ we define $d_L(x, y)$ as above.

Let c be the length of a side of P. Let $[r]_C$ be the *cut-off* function: $[r]_C = 0$ if $r < C$ and $[r]_C = r$ if $r \geq C$.

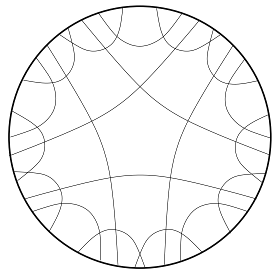

FIGURE 1. A terrible picture of the pentagonal tiling of $\mathbb{H}^2$. Do the world a favor and write a program on your computer to draw a better picture.

EXERCISE 3.2. Prove that, for any $C > 10c$ there are constants (K, E) so that

$$d_{\mathbb{H}^2}(x, y) \stackrel{K,E}{=} \sum [d_L(x, y)]_C.$$

The sum ranges over the lines formed by the tiling. (A good first step would be to show that only finitely many terms of the sum are non-zero.)

These examples may help to make the next section clearer.

2. Hierarchies, holes, and the pants complex

Let us try to make the ideas behind Claim 2.41 a bit more general. Suppose that $\mathcal{G}(S)$ is a simplicial complex, where the vertices are collections of essential non-peripheral curves or arcs in S, perhaps with some additional restrictions (*eg* separating). The edges and higher dimensional simplices of $\mathcal{G}(S)$ come from a relation of the form "α and β have small geometric intersection." We further insist that there be a natural map $\nu\colon \mathcal{G}(S) \to \mathcal{C}(S)$. Often this map is $\mathcal{MCG}(S)$ equivariant. For example: the subcomplex $\mathcal{G}(S) = \mathrm{Sep}(S) \subset \mathcal{C}(S)$.

DEFINITION 3.3. A non-pants essential subsurface $X \subset S$ is a *hole* for $\mathcal{G}(S)$ if every vertex of $\mathcal{G}(S)$ cuts X.

Generalizing Lemma 2.28 we have:

LEMMA 3.4. *For any $\mathcal{G}(S)$ there is a constant K so that if X is a hole for $\mathcal{G}(S)$ then the projection map $\pi_X\colon \mathcal{G}(S) \to \mathcal{C}(X)$ is K-Lipschitz.* ☐

DEFINITION 3.5. Suppose X is a hole for $\mathcal{G}$. Define the *diameter* of X to be $\mathrm{diam}(\pi_X(\mathcal{G}))$.

Often we are only interested in holes where the diameter is sufficiently large, say bigger than $M_2 > M_1$, the latter being the constant of Theorem 2.34. In these cases it often follows that the diameter is infinite.

Note that if $\mathcal{MCG}(S)$ acts naturally on $\mathcal{G}(S)$ then all holes have infinite diameter. This is due to the existence of partial pseudo-Anosov mappings with support equal to X. When this happens we find that if X is a hole and $X \subset Y$ then Y is a hole; it also follows that the entire surface S is a hole.

Note that it is *not* always the case that $\mathcal{MCG}(S)$ acts naturally on $\mathcal{G}(S)$: see our discussion of the disk complex $\mathcal{D}(V) \subset \mathcal{C}(S)$ in Chapter 4.

EXERCISE 3.6. Find all holes for $\mathrm{Nonsep}(S_g)$.

EXERCISE 3.7. Find all holes for $\mathrm{Sep}(S_2)$ and $\mathrm{Sep}(S_3)$.

EXERCISE 3.8. Suppose that $\mathcal{MCG}(S)$ acts naturally on $\mathcal{G}(S)$. Show that if $\mathcal{G}(S)$ has disjoint holes X and Y, then there is a quasi-isometrically embedded $\mathbb{Z}^2$ in $\mathcal{G}(S)$. (This generalizes Exercise 2.42, above.)

Here is another concrete example. Recall that a *pants decomposition* of S is a collection of disjoint curves in S cutting S into a collection of essential *pants*: three-holed spheres. Two pants decompositions $\{\alpha_i\}_{i=1}^{\zeta(S)}$ and $\{\beta_i\}_{i=1}^{\zeta(S)}$ are connected by an *elementary move* if $\alpha_i = \beta_i$ for all $i > 1$ and $0 < \iota(\alpha_1, \beta_1) < 3$. That is, α_1 and β_1 are at distance one in the Farey graph of the surface they fill. Let $\mathcal{P}(S)$ be the *pants graph*: vertices are pants decompositions and edges are elementary moves.

It follows that X is a hole for $\mathcal{P}(S)$ as long as X is not a simple surface. We have a remarkable theorem of Masur and Minsky (see Section 8 of [**31**]):

THEOREM 3.9 (Masur-Minsky). *There is a constant $C' = C'(S)$ where, for any $C > C'$, there are constants (K, E) so that*

$$d_{\mathcal{P}}(P, P') \stackrel{K,E}{=} \sum [d_X(P, P')]_C$$

where the sum is taken over all holes X for $\mathcal{P}(S)$.

This generalizes their estimate of *word length* in $\mathcal{MCG}(S)$. Equivalently, they estimate distance in the *marking complex* $\mathcal{MC}(S)$:

THEOREM 3.10 (Masur-Minsky). *There is a constant $C' = C'(S)$ where, for any $C > C'$, there are constants (K, E) so that*

$$d_{\mathcal{P}}(\mu, \mu') \stackrel{K,E}{=} \sum [d_X(\mu, \mu')]_C$$

where the sum is taken over all holes X for $\mathcal{MC}(S)$.

Of course all essential subsurfaces $X \subset S$ (except for pants) are holes for $\mathcal{MCG}(S)$ and the marking complex. See the discussion of *heirarchies*, below, for a brief description of the marking complex.

Now, if $\mathcal{G}(S)$ is fairly well behaved, we can hope that the "distance estimate" of Theorem 3.9 holds:

CONJECTURE 3.11. *There is a constant $C' = C'(\mathcal{G}(S))$ where, for any $C > C'$, there are constants (K, E) so that*

$$d_{\mathcal{G}}(x, y) \overset{K,E}{=} \sum [d_X(x, y)]_C$$

where the sum is taken over all holes X for $\mathcal{G}(S)$.

In particular, when $\mathcal{G}(S)$ equals either the arc complex of a surface with boundary or the disk complex of a handlebody (both defined below) then the above conjecture is work-in-progress of Howard Masur and myself. On the other hand, if $\mathcal{G}$ is not quasi-convex (say, $\mathcal{C}(S)$ minus a sequence of disjoint metric balls which increase in size) then the distance estimate will *not* hold.

Trivially, we have

EXERCISE 3.12. Conjecture 3.11 holds for $\mathrm{Nonsep}(S_g)$.

Here is a general scheme for verifying Conjecture 3.11. We follow Masur and Minsky's proof of Theorem 3.9.

The crucial tool they introduce is the notion of a *hierarchy*. Suppose that x and y are vertices of $\mathcal{G}(S)$. These give simplices in $\mathcal{C}(S)$. Choose a *tight* geodesic $g \subset \mathcal{C}(S)$ from the simplex x to the simplex y. The "vertices" of this geodesic may be simplices in $\mathcal{C}(S)$. Write $g = \{v_i\}$. In the simplest case tightness tells us that $v_{i\pm1}$ fill some component X_i of the complement of v_i (and in fact v_i equals the essential non-peripheral components of the boundary of $N(v_{i-1} \cup v_{i+1})$). For each i we choose a tight geodesic in $\mathcal{C}(X_i)$ connecting v_{i-1} to v_{i+1} and for each of these we do the same, and so forth. (There are delicate issues to cope with when switching from one X_i to another.) We note that, by Lemma 6.2 of [**31**], the length of the geodesic we chose in a subsurface X is quasi-equal to the size of the projection $d_X(x, y)$.

This entire structure is the hierarchy connecting x to y. The hierarchy can be broken into a sequence of *markings*: in the generic case this is a pants decomposition Q of the surface S together with *transversals* for each $\alpha \in Q$. (It is not too far wrong to think of the transversal for α as being a curve β which meets α exactly once or twice and which is disjoint from all curves in $Q \setminus \alpha$. Strictly speaking, what we have set forth is called a *complete clean marking* by Masur and Minsky.)

