CURVE COMPLEXES OF NON-ORIENTABLE SURFACES

ALEX PIELOCH

Abstract. We explore non-orientable surfaces and their associated curve complexes. Studying the combinatorics modeled by the curve complex of a surface helps elucidate the algebraic properties of the mapping class group of the surface. We begin by studying geometric properties of the curve complexes of non-orientable surfaces and the geometric properties of natural sub-complexes of the curve complex. Finally, we prove that the curve complex of a non-orientable surface is homotopy equivalent to a wedge of spheres of possibly different dimensions.

Contents

1. Preliminaries 1
2. Geometry of the Curve Complex 3
3. Homotopy Type of the Curve Complex 13
Acknowledgments 16
References 16

1. Preliminaries

In this section, we give some necessary background knowledge and notation used in this paper. For a more in-depth treatment of the ideas in this section, see [1].

In this paper, we will use $F$ to denote a surface that may be either orientable or non-orientable. We will use $S$ to denote an orientable surface. And we will use $N$ to denote a non-orientable surface. By the Classification of Surfaces Theorem [5], every connected, orientable, compact surface can be identified by three different topological invariants: the genus $g$, the number of boundary components $s$, and the number of marked points $n$. We will denote a specific orientable surface of genus $g$ with $s$ boundary components and $n$ punctures by $S_{g,s}^n$. By the Classification of Surfaces Theorem [5], every connected, non-orientable, compact surface can be identified by three different topological invariants: the number of cross caps $c$, the number of boundary components $s$, and the number of marked points $n$. Recall, a cross cap in a surface $F$ is a two-dimensional surface in 3-space that is one-sided and the continuous image of a Möbius strip that intersects itself in an interval. Equivalently, a cross cap in a surface $F$ is the complement of an open disk in the surface with the antipodal points on its boundary identified. We will denote a specific non-orientable surface with $c$ cross caps, $s$ boundary components, and $n$ punctures by $N_{c,s}^n$. Notice that marked points in a surface can equivalently be...
viewed as *punctures* in a surface. In light of this, we will freely switch between these two formulations.

The primary objects that we are interested in studying are the isotopy classes of simple closed curves in a surface. We say that a simple closed curve is *essential* if it is not isotopic to a curve in a regular neighborhood of a point, puncture, or boundary component. A curve which is not essential is said to be *non-essential*. A curve $\mu$ in a surface $F$ is called a *Möbius curve* if one component of $F - \mu$ is a Möbius band. This brings us to the definition of the curve complex of a surface.

**Definition 1.1.** The *curve complex* associated to a surface $F$, denoted $\mathcal{C}(F)$, is the simplicial complex whose vertices are the isotopy classes of essential simple closed curves in $F$, excluding Möbius curves. A set of $k + 1$ vertices $\{v_0, \ldots, v_k\}$ defines a $k$-simplex if the geometric intersection number $i(v_i, v_j)$ is zero for all $i$ and $j$.

Notice that $\mathcal{C}(F)$ is a flag complex and thus the combinatorial and geometric information is encoded completely in its 1-skeleton. Moreover, from the point of view of the curve complex, boundary components and punctures are the same, replacing one by the other does not change the isomorphism type of the curve complex. In the case of a non-orientable surface, the lack of orientation gives rise to a notion of orientable and non-orientable curves.

**Definition 1.2.** Let $\alpha$ be a simple closed curve in a surface $F$.

- We say that $\alpha$ is *one-sided* if a regular neighborhood of $\alpha$ is topologically a Möbius band.
- We say that $\alpha$ is *two-sided* if a regular neighborhood of $\alpha$ is topologically an annulus.

Definition 1.2 motivates the definitions of the following sub-complexes of the curve complex of a surface.

**Definition 1.3.** Let $\mathcal{C}_1(F)$ denote the full sub-complex of $\mathcal{C}(F)$ spanned by vertices given by one-sided curves. Let $\mathcal{C}_2(F)$ denote the full sub-complex of $\mathcal{C}(F)$ spanned by vertices given by two-sided curves.

When $F$ is an orientable surface, we have that $\mathcal{C}_2(F)$ is equal to $\mathcal{C}(F)$. We will also consider the following sub-complex.

**Definition 1.4.** Let $\mathcal{C}_{NS}(F)$ denote the full sub-complex of $\mathcal{C}(F)$ spanned by vertices given by curves $\alpha$ such that $F - \alpha$ is connected.

The non-separating curve complex $\mathcal{C}_{NS}(F)$ is also commonly denoted by $\mathcal{N}(F)$ and $\text{NonSep}(F)$.