Let $\{\mu_i\}$ denote this sequence of markings. Consecutive markings are related by two moves "Flip" (switch a pants curve with its transversal) and "Twist" (Dehn twist a transversal once about its pants curve). We can use the hierarchy to determine a small collection of subsurfaces in which each of these moves occurs – this is the drift of pages 962 to 963 of [**31**]. For each move which occurs in a hole X we must find a vertex $\alpha_i \in \mathcal{G}(S)$ which meets α_{i-1} in a uniformly bounded number of points. (As we shall see, it is not always possible to choose a vertex of $\mathcal{G}(S)$ which lies completely inside of the hole X.) For moves not occuring in a hole retain the previously chosen vertex of $\mathcal{G}(S)$. Consecutively chosen vertices of $\mathcal{G}(S)$, say α_i and α_{i+1}, may not be adjacent in $\mathcal{G}(S)$ but, as they have uniformly bounded intersection they have uniformly bounded distance in $\mathcal{G}(S)$.

This construction should give the upper bound: $d_{\mathcal{G}}(x, y)$ is less than the sum of the subsurface projections to the holes (with some choice of (K, E), as usual). The lower bound follows from the fact that the marking moves cannot occur in more than a small collection of holes simultaneously.

We end this section with another simple conjecture:

CONJECTURE 3.13. *Suppose that any pair of infinite diameter holes X and Y for $\mathcal{G}(S)$ intersect. Then $\mathcal{G}(S)$ is Gromov hyperbolic.*

Examples of this may be found in [**7**] or [**2**]. Both rely heavily on [**31**]. It would follow, for example, that $\mathrm{Sep}(S_2)$ is Gromov hyperbolic.

EXERCISE 3.14. Suppose that $\mathcal{G}(S)$ has n disjoint holes, and $\mathcal{MCG}(S)$ acts naturally on $\mathcal{G}(S)$. Show that there is a quasi-isometric embedding of $\mathbb{Z}^n$ in $\mathcal{G}(S)$.

The image of a quasi-isometric embedding of $\mathbb{Z}^m$, for $m > 1$, is called a *quasi-flat*. The maximal possible rank of a quasi-flat in $\mathcal{G}(S)$ is called the *rank* of $\mathcal{G}(S)$. Generally it is somewhat difficult to compute this quantity. There is a preprint of Brock and Masur showing that the rank of $\mathcal{P}(S_2)$ is two.

Ask one of them for a copy of their paper and imitate the techniques within to prove that the rank of $\mathrm{Sep}(S_g)$ is two, for all $g > 2$. In particular, generalize Hruska's *isolated flats property* (see [**21**]) to this context and prove that the property holds for $\mathrm{Sep}(S_g)$, if $g > 2$.

3. A trip to the zoo: the arc complex

Define the *arc complex* $\mathcal{A}(S)$, when $\partial S \neq \emptyset$, as follows: vertices are essential arcs in S. A collection of $k + 1$ vertices span a k-simplex if all $k+1$ of the arcs can be realized disjointly. As usual, we generally restrict attention to the one-skeleton. The distance in $\mathcal{A}(S)$ between two vertices is the minimal possible number of edges in an edge path between them.

There is a natural map $\pi_S\colon \mathcal{A}(S) \to \mathcal{C}(S)$ defined exactly as the subsurface projections were defined, above: let $\pi_S(\alpha)$ be any component of the boundary of $N(\alpha \cup \partial S)$, a regular neighborhood of $\alpha \cup \partial S$.

In order to understand π_S a bit better we define the *arc and curve complex* $\mathcal{AC}(S)$, when $\partial S \neq \emptyset$ as follows: vertices are essential arcs or curves in S. As usual, a collection of $k+1$ vertices span a k-simplex if all $k+1$ of the arcs and curves can be realized disjointly.

Exercise 3.15. Show that the inclusion of $\mathcal{C}(S)$ into $\mathcal{AC}(S)$ is a quasi-isometry. The map π_S and the proof of Lemma 2.27 will be useful.

However we also find:

Claim 3.16. Suppose that S is not planar, S has at least one boundary component, and $S \neq S_{1,1}$. Then the map $\pi_S\colon \mathcal{A}(S) \to \mathcal{C}(S)$ is not a quasi-isometry.

Proof of Claim 3.16. The proof follows the outline provided by Claim 2.41. The only change is the set of holes: let X be any essential non-simple subsurface of S, not equal to all of S, but which contains all of the boundary components of S. To be concrete, let X be the complement of a non-separating curve in S. It is clear that every essential arc cuts X. We can now choose a partial pseudo-Anosov supported on X and the proof goes through as before. □

Note that Claim 3.16 lends context to a question asked by Brian Bowditch:

Question 3.17. As g and b vary how do the quasi-isometry types of $\mathcal{C}(S_{g,b})$ and $\mathcal{A}(S_{g,b})$ change?

We may look to Conjectures 3.11 and 3.13 for other facts which may be true about $\mathcal{A}(S)$.

We end this section by sketching an elegant argument shown to me by Feng Luo, adapting an argument of Darryl McCullough [**33**]:

Claim 3.18 (Harer [**15**]). The simplicial complex $\mathcal{A}(S)$ is contractible.

This contrasts with $\mathcal{C}(S)$ which, as remarked above, is typically an infinite wedge of spheres [**15**].

PROOF OF CLAIM 3.18. Fix attention on a single arc α. Suppose that K is a finite simplicial complex and $f\colon K \to \mathcal{A}(S)$ is a simplicial map. We will show that f contracts to a point. Note that if $f(K)$ is contained in the one-neighborhood of α there is a homotopy of f to the constant map $f\colon K \to \alpha$ as desired.

Suppose not. Choose a generic hyperbolic metric on S so that the boundary of S is geodesic. For each vertex $f(v)$ of the image, straighten $f(v)$ to be a geodesic and do the same for α. Define $|f| = \sum \iota(\alpha, f(v))$ where the sum ranges over the vertices of K. Orient α and let v be the vertex of K so that $f(v)$ is the first arc you meet while traveling along α in the direction of the orientation. Let β be the arc connecting the beginning of α to this first intersection point of α and $f(v)$.

We can surger $f(v)$ along β to obtain two essential arcs: form $\beta \cup f(v)$ and take a regular neighborhood. This is a hexagon in S and two of the three sides, γ and γ', are the desired arcs. (Check that these are both essential.) Finally, define $f'\colon K \to \mathcal{A}(S)$ to agree with f on all of $K^0 \setminus v$ and set $f'(v) = \gamma$. This is again a simplicial map and $|f'| < |f|$. The induction is complete.

We have shown that any map of a finite simplicial complex into $\mathcal{A}(S)$ may be homotoped to α. By Whitehead's Theorem [**19**] $\mathcal{A}(S)$ is contractible. □

REMARK 3.19. There is a subtle point here – Whitehead's Theorem is usually stated in terms of CW complexes. However, $\mathcal{A}(C)$ thought of as a CW complex is not metrizable! This is somewhat uncomfortable, as we have been thinking about $\mathcal{A}(S)$ as a metric space. One solution is to prove Whitehead's Theorem for simplicial complexes, with metric induced by taking every simplex to be isometric to the Euclidean simplex with side lengths equal to one. (See Bridson's paper [**5**].) We must then verify that Whitehead's Theorem in fact holds for $\mathcal{A}(S)$ equipped with this metric topology.

EXERCISE 3.20. Can you improve the argument to give an explicit deformation retraction to α?

EXERCISE 3.21. Show that the complex $\mathcal{AC}(S)$ is contractible.

Even better than contractibility or Gromov hyperbolicity would be a control over the global curvature of $\mathcal{A}(S)$. Our combinatorial complexes are not manifolds, so we do not expect any kind of Riemannian curvature. However there are synthetic geometry definitions of curvature, for example the notion of a CAT(κ) space (see Bridson's and Haefliger's book [**6**]).

QUESTION 3.22. Is $\mathcal{A}(S)$ a CAT(κ) space for some $\kappa \leq 0$?

CHAPTER 4

Handlebodies

In this chapter we turn from the study of the curve complex and directly related objects in order to begin our study of Heegaard splittings. We refer the reader to Scharlemann's survey article [**37**] for a detailed treatment.

Following the work of Minsky and others (Brock, Souto, Namazi...) we expect that the way that a Heegaard splitting interacts with the curve complex will inform the geometry of the underlying manifold.

Our immediate goal is more modest: What is a "good" Heegaard diagram and, if given a bad Heegaard diagram, how can we find such a good diagram?

1. Basic definitions

Recall that a *handlebody* V is a compact three-manifold which is homeomorphic to a closed regular neighborhood of a finite, connected, polygonal graph in $\mathbb{R}^3$. The graph is called a *spine* for V. The *genus* of V is the genus of ∂V.

It is a basic theorem of three-manifold topology (see Rolfsen [**35**]) that any closed orientable three-manifold M can be obtained by taking a pair of handlebodies V and W, of the same genus, and gluing them together by a homeomorphism $f\colon \partial V \to \partial W$. We usually denote the image of ∂V inside of M as the surface S and call S a *Heegaard splitting surface*. The triple (S, V, W) completely determines M and is called a *Heegaard splitting* of M.

We note that a three-manifold never has only one splitting: new ones may be obtained from old by a process called *stabilization*. This is defined as follows: Fix attention on a splitting $S \subset M$ and let B be a small ball in M with $B \cap S = D$ being an equatorial disk. Remove two smaller disks from D and add an unknotted tube in B to create a surface S' with genus one higher. The surface S' cuts M into two handlebodies V' and W' and is the *stabilization* of S.

EXERCISE 4.1. Stabilization is *unique*: that is, the splitting S' depends only on S and not on the choices made in the stabilization procedure.

Theorem 4.2 (Reidemeister and Singer). *Any two Heegaard splittings of a closed three-manifold have a common stabilization.*

A proof can be derived from the fact that two triangulations of M have a common subdivision.

We now must understand how the handlebodies V and W interact with the curve complex. Recall that a properly embedded disk $D \subset V$ is *essential* in the handlebody V if $\partial D \subset S = \partial V$ is an essential curve. Define the *disk complex* $\mathcal{D}(V) \subset \mathcal{C}(S)$ to be the subcomplex spanned by the boundaries of all essential disks.

Exercise 4.3. Show that $\mathcal{D}(V)$ is connected. Even better: show that $\mathcal{D}(V)$ is contractible. (Or read McCullough's paper [**33**].)

Exercise 4.4. Show that a splitting (S, V, W) is stabilized if and only if there are essential disks $D \subset V$ and $E \subset W$ so that $\iota(\partial D, \partial E) = 1$.

Define the distance between subcomplexes $\mathcal{X}, \mathcal{Y}$ of the curve complex to be the minimal possible number of edges in an edge path connecting a vertex of $\mathcal{X}$ to $\mathcal{Y}$. We denote this distance by $d_S(\mathcal{X}, \mathcal{Y})$.

Hempel [**20**] defines the *distance*, $d_S(V, W)$, of a Heegaard splitting (S, V, W) to be the number $d_S(\mathcal{D}(V), \mathcal{D}(W))$. This generalizes several more classical definitions:

- $S \subset M$ is *reducible* if and only if $d_S(V, W) = 0$.
- $S \subset M$ is *weakly reducible* if and only if $d_S(V, W) \leq 1$ (Casson and Gordon [**9**]).
- $S \subset M$ has the *disjoint curve property* if and only if $d_S(V, W) \leq 2$ (Thompson [**39**]).

The negations are called *irreducible*, *strongly irreducible*, and *full*, respectively.

We also recall the definition of the *handlebody group*. Let $\mathcal{MCG}(V)$ be the group of proper isotopy classes of homeomorphisms of V. Note that, for every $f \in \mathcal{MCG}(V)$ there is a mapping class $\partial f \in \mathcal{MCG}(S) = \mathcal{MCG}(\partial V)$.

Exercise 4.5. Check that $\mathcal{MCG}(V)$ is in fact a group. Prove that $\partial \colon \mathcal{MCG}(V) \to \mathcal{MCG}(S)$ is an injection.

Exercise 4.6. Does every torsion element in $\mathcal{MCG}(V)$ arise as a symmetry of some spine for V?

There is a purely "group-theoretic" approach to Heegaard splittings: they correspond to *double cosets* of $\mathcal{MCG}(S)$ by $\mathcal{MCG}(V)$, the handlebody group.

2. The disk complex is quasi-convex

Recall that a subset $\mathcal{Y}$ of a geodesic metric space $\mathcal{X}$ is *convex* if every geodesic with endpoints in $\mathcal{Y}$ is contained in $\mathcal{Y}$.

It is possible to coarsen this idea: we say that $\mathcal{Y}$ in *quasi-convex* in $\mathcal{X}$, with constant R, if every geodesic with endpoints in $\mathcal{Y}$ is contained in an R neighborhood of $\mathcal{Y}$.

We can now state another striking result of Masur and Minsky [**29**]:

THEOREM 4.7. *The disk complex $\mathcal{D}(V)$ is a quasi-convex subset of the curve complex $\mathcal{C}(\partial V)$.*

As usual we only hint at the proof: Masur and Minsky find a sequence of essential disks D_i joining D to E by doing a sequence of *disk surgeries*. (See below.) Furthermore, they arrange that the ∂D_i occur as the vertices of a nested sequence of train tracks. Thus the sequence D_i is quasi-convex and the theorem follows.

COROLLARY 4.8. *Fix M_1 as in Theorem 2.34. Suppose that X is an essential, non-simple subsurface of S. Suppose also that $d_S(\partial X, \mathcal{D}(V))$ is greater than the constant of quasi-convexity of $\mathcal{D}(V)$. Then X is a hole for $\mathcal{D}(V)$ with diameter bounded by M_1.*

PROOF. Suppose X is as given by hypothesis. X is a hole because $d_S(D, \partial X) > 1$ for any disk D. Suppose now that D and E are essential disks in V and g is a geodesic in $\mathcal{C}(S)$ connecting ∂D to ∂E. Note that every vertex of g cuts X, so we may apply Theorem 2.34 to find that $d_X(D, E)$ is at most M_1. As the same holds for any pair of disks, the diameter of $\pi_X(\mathcal{D}(V))$ is bounded by M_1. □

Thus, if large diameter holes do exist for $\mathcal{D}(V)$, they must be relatively close to $\mathcal{D}(V)$.

EXERCISE 4.9. Before reading the next section: can you find a large diameter hole for $\mathcal{D}(V)$?

3. I-bundles in $\mathcal{D}(V)$

Here we present a "standard example" (shown to me by Hossein Namazi) which will inform the rest of our discussion of $\mathcal{D}(V)$. Begin as follows:

DEFINITION 4.10. A curve $\alpha \in \mathcal{C}(S)$ is a *dead end* for a subset $\mathcal{X} \subset \mathcal{C}(S)$ if, for all β such that $d_S(\alpha, \beta) = 1$, we have $d_S(\beta, \mathcal{X}) + 1 = d_S(\alpha, \mathcal{X})$.

Now fix $F = S_{1,1}$ and take $V = F \times [0,1]$. Note that V is the genus two handlebody. As usual set $S = \partial V$. As a bit of notation, take $\alpha = \partial F \times \{1/2\}$ and take $X = F \times \{0\}$, $Y = F \times \{1\}$. Let $\mathrm{proj}_F \colon V \to F$ be projection onto the first factor.

Proposition 4.11 (Namazi). *The curve α is a dead end for $\mathcal{D}(V)$.*

We require one definition: if β is a curve or arc in F then the surface $\mathrm{proj}_F^{-1}(\beta)$ is a *vertical* surface.

Proof. Suppose that β is any essential non-peripheral curve in X. Suppose that $\delta \subset X$ is an essential arc disjoint from β. Let B be the vertical annulus having β as a boundary component. Let D be the vertical disk with $\delta \subset \partial D$.

Note that B is disjoint from D. Also, α is disjoint from β, which is disjoint from ∂D. The proposition follows. □

We can adapt the standard example above to prove:

Claim 4.12. The inclusion of the disk complex $\mathcal{D}(V)$ into $\mathcal{C}(S)$ is not a quasi-isometric embedding.

Proof. Choose a compact, connected orientable surface F with one or two boundary components so that $V \cong F \times [0,1]$. As above, set $X = F \times \{0\}$ and $Y = F \times \{1\}$. We will show that X is an infinite diameter hole for $\mathcal{D}(V)$

First note that Y is incompressible in V. So every disk meets X and so X is a hole. Second, fix δ, an essential arc in F. Fix $f \colon F \to F$, a pseudo-Anosov map on F. Let $\epsilon = f^n(\delta)$ where n is arbitrary. Let $D = \mathrm{proj}_F^{-1}(\delta)$. Let $E = \mathrm{proj}_F^{-1}(\epsilon)$. Then $\pi_X(D, E) \geq n/2$ and we are done. □

4. Holes in $\mathcal{D}(V)$

Here we sketch a classification of all holes for $\mathcal{D}(V)$, with diameter bounded below. Please note that this is a work in progress.