The remainder of this paper is organized as follows. In Section 2, we study the geometric properties of the curve complex and its natural sub-complexes. Specifically, we will study the connectivity of $\mathcal{C}_1(F)$, $\mathcal{C}_2(F)$, and $\mathcal{C}_{NS}(F)$, investigate if any of these sub-complexes are quasi-isometric to the full complex, and look at the structure of maximal simplices. In Section 3, we study the homotopy type of the curve complex of a surface $F$. In the case of $F$ orientable, Harer showed that $\mathcal{C}(S^n_{g,s})$ is homotopy equivalent to a wedge of spheres of dimension $2g + s + n - 3$, barring some exceptional cases [2]. We generalize Harer’s argument to non-orientable surfaces and show that $\mathcal{C}(N^n_{c,s})$ is homotopy equivalent to a wedge of spheres each with dimension fewer than or equal to $c + s + n - 3$, barring some exceptional cases.
2. Geometry of the Curve Complex

We begin this section by showing that $C(N_{c,s}^n)$ is connected. This proposition is left as an exercise in Schleimer’s notes on the curve complex, see [9]. Our proof is a generalization of the proof for the case of orientable surfaces, which is originally due to Lickorish [4].

**Proposition 2.1.** Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. If $\frac{3c}{2} + s \geq 5$, then for all $\alpha$ and $\beta$ in $C(N)$ the following inequality holds

$$d(\alpha, \beta) \leq 2 \log_2(i(\alpha, \beta)) + 2.$$  

*Proof.* To establish the result, we use a curve surgery argument. We will first show that the result holds for any $\alpha$ and $\beta$ in $C_2(N) \cup M$, where $M$ denotes the collection of Möbius curves in $N$. We will induct on the geometric intersection number of $\alpha$ and $\beta$.

First, suppose that $i(\alpha, \beta) = 1$. Then a regular neighborhood of $\alpha \cup \beta$ is homeomorphic to a punctured torus, say $Y$. If $N - Y$ is empty or contractible, then $N$ is topologically a punctured torus or a torus, which contradicts our assumption of $\frac{3c}{2} + s \geq 5$. Selecting the boundary curve $\gamma$ of $Y$, we have that

$$i(\alpha, \gamma) = 0 = i(\gamma, \beta).$$

It follows that

$$d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta) = 2 \leq 2 \log_2(i(\alpha, \beta)) + 2.$$  

Next, suppose that $i(\alpha, \beta) = 2$. Notice that $\alpha$ must consecutively intersect $\beta$ with one of two different orientations as shown in the left two configurations in Figure 1.

![Figure 1. Curve surgeries used in the proof of Proposition 2.1.](image)

Let's first consider the configuration on the far left of Figure 1. Notice that $\gamma_1$ and $\gamma_2$ are both disjoint from $\alpha$ and $\beta$. If both $\gamma_1$ and $\gamma_2$ are non-essential, then we may consider the pair of curves in the top and bottom of the configuration that are disjoint from $\gamma_1$ and $\gamma_2$. If both of these curves are also non-essential, then $N$ is topologically a four times punctured sphere, which contradicts our assumption of $N$. Hence, at least one of these curves is essential, say $\gamma$, and yields

$$i(\alpha, \gamma) = 0 = i(\gamma, \beta).$$

It follows that

$$d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta) = 2 \leq 2 \log_2(i(\alpha, \beta)) + 2.$$
Now consider the configuration in the center of Figure 1. Since
\[ i(\alpha, \gamma_1) = 1 = i(\alpha, \gamma_2) \]
we know \( \gamma_1 \) is essential. If \( \gamma_1 \) was non-essential, then we could resolve the intersection with \( \alpha \). However, we can only resolve trivial intersections in pairs of two since a trivial intersection always forms a bigon [1]. Consequently, we can see that no such bigon is formed and \( \gamma_1 \) is essential. Moreover, both \( \gamma_1 \) and \( \gamma_2 \) intersect \( \beta \) strictly fewer than two times. We select one of these curves to be \( \gamma \). It follows that
\[ d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta) = 4 \leq 2 \log_2(i(\alpha, \beta)) + 2. \]