We begin with a few pieces of notation for I-bundles. In what follows we only consider I-bundles $I \to T \overset{\mathrm{proj}_F}{\to} F$ with *total space* T being orientable. So, given the *base surface* F which must be a compact, with boundary, connected surface, we take T to be the *orientation I-bundle*: T is a product, $F \times I$, if F is orientable and T is twisted, $F \tilde{\times} I$, if F is non-orientable. Recall that the vertical surface in T above a curve $\alpha \subset F$ is an annulus or Möbius band exactly as α preserves or does not preserve orientation in F.

We call the vertical surface in T lying above ∂F the *vertical boundary* of T. Denote this as $\partial_v T = \mathrm{proj}_F^{-1}(\partial F)$. The closure of the complement of of $\partial_v T$ is the horizontal boundary. We write this as $\partial_h T = \overline{\partial T \smallsetminus \partial_v T}$.

We may now state our main theorem:

THEOREM 4.13. *Fix $V = V_g$. There is a constant $M_2 = M_2(V)$ with the following property: Suppose that $X \subset S = \partial V$ is a hole for $\mathcal{D}(V)$ with diameter at least M_2. Then either X compresses in V or X is incompressible in V and there is an I-bundle $T \subset V$ with the following properties:*

- *the surface X is a component of $\partial_h T$,*
- *$\partial_h T \subset S$, and*
- *at least one component of $\partial_v T$ is contained in S.*

In any case, the surface X is not simple.

REMARK 4.14. Note that there is a marked similarity in the above theorem to some of the statements of JSJ theory for pared handlebodies. The same similarity (but not identity!) holds for the proofs. This line of thought was inspired by Oertel's [**34**] investigation of handlebody automorphisms.

We proceed via a sequence of claims.

CLAIM 4.15. If $X \subset S$ is a hole for $\mathcal{D}(V)$ then $Y = S \smallsetminus X$ is incompressible.

This follows directly from the definitions.

CLAIM 4.16. If $X \subset S$ is an essential subsurface and X compresses in V then $d_S(\partial X, \mathcal{D}(V)) \leq 1$.

This is obvious, but provides a bit of context for Corollary 4.8. We may now turn to the incompressible case.

We begin with a standard definition:

DEFINITION 4.17. Suppose that D is an essential disk in V. Suppose that that E is a disk in V so that $\partial E = \alpha \cup \beta$ with

- $\alpha \cap \beta = \partial \alpha = \partial \beta$,
- $E \cap S = \alpha$,
- $E \cap D = \beta$, and
- the arc α is essential in $S \smallsetminus \partial D$.

Then E is a *boundary compression* for D and we may *surger* D along E as follows: Let E' and E'' be two parallel copies of E. Set $D' \cup D'' = (D \smallsetminus N(\beta)) \cup E' \cup E''$. Then D' and D'' are the surgered disks. Note that these are both essential in V.

We specialize this definition as follows:

Definition 4.18. Suppose that D is an essential disk in V and that X is an essential subsurface in S. Suppose that E is a boundary compressing disk for D with $E \cap S = \alpha \subset X$. Then we call E a ∂_X *compressing disk*.

Claim 4.19. Suppose that D is an essential disk in an I-bundle $T \to F$ and D has been isotoped to minimize $\partial D \cap \partial_v T$. Suppose F is orientable. Let $X \cup Y = \partial_h T$. We have $\operatorname{proj}_F(D \cap X)$ is at most distance one from $\operatorname{proj}_F(D \cap Y)$ in $\mathcal{A}(F)$.

Here is a sketch: The claim is true for vertical disks. Now induct on the number of ∂_X compressions required to make D vertical. The induction hypothesis needs to be a bit stronger than that stated in the claim: instead of a single pair of arcs $(\alpha, \beta) \subset (D \cap X, D \cap Y)$ with disjoint projection to F several such pairs are needed.

Remark 4.20. If F is nonorientable let $X = \partial_h T$ and let $\tau \colon \mathcal{A}(X) \to \mathcal{A}(X)$ be the natural involution. In this case if $D \subset V$ is an essential disk then $D \cap X$ is at most distance one from the fixed point set of τ.

Claim 4.21. Suppose X is an incompressible hole for $\mathcal{D}(V)$ and D is an essential disk. Then there is an essential disk D'' in V which is ∂_X incompressible, is $\partial_{S \setminus X}$ incompressible, and has $d_X(D, D'') < M_3$. Here M_3 is a constant independent of X, D, and S.

A sketch: Given D do $\partial_{S \setminus X}$ compressions until no more are possible to obtain a disk D'. Note that $D' \cap X$ is nonempty, as X is incompressible. As all of the boundary compressions are disjoint from X we may assume that the projections $\pi_X(D)$ and $\pi_X(D')$ are identical.

Now let N be a regular neighborhood of X in V. Then $I \to N \to X$ is a product I-bundle. For every ∂_X compressing disk E of D' use E to guide an isotopy of D' which is the identity outside of a regular neighborhood of E and which isotopes $E \cap D'$ into N. After all of these isotopies we may apply Claim 4.19 and Lemma 2.27 (several times) to obtain the claim.

We are now equipped to give a sketch of the proof of Theorem 4.13: Suppose that X is an incompressible hole for $\mathcal{D}(V)$. Suppose that D and E are essential disks in V so that $d_X(D, E)$ is larger than M_2. Applying Claim 4.21 twice we may assume that D and E are both ∂_X and $\partial_{S \setminus X}$ incompressible, while maintaining the fact that $d_X(D, E)$ is quite large.

We now regard D and E as polygons, with vertices being the points of $\partial D \cap \partial X$ and $\partial E \cap \partial X$. Note that the arcs of $D \cap E$, lying in D, form a collection of diagonals of D. A combinatorial argument proves that there is an arc $\alpha \subset \partial D \cap X$ which has at most four types of diagonal adjacent to it. Note that there is at least one type of diagonal adjacent to α as $d_X(\alpha, \partial E)$ is quite large. Identically there is an arc $\beta \subset \partial E \cap X$ with similar properties.

We deduce the existence of a pair of rectangles $Q \subset D$ and $R \subset E$ with the following properties:

- exactly one side $\alpha' \subset \partial Q$ ($\beta' \subset \partial R$) is contained in α (β)
- the two sides of Q (R) adjacent to α' (β') are parallel diagonals in D (E)
- the number of intersections of α' and β' is at least one-sixteenth of $\iota_X(\alpha, \beta)$.

Finally, we claim that α' and β' fill X. Thus we may take a regular neighborhood of $\alpha' \cup \beta'$, add vertical three-balls as necessary, and find the desired I-bundle T having X as a component of $\partial_h T$.

Remark 4.22. This completes the proof of Theorem 4.13, except for the case of annuli. This case is covered in a forth-coming paper with H. Masur.

Remark 4.23. We also note that Conjectures 3.11 and 3.13 should also apply to $\mathcal{D}(V)$, suggesting that $\mathcal{D}(V)$ is Gromov hyperbolic.

The applicability of Conjecture 3.13 seems puzzling at first – for instance in the standard example $V = F \times I$ the top and bottom surfaces X and Y are disjoint and are both holes. However, by Claim 4.19 the projection of a disk D into X is (for our purposes) identical to the projection of D into Y. Thus X and Y are not really "disjoint."

5. Heegaard diagrams

We now recall another piece of standard terminology:

Definition 4.24. Suppose that S is a closed connected orientable surface with genus g. A collection Δ of g disjoint curves $\{\alpha_i\}$, so that $S \smallsetminus \Delta$ is homeomorphic to $S_{0,2g}$, is called a *cut system* for S.

Exercise 4.25. Define the *Hatcher-Thurston* graph $\mathcal{HT}(S)$ as follows: vertices are cut systems for S and two such, $\{\alpha_i\}$ and $\{\beta_i\}$, are connected by an edge if firstly $\alpha_i = \beta_i$ for $i \neq 1$ and secondly $\iota(\alpha_1, \beta_1) = 1$. (This graph was introduced in the course of their proof that the mapping class group is finitely presented [**18**]. Their paper also introduces the pants complex, $\mathcal{P}(S)$, in an appendix.)

Find the holes for $\mathcal{HT}(S)$. You might also prove that $\mathcal{HT}(S)$ is connected. What is the maximal size of a complete subgraph of $\mathcal{HT}(S)$?

Note that a cut system $\Delta(V)$ uniquely determines a handlebody V with boundary S. The converse is far from true; as long as $g > 1$ the handlebody V has infinitely many cut systems. As $\mathcal{MCG}(V)$ contains maps f with ∂f pseudo-Anosov there are, for any n, cut systems $\Delta(V)$ and $\Delta'(V)$ at distance more than n in the curve complex $\mathcal{C}(S)$.

Exercise 4.26. Find an explicit $f \in \mathcal{MCG}(V)$ so that ∂f is pseudo-Anosov.

Exercise 4.27. Wajnryb [**41**] has introduced a graph similar to that of Hatcher and Thurston in his study of the handlebody group: Fix a handlebody V with boundary S and define a graph with vertices being cut systems for V. Connect two systems $\{\alpha_i\}$ and $\{\beta_i\}$ by an edge if firstly $\alpha_i = \beta_i$ for $i \neq 1$ and secondly $\iota(\alpha_1, \beta_1) = 0$. We call this graph $\mathrm{Waj}(V)$

What are the maximal complete subgraphs for $\mathrm{Waj}(V)$? (Careful: there are two "kinds.") What are the holes for $\mathrm{Waj}(V)$? (Careful: the symmetry group of $\mathrm{Waj}(V)$ is $\mathcal{MCG}(V)$, not $\mathcal{MCG}(S)$.) Is it connected? Can you add cells to make it simply connected?

Definition 4.28. Suppose that (S, V, W) is a Heegaard splitting and that $\Delta(V)$ and $\Delta(W)$ are cut systems determining V and W. Then the triple $(S, \Delta(V), \Delta(W))$ is a *Heegaard diagram* for the splitting S.

The obvious question immediately arises: given a Heegaard diagram, what can we deduce about the underlying splitting S or, better yet, about the underlying manifold $M = V \cup_S W$?

A great deal of work has gone into this question as it has connections to problems such as three-sphere recognition, deciding (strong) irreducibility of splittings, and the Poincare Conjecture.

6. Almost computing the Hempel distance

We end by sketching a possible application of our work to a closely related question. Recall that the Hempel distance of a splitting (S, V, W) is $d_S(V, W)$: the minimal possible number of edges in an edge path from $\mathcal{D}(V)$ to $\mathcal{D}(W)$ inside of $\mathcal{C}(S)$.

Question 4.29. Find an algorithm which, given a Heegaard diagram $(S, \Delta(V), \Delta(W))$, computes the distance $d_S(V, W)$.

This question seems somewhat out of reach, as least for our coarse geometric techniques. Instead we consider:

Question 4.30. Find an algorithm which, given a Heegaard diagram $(S, \Delta(V), \Delta(W))$, computes the distance $d_S(V, W)$ up to an error term which is *a priori* bounded by some function of $\zeta(S)$.

This appears to be more tractable. We proceed as follows: Suppose that S is a closed, orientable, connected surface with genus two or larger.

CONJECTURE 4.31. *There is an algorithm which, given a vertex $\alpha \in \mathcal{C}(S)$ and a cut system $\Delta(V)$ in S, finds a disk $D \subset V$ so that*

$$d_S(\alpha, D) \leq d_S(\alpha, \mathcal{D}(V)) + M_4$$

where M_4 is a constant depending only on the topological type of S.

Note that an answer to Question 4.30 follows by applying the hyperbolicity of $\mathcal{C}(S)$ and the quasi-convexity of $\mathcal{D}(V)$.

Here is a "work-in-progress" approach to Conjecture 4.31: Build algorithmically a hierarchy H between α and $\Delta(V)$. (See recent work of Shackleton [**38**].) Let D be a disk of $\mathcal{D}(V)$ which is as close as possible to α. (We are trying to construct D or in fact any disk which is within a bounded distance of D.) Consider H', a hierarchy between $\Delta(V)$ and D. By hyperbolicity of $\mathcal{C}(S)$ the hierarchies H and H' should fellow-travel until H "turns" and moves directly away from $\mathcal{D}(V)$ towards α.

Now, the large links along H' should only occur in holes for $\mathcal{D}(V)$. Note that if we are given a subsurface $X \subset V$ we can algorithmically decide if X is a hole for $\mathcal{D}(V)$. As H and H' fellow-travel they should have identical large links. So the large links along H should be holes for $\mathcal{D}(V)$ until H turns towards α. To find a disk near D it suffices to find this corner which is almost equivalent to finding the last time a large link in H is a hole for $\mathcal{D}(V)$.

7. A trip to the zoo: Scharlemann's complex

We end the chapter with a graph recently introduced by Scharlemann [**36**]. Let S_2 be the standard genus 2 Heegaard splitting for the three-sphere, S^3. We say that a separating curve $\alpha \subset S$ is a *reducing curve* if α bounds a disk in both of the handlebodies $V \cup W = S^3 \setminus S$.

We define a complex $\mathcal{MS}(S_2)$ as follows: the vertices are reducing curves for S_2 in S^3. Two such are connected by an edge if they intersect exactly four times. We ask our now standard list of questions about $\mathcal{MS}(S_2)$:

- What is the maximal complete subgraph?
- It the graph connected? (See Scharlemann's paper [**36**] for answers to both of these questions.)
- What cells need we add to make it simply connected?
- Finally, what are the holes for $\mathcal{MS}(S_2)$?

REMARK 4.32. A student of Culler's, Erol Akbas, has found a proof that $\mathcal{MS}(S_2)$ is quasi-isometric to a tree: in fact, the longest simple loop in $\mathcal{MS}(S_2)$ has length three.

The motivated reader will find several interesting questions posed in Scharlemann's paper. In particular, what is the correct approach to the group of automorphisms of the genus g Heegaard splitting of S^3, when $g > 2$?

CHAPTER 5

Ends and boundaries

The main result we wish to reach is the following:

THEOREM 5.8. *Fix $g \geq 2$. For any $\omega \in \mathcal{C}^0(S_{g,1})$ and for any $r \in \mathbb{N}$ the complex $\mathcal{C}(S_{g,1}) \smallsetminus B(\omega, r)$ is connected.*

Here $B(\omega, r)$ is the closed ball of radius r about ω. This answers a question of Masur's, as least for $S = S_{g,1}$ with $g \geq 2$. The theorem is perhaps unexpected when compared to Remark 2.32 or compared with the unsettled status of Storm's:

QUESTION 2.16. Is the Gromov boundary of the curve complex, $\partial_\infty \mathcal{C}(S)$, connected?

1. Proof sketch

We prove Theorem 5.8 in two steps. Fix a basepoint $\omega \in \mathcal{C}^0(S)$. We first show:

PROPOSITION 5.1. *The curve complex has no dead ends with respect to ω.*

Recall that by Definition 4.10 a curve $\alpha \in \mathcal{C}(S)$ is a *dead end* for ω if, for all β such that $d_S(\alpha, \beta) = 1$, we have $d_S(\beta, \omega) + 1 = d_S(\alpha, \omega)$.

The next step is to investigate the natural map $\pi_* \colon \mathcal{C}(S_{g,1}) \to \mathcal{C}(S_g)$ which "caps-off" the boundary component. Fix $\tau \in \mathcal{C}(S_g)$ and let $\mathcal{F}_\tau = \pi_*^{-1}(\tau)$. We now have a remarkable collection of observations due to Behrstock and Leininger:

PROPOSITION 5.4. *The subcomplex $\mathcal{F}_\tau$*

- *is not R-dense, for any R,*
- *is dense in $\partial_\infty \mathcal{C}(S_{g,1})$, and*
- *is connected.*

Their original interest in $\mathcal{F}_\tau$ was to give a "natural" subcomplex of $\mathcal{C}(S)$ which is not quasi-convex: this is implied by the first two properties.

Proposition 5.4 and Proposition 5.1, together with a discussion of Dehn twists, will finish the proof of Theorem 5.8.

I take this opportunity to again thank Jason Behrstock and Chris Leininger for many interesting conversations and for showing me the proof of Proposition 5.4. I also thank Ken Bromberg for showing me his great simplification of my proof of Theorem 5.8. The shorter version is presented below.

2. Dead ends

Let $S = S_{g,b}$ be a non-pants surface. Fix $\omega \in \mathcal{C}^0(S)$. We now prove:

PROPOSITION 5.1. *The curve complex $\mathcal{C}(S)$ has no dead ends with respect to ω.*

PROOF. If $\mathcal{C}(S)$ is a copy of the Farey graph then the claim is trivial. Likewise when $S = S_{0,2}$. So suppose that S is non-sporadic, as well as non-simple.

In the first case we will suppose that $\alpha \in \mathcal{C}^0(S)$ is either non-separating or cuts off a pair of pants from S. Thus $S \smallsetminus \alpha$ has one component, X, which is not a pair of pants. In the second case we will suppose that α cuts S into a pair of surfaces X and Y, neither of which is simple.

Suppose we are in the first case. Thus $\operatorname{link}(\alpha) = B(\alpha, 1) \smallsetminus \alpha \cong \mathcal{C}(X)$ has infinite diameter in its intrinsic metric. (If X is sporadic then recall that $\mathcal{C}(X)$ is instead *defined* to be a copy of the Farey graph.) Set $n = d_S(\alpha, \omega)$. Choose some $\beta \in \mathcal{C}(X)$ so that $d_X(\beta, \omega) \geq 6n + 1$. It follows from Lemma 2.28 that the geodesic from ω to β misses X. But the only essential non-peripheral curve in the complement of X is α itself. Thus $d_S(\beta, \omega) = n + 1$ and α is not a dead end.

Suppose now we are in the second case. Thus $\operatorname{link}(\alpha) = B(\alpha, 1) \smallsetminus \alpha$ is the join of $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ and has diameter equal to two in both its intrinsic and extrinsic metrics. (Again, if X or Y is sporadic we replace the infinite disconnected set of vertices of $\mathcal{C}(X)$ or $\mathcal{C}(Y)$ by a copy of the Farey graph.) Set $n = d_S(\alpha, \omega)$. If ω is contained in X, say, the claim is trivial – simply choose $\beta \subset X$ to intersect ω. We may thus assume that both $\pi_X(\omega)$ and $\pi_Y(\omega)$ are nonempty. This is equivalent to saying that $n \geq 2$.

So choose some $\beta \in \mathcal{C}(X)$ so that $d_X(\beta, \omega) \geq 6n + 1$. We may assume that β is either nonseparating in S or cuts off a pair of pants from S. It again follows from Lemma 2.28 that the geodesic from ω to β misses X. If the geodesic goes thru α we are done, as before. So assume that the geodesic visits a curve γ in the subsurface Y. If $d_S(\gamma, \omega) < n-1$ we contradict the fact that $d_S(\alpha, \omega) = n$. If $d_S(\gamma, \omega) > n - 1$ then $d_S(\beta, \omega) > n$ and we are done. So assume that $d_S(\gamma, \omega) = n - 1$. It follows that $d_S(\beta, \omega) = n$. Finally, by the first case the curve β is not a dead end. We find a curve β' so that $d_S(\beta', \beta) = 1$ and $d_S(\beta', \omega) = n+1$. The path from α to β to β' shows that α is not a dead end and the proof is complete. $\square$

3. The Birman short exact sequence

Recall the *Birman short exact sequence*:

$$\pi_1(S_g, x_0) \to \mathcal{MCG}(S_{g,1}) \to \mathcal{MCG}(S_g)$$

for $g \geq 2$. Here we think of $S_{g,1}$ as being a copy of the closed genus g surface equipped with a basepoint x_0 which all curves avoid and which all isotopies fix. The inclusion is defined by sending $\gamma \in \pi_1(S_g, x_0)$ to the homeomorphism which drags x_0 along the path γ. The surjection is defined by forgetting the point x_0. See Birman's book [**3**] for details.

EXERCISE 5.2. If γ and δ are paths in $\pi_1(S_g, x_0)$ we write their composition by $\delta \circ \gamma$. This is the path obtained by first following γ and then δ. Check that the inclusion of the Birman short exact sequence is in fact a homomorphism. Can you write the image of γ in $\mathcal{MCG}(S_{g,1})$ as a composition of Dehn twists? (Hint: what about when γ is a simple closed curve?)

EXERCISE 5.3. Show that $\pi_1(S_g, x_0) \subset \mathcal{MCG}(S_{g,1})$ is of infinite index, normal, and finitely generated.

Corresponding to the Birman short exact sequence there is a "fibration" of curve complexes:

$$\mathcal{F}_\tau \to \mathcal{C}(S_{g,1}) \to \mathcal{C}(S_g).$$

Here the fibre map (on the right) is the map π_* which caps-off the boundary component with a disk (or, equivalently, forgets the marked point). The properties obtained in Exercise 5.3 become:

PROPOSITION 5.4 (Behrstock-Leininger). *The subcomplex $\mathcal{F}_\tau \subset \mathcal{C}(S_{g,1})$*

- *is not R-dense, for any R,*
- *is* dense *at infinity, and*
- *is connected.*

We first prove a lemma:

LEMMA 5.5. *The fibre map π_* is 1-Lipschitz*

PROOF. Every curve which is essential and non-peripheral in $S_{g,1}$ remains so in S_g. Also if $\alpha, \beta \subset S_{g,1}$ are disjoint then their images in S_g are either disjoint or are equal. Thus π_* does not increase distance. $\square$

Proof of Proposition 5.4. We take the assertions in turn. As $\mathcal{C}(S_g)$ has infinite diameter the first assertion follows from Lemma 5.5.

Pick now some point Σ in $\partial_\infty \mathcal{C}$ and choose a sequence of curves $\{\sigma_n\}$ converging to Σ. For convenience pick $\alpha \in \mathcal{F}_\tau$ as the basepoint for the Gromov product. Our goal is to produce a sequence $\alpha_n \in \mathcal{F}_\tau$ which also converges to Σ.

For each n choose a simple arc δ_n connecting the puncture x_0 to the curve σ_n. Let $g_n \in \mathcal{MCG}(S_{g,1})$, in the image of $\pi_1(S_g)$, be the result of dragging the puncture along δ_n, around σ_n, and back to x_0 along δ_n. (This map is isotopic to the identity when S has genus one.) See Figure 1.

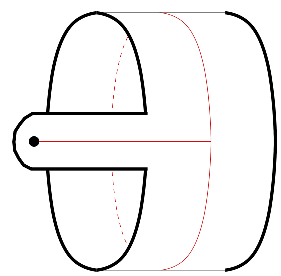

Figure 1. The dot is the puncture. The map g_n is isotopic to doing oppositely oriented Dehn twists on the two boundary components of the annulus.

Choose m to be a large multiple of twelve times the distance between α to σ_n. Take $\alpha_n = g_n^m(\alpha)$. Clearly $\alpha_n \in \mathcal{F}_\tau$. Take Y_n to be an annular neighborhood of σ_n. Then $d_{Y_n}(\alpha, \alpha_n) \geq 6 \cdot d_{S_{g,1}}(\alpha, \alpha_n)$ by the Dehn twist case of Theorem 2.24 and by the choice of m. Using Theorem 2.34 or the more elementary Lemma 2.28 it is an easy to show that the geodesic from α to α_n has a vertex which is disjoint from σ_n. It follows from the definition of the Gromov norm that α_n converges to Σ.

To show $\mathcal{F}_\tau$ is connected in $\mathcal{C}(S_{g,1})$ fix distinct curves α and β in $\mathcal{F}_\tau$. These are isotopic in S_g but not in $S_{g,1}$. We induct on the intersection number $\iota(\alpha, \beta)$, measured in $S_{g,1}$. Suppose the intersection number is zero. Then α and β are disjoint and we are done. Suppose that the intersection number is non-zero. There is a bigon $B \subset S_g \setminus (\alpha \cup \beta)$. It follows that the puncture resides in B, in the surface $S_{g,1}$. Now construct a curve β' by replacing the arc $\beta \cap B$ by the arc $\alpha \cap B$ and performing a small isotopy. Now $\beta' \in \mathcal{F}_\tau$ because β' is isotopic to β in S_g. Finally, $\iota(\alpha, \beta') \leq \iota(\alpha, \beta) - 2$. □

EXERCISE 5.6. Show that $\pi_1(S_g, x_0)$ acts transitively on $\mathcal{F}_\tau$.

QUESTION 5.7. The proof of Proposition 5.4 shows that all points of $\partial_\infty \mathcal{C}$ are tangential limit points of the fibre $\mathcal{F}_\tau$. Is it possible to describe the subset of the boundary which can be reached conically?

Recall our notation $\pi_*\colon \mathcal{C}(S_{g,1}) \to \mathcal{C}(S_g)$ which sends a curve in $S_{g,1}$ to its isotopy class in S_g. As above we use $\mathcal{F}_\tau = \pi_*^{-1}(\tau)$ to denote the fibre over τ. We may now prove:

THEOREM 5.8. *Fix $g \geq 2$. For any $\omega \in \mathcal{C}^0(S_{g,1})$ and for any $r \in \mathbb{N}$ the complex $\mathcal{C}(S_{g,1}) \smallsetminus B(\omega, r)$ is connected.*

PROOF. Choose α and β vertices of $\mathcal{C}(S_{g,1}) \smallsetminus B(\omega, r)$. By Proposition 4.11 the curve complex has no dead ends. So we may connect α and β, by paths disjoint from $B(\omega, r)$, to points outside of $B(\omega, 2r+2)$. Call these new points α' and β'.

Choose now a point $\tau \in \mathcal{C}(S_g)$ so that $d_{S_g}(\pi_*(\omega), \tau) > r$. Pick any point $\gamma \in \mathcal{F}_\tau$. As in the proof of the second property of Proposition 5.4 build a curve $\gamma' \in \mathcal{F}_\tau$: γ' which is the image of γ under the mapping class which drags the puncture x_0 many times around α'. As before we may assume that the geodesic, call it g, between γ and γ' contains a vertex α'' which is distance at most one away from α'.

We claim that at least one of the two segments in $g \smallsetminus \alpha''$ avoids the ball $B(\omega, r)$. For suppose not: Then there are vertices $\mu, \mu' \in g$ on opposite sides of α'' which both lie in $B(\omega, r)$. Thus $d_{S_{g,1}}(\mu, \mu') \leq 2r$. It follows that the length along g between μ and μ' is at most $2r$. So $d_{S_{g,1}}(\omega, \alpha') \leq 2r + 1$. This is a contradiction.

Thus we can connect α' to a vertex of $\mathcal{F}_\tau$ (namely, γ or γ') avoiding $B(\omega, r)$. Identically, we can connect β' to, say, $\delta \in \mathcal{F}_\tau$ while avoiding $B(\omega, r)$. Finally, by Proposition 5.4 the fibre $\mathcal{F}_\tau$ is connected. As $B(\omega, r) \cap \mathcal{F}_\tau = \emptyset$ this completes the proof. ∎

4. Difficulties finding ends of other curve complexes

We first note that no straight-forward extension of Theorem 5.8, to $S_{g,n}$ is possible. Although the Birman sequence remains the same:

$$1 \to \pi_1(S_{g,n}) \to \mathcal{MCG}(S_{g,n+1}) \to \mathcal{MCG}(S_{g,n}) \to 1$$

the map $\mathcal{C}(S_{g,n+1}) \to \mathcal{C}(S_{g,n})$ is no longer defined. Consider a loop in $S_{g,n+1}$) which encloses the "to-be-forgotten" puncture and exactly one other. This loop has no possible image in $\mathcal{C}(S_{g,n})$. Ignoring this problem, one might examine an orbit of $\Gamma = \pi_1(S_{g,n})$, as it acts on $\mathcal{C}(S_{g,n})$. The orbit in fact is dense at infinity and is again connected. The proof is identical to the proof of Proposition 5.4. However:

Exercise 5.9. Fix $\alpha \in \mathcal{C}(S_{g,n+1})$. Then $\Gamma \cdot \alpha$ is 3-dense in $\mathcal{C}$.

In a more general vein, we might consider any subgroup $H < \mathcal{MCG}(S)$ which is infinite index, normal, and finitely generated. Hopefully, H orbits will have the properties set out in Proposition 5.4. The Torelli group I_g springs to mind. Unfortunately, we have:

Lemma 5.10. *Any orbit of the Torelli group is 5-dense in* $\mathcal{C}(S)$.

Proof. Fix, once and for all, a standard homology basis $B = \{x_1, y_1, \ldots, x_g, y_g\}$. Let γ be an arbitrary curve. We will find an element of the Torelli group taking B to a curve system within distance four of γ.

Replace γ by a disjoint curve a_1 which is nonseparating. Extend a_1 to a homology basis $\{a_1, b_1, \ldots, a_g, b_g\}$. Express x_1 in terms of this basis: $x_1 = \sum(p_i a_i + q_i b_i)$. In the once-holed torus spanned by a_i and b_i take δ_i to be the curve of slope p_i/q_i. Band-sum the curves δ_i together to get a curve x_1' which is homologous to x_1. We note that $d_{\mathcal{C}}(\gamma, x_1') \leq 4$.

We must now produce a curve y_1' which meets x_1' exactly once and which is homologous to y_1. Consider the original curve y_1. Consider the intersection points $x_1' \cap y_1$ and notice that the algebraic intersection is exactly one. Order the intersection points $\{c_j\}$ using the orientation on x_1'. If there is only one such point, take $y_1' = y_1$. If there are many, then there is a pair of points c_j and c_{j+1} of opposite sign. Let δ be the subarc of x_1' with $\partial\delta = \{c_j\} \cup \{c_{j+1}\}$ so that the interior of δ is disjoint from y_1. Surger y_1 along δ. (See Figure 2.) This produces an oriented multi-curve which, in homology, sums to $[y_1]$.

Figure 2.

We repeatedly surger the multicurve. Every time we surger we produce a new oriented multi-curve which intersects x_1' in two fewer points and which sums, in homology, to $[y_1]$. This procedure halts when the number of intersection points falls to one. Now band sum the resulting multi-curve in the complement of x_1'. Note that the component meeting x_1' is non-separating, so it is possible to perform the band sum while preserving orientations. The resulting single curve is the desired y_1'.

In essentially identical fashion we may surger and band-sum the curves x_i and y_i, $i > 1$, to produce a standard homology basis $B' = \{x_i', y_i'\}$ where $x_i' \in [x_i]$ and $y_i' \in [y_i]$. It is easily checked that any mapping class taking the basis B to the basis B' lies in the Torelli group, and this completes the proof. □

5. A trip to the zoo: the sphere complex

The *double*, $\mathrm{Doub}(N)$ of a manifold N is obtained by taking two copies of N, say $N\times\{0\}$ and $N\times\{1\}$, and gluing them via the identity map on ∂N.

EXERCISE 5.11. Check that $\mathrm{Doub}(S_{g,b}) \cong S_{2g+k-1}$. Verify that the double of a handlebody $\mathrm{Doub}(V_g)$ is homeomorphic to M_g, the connect sum of g copies of $S^2\times S^1$.

Similarly, a properly embedded submanifold $F \subset N$ gives rise to an embedded submanifold $\mathrm{Doub}(F) \subset \mathrm{Doub}(N)$. For example, if E is a disk in V_g then its double is a sphere in M_g.

As usual we have a related complex. Define the *sphere complex* $\mathcal{S}(M_g)$ as follows: vertices are essential two-spheres in M_g. A collection of $k + 1$ vertices span a k-simplex if all $k + 1$ of the spheres can be isotoped to be disjoint.

We have a natural map $\mathrm{Doub}(\cdot)\colon \mathcal{D}(V_g) \to \mathcal{S}(M_g)$ taking disks to spheres.

EXERCISE 5.12. Suppose that $g \geq 2$. Before reading on show that this map is not one-to-one.

EXERCISE 5.13. Suppose that $g \geq 2$. Check that this map is onto. (This is much harder.)

There are several ways to understand the fibre of the map $\mathcal{D}(V_g) \to \mathcal{S}(M_g)$. The most obvious would be to consider the corresponding mapping class groups: $\mathrm{Doub}(\cdot)\colon \mathcal{MCG}(V_g) \to \mathcal{MCG}(M_g)$. Just as we defined a Dehn twist, τ_α, on a curve α there is a notion of a Dehn twist on a disk $D \subset V$ or a sphere $S \subset M$. Naturally enough, we find that the double of a disk twist τ_D gives a sphere twist about $\mathrm{Doub}(D)$. Note that, for any essential sphere $S \subset M$ the map τ_S^2 is trivial. That is, twisting a sphere twice gives a map isotopic to the identity. This is the famous *plate trick*.

In fact, both disk and sphere twists act trivially on the fundamental group of the underlying manifold. Therefore it algebraically nicer to consider the homomorphism $\mathcal{MCG}(V_g) \to \mathrm{Out}(F_g)$. The kernel is generated by disk twists and accordingly it is called the *twist group*:

$$\mathrm{Twist}_g \to \mathcal{MCG}(V_g) \to \mathrm{Out}(F_g).$$

There is a marked resemblance to the Birman short exact sequence. As expected there is a fibration:

$$\mathcal{F}_S \to \mathcal{D}(V_g) \to \mathcal{S}(M_g).$$

However, much less is know about the fibre of this map.

QUESTION 5.14. Is $\mathcal{F}_S$ equal to an orbit of Twist_g acting $\mathcal{D}(V)$? Is $\mathcal{F}_S$ coarsely connected? Is the complement of a ball in $\mathcal{D}(V)$ still connected? Does this give information about $\mathcal{C}(S_g)$?

APPENDIX A

Hints for some exercises

This chapter is under construction.

HINT 1.3. The only orientable simple surfaces are the annulus and the pants.

HINT 1.4. The sporadic surfaces are the closed and once holed tori, as well as the four holed sphere.

HINT 1.5. Use the classification of surfaces. What do you get when you cut S along α?

HINT 1.6. Take a regular $4g$-gon and identify opposite sides. This gives a genus g surface with a rotation symmetry of order $4g$. This symmetry cannot be obtained from the motion of a graph in $\mathbb{R}^3$.

HINT 1.14. First show that if a mapping class f fixes, up to isotopy, a collection of curves (or arcs) cutting S into a bunch of disks, then f is the identity element of $\mathcal{MCG}(S)$. (You will need the Alexander Trick: any homeomorphism of a disk which fixes the boundary pointwise is isotopic to the identity, via an isotopy fixing the boundary pointwise.)
Now notice that to answer the question it is enough to examine a regular neighborhood of $\alpha \cup \beta$, which is a once holed torus.

HINT 1.15. There are several ways to do this – for example lift everything to the universal cover and "straighten".

HINT 1.16. See the hint for Exercise 1.6.

HINT 1.26. Perhaps a partial mapping will be useful.

HINT 1.27. For curves at distance three, it may help to notice that two essential arcs can fill the surface $S_{1,1}$ and that you can glue together two copies of $S_{1,1}$ to obtain S_2.

Bibliography

[1] J. M. Alonso and et al. Notes on word hyperbolic groups. In *Group theory from a geometrical viewpoint (Trieste, 1990)*, pages 3–63. World Sci. Publishing, River Edge, NJ, 1991. Edited by H. Short.

[2] Jason Behrstock. *Asymptotic geometry of the mapping class group and Teichmüller space*. PhD thesis, SUNY Stony Brook, 2004. http://www.math.columbia.edu/∼jason/thesis.pdf.

[3] Joan S. Birman. *Braids, links, and mapping class groups*. Princeton University Press, Princeton, N.J., 1974. Annals of Mathematics Studies, No. 82.

[4] Brian H. Bowditch. Intersection numbers and the hyperbolicity of the curve complex. http://www.maths.soton.ac.uk/staff/Bowditch/bhb-curvecomplex.pdf.

[5] Martin R. Bridson. Geodesics and curvature in metric simplicial complexes. In *Group theory from a geometrical viewpoint (Trieste, 1990)*, pages 373–463. World Sci. Publishing, River Edge, NJ, 1991.

[6] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*. Springer-Verlag, Berlin, 1999.

[7] Jeffrey Brock and Benson Farb. Curvature and rank of Teichmüller space. arXiv:math.GT/0109045.

[8] Andrew J. Casson and Steven A. Bleiler. *Automorphisms of surfaces after Nielsen and Thurston*. Cambridge University Press, Cambridge, 1988.

[9] Andrew J. Casson and Cameron McA. Gordon. Reducing Heegaard splittings. *Topology Appl.*, 27(3):275–283, 1987.

[10] M. Coornaert, T. Delzant, and A. Papadopoulos. *Géométrie et théorie des groupes*. Springer-Verlag, Berlin, 1990. Les groupes hyperboliques de Gromov.

[11] Max Dehn. *Papers on group theory and topology*. Springer-Verlag, New York, 1987. Translated from the German and with introductions and an appendix by John Stillwell, With an appendix by Otto Schreier.

[12] Fathi, Laudenbach, and Poénaru, editors. *Travaux de Thurston sur les surfaces*. Société Mathématique de France, Paris, 1991. Séminaire Orsay, Reprint of *Travaux de Thurston sur les surfaces*, Soc. Math. France, Paris, 1979 [MR 82m:57003], Astérisque No. 66-67 (1991).

[13] É. Ghys and P. de la Harpe, editors. *Sur les groupes hyperboliques d'après Mikhael Gromov*, volume 83 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1990. Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988.

[14] Mikhael Gromov. Hyperbolic groups. In *Essays in group theory*, pages 75–263. Springer, New York, 1987.

[15] John L. Harer. Stability of the homology of the mapping class groups of orientable surfaces. *Ann. of Math. (2)*, 121(2):215–249, 1985.

[16] John L. Harer. The virtual cohomological dimension of the mapping class group of an orientable surface. *Invent. Math.*, 84(1):157–176, 1986.

[17] Willam J. Harvey. Boundary structure of the modular group. In *Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978)*, pages 245–251, Princeton, N.J., 1981. Princeton Univ. Press.

[18] A. Hatcher and W. Thurston. A presentation for the mapping class group of a closed orientable surface. *Topology*, 19(3):221–237, 1980.

[19] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[20] John Hempel. 3-manifolds as viewed from the curve complex. *Topology*, 40(3):631–657, 2001. arXiv:math.GT/9712220.

[21] G. Christopher Hruska. Nonpositively curved 2-complexes with isolated flats. *Geom. Topol.*, 8:205–275 (electronic), 2004.

[22] Nikolai V. Ivanov. *Subgroups of Teichmüller modular groups*, volume 115 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1992. Translated from the Russian by E. J. F. Primrose and revised by the author.

[23] Nikolai V. Ivanov. Mapping class groups. In *Handbook of geometric topology*, pages 523–633. North-Holland, Amsterdam, 2002.

[24] Michael Kapovich. *Hyperbolic manifolds and discrete groups*. Birkhäuser Boston Inc., Boston, MA, 2001.

[25] Erica Klarreich. The boundary at infinity of the curve complex and the relative Teichmüller space. http://nasw.org/users/klarreich/research.htm.

[26] Tsuyoshi Kobayashi. Heights of simple loops and pseudo-Anosov homeomorphisms. In *Braids (Santa Cruz, CA, 1986)*, pages 327–338. Amer. Math. Soc., Providence, RI, 1988.

[27] W. B. R. Lickorish. A representation of orientable combinatorial 3-manifolds. *Ann. of Math. (2)*, 76:531–540, 1962.

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

[29] Howard A. Masur and Yair N. Minsky. Quasiconvexity in the curve complex. arXiv:math.GT/0307083.

[30] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.*, 138(1):103–149, 1999. arXiv:math.GT/9804098.

[31] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000. arXiv:math.GT/9807150.

[32] Howard A. Masur and Saul Schleimer. The pants complex has one end. arXiv:math.GT/0312385.

[33] Darryl McCullough. Virtually geometrically finite mapping class groups of 3-manifolds. *J. Differential Geom.*, 33(1):1–65, 1991.

[34] Ulrich Oertel. Automorphisms of three-dimensional handlebodies. *Topology*, 41(2):363–410, 2002.

[35] Dale Rolfsen. *Knots and links*. Publish or Perish Inc., Houston, TX, 1990. Corrected revision of the 1976 original.

[36] Martin Scharlemann. Automorphisms of the 3-sphere that preserve a genus two Heegaard splitting. UCSB Math 2003-14. arXiv:math.GT/0307231.

[37] Martin Scharlemann. Heegaard splittings of compact 3-manifolds. In *Handbook of geometric topology*, pages 921–953. North-Holland, Amsterdam, 2002. arXiv:math.GT/0007144.

[38] Kenneth J. Shackleton. Tightness and computing distances in the curve complex. arXiv:math.GT/0412078.

[39] Abigail Thompson. Thin position and bridge number for knots in the 3-sphere. *Topology*, 36(2):505–507, 1997.

[40] William P. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Bull. Amer. Math. Soc. (N.S.)*, 19(2):417–431, 1988.

[41] Bronisław Wajnryb. Mapping class group of a handlebody. *Fund. Math.*, 158(3):195–228, 1998.