Inductively, suppose that \( i(\alpha, \beta) > 2 \). We need to handle two different cases. First, suppose that \( \alpha \) has two consecutive points of intersection along \( \beta \) oriented as shown in the center configuration of Figure 1. Since
\[ i(\alpha, \gamma_1) = 1 = i(\alpha, \gamma_2) \]
we know \( \gamma_i \) is essential. Moreover, at least one of \( \gamma_1 \) or \( \gamma_2 \) intersects \( \beta \) fewer than or equal to \( n/2 \) times. We select said curve to be \( \gamma \). If \( \alpha \) does not have such consecutive points of intersection along \( \beta \), then we must have that \( \alpha \) intersects \( \beta \) as shown in the right configuration of Figure 1. Up to changing the orientation of \( \alpha \), we may assume that \( \alpha \) intersects \( \beta \) at \( x_1, x_2, x_3 \) in that respective order. Notice that a regular neighborhood of \( \alpha \cup \gamma_1 \cup \gamma_2 \) is topologically a four times punctured sphere. To see this, note that \( \alpha \cup \gamma_1 \cup \gamma_2 \) is isotopically two cords glued onto \( \alpha \) such that the cords share a common point of intersection with \( \alpha \). Hence, \( \alpha \cup \gamma_1 \cup \gamma_2 \) is isotopically a wedge of three circles and its regular neighborhood is a four times punctures sphere. Let’s denote this four times punctured sphere by \( Y \). If \( N - Y \) is topologically a disjoint collection of disks, then we have that \( c = 0 \) and \( s \leq 4 \), which contradicts our assumption of \( N \). Selecting some boundary curve of \( Y \), which is disjoint from \( \alpha, \gamma_1, \) and \( \gamma_2 \) and essential in \( N \), we have that
\[ d(\alpha, \gamma_1) = 2 = d(\alpha, \gamma_2). \]
Moreover, at least one of \( \gamma_1 \) or \( \gamma_2 \) intersects \( \beta \) fewer than or equal to \( n/2 \) times. We select said curve to be \( \gamma \). Using our selected \( \gamma \), we have
\[ d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\beta, \gamma) \]
\[ \leq 2 + 2 \log_2(i(\beta, \gamma)) + 2 \]
\[ \leq 2 + 2 \log_2(i(\alpha, \beta)/2) + 2 \]
\[ = 2 \log_2(i(\alpha, \beta)) + 2. \]

To complete the proof, notice that each \( \gamma \) we constructed was two-sided as an orientation of each regular neighborhood is obtained from the orientations of the regular neighborhoods of \( \alpha \) and \( \beta \). However, \( \gamma \) may have been a Möbius curve. In this case, we may replace \( \gamma \) by the core circle \( \gamma' \) of the Möbius band that \( \gamma \) bounds. This is because \( i(\alpha, \gamma) = 0 \) if and only if \( i(\alpha, \gamma') = 0 \) for all \( \alpha \). After this replacement, the result follows. \( \square \)

**Corollary 2.2.** Let \( N \) be a non-orientable surface with \( c \) cross caps and \( s \) boundary components. If \( \frac{5c}{2} + s \geq 5 \), then \( C(N) \) is connected.

Now we turn our attention to the connectivity properties of \( C_1(N) \) and \( C_2(N) \). Then connectivity of \( C_2(N) \) follows quickly from Corollary 2.2. If two vertices of
... connected by a path that contains a one-sided curve, then we can push this path off of the one-sided curve into $C_2(N)$.

**Proposition 2.3.** Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. If $c + s \geq 5$, then $C_2(N)$ is connected.

**Proof.** Fix some $\alpha$ and $\beta$ in $C_2(N)$. Suppose that $P = (\alpha = \gamma_0, \ldots, \gamma_i, \ldots, \gamma_n = \beta) \subset C(N)$ is a path from $\alpha$ to $\beta$ in $C(N)$. To show that $C_2(N)$ is connected, we push this path in $C(N)$ off any one-sided curves, producing a path from $\alpha$ to $\beta$ in $C_2(N)$. Suppose that $\gamma_i$ is a one-sided curve. Consider the surface $X = N - \gamma_i$. If $X$ is orientable, then $X$ is homeomorphic to $S^{c-1}_{s+1}$, which is connected when $c + s \geq 5$. By the connectivity of $C(X)$ there exists a path $P'$ in $C(X)$ from $\gamma_{i-1}$ to $\gamma_{i+1}$. This path $P'$ from $\gamma_{i-1}$ to $\gamma_{i+1}$ in $C(X)$ lies in $C(N)$ by the injectivity of $\pi_1(X)$ into $\pi_1(N)$. When $X$ is orientable, $P'$ is composed entirely of two-sided curves and we have resolved the issue caused by $\gamma_i$. If $X$ is non-orientable, then $X$ is homeomorphic to $N^{c-1}_{s+1}$, which is connected when $c + s \geq 5$. The same argument above gives a path $P'$ from $\gamma_{i-1}$ to $\gamma_{i+1}$ in both $C_1(X)$ and $C_2(N)$. However, $P'$ may contain one-sided curves. Since $N$ has a finite number of disjoint one-sided curves, we may repeatedly apply this argument to $X$ and its subsurfaces until we obtain an orientable surface and a path composed of two-sided curves. Repeatedly applying this argument to each element of $P$ that is one-sided, we obtain the desired result. □

Unlike the connectivity of $C_2(N)$, the connectivity of $C_1(N)$ depends on the number of cross caps in $N$. If the number of cross caps is even, then $C_1(N)$ is connected. If the number of cross caps is odd, then $C_1(N)$ is not connected. The proof of the latter statement is simple.

**Proposition 2.4.** Let $N$ be a non-orientable surface with an odd number of cross caps. Then $C_1(N)$ is not connected.

**Proof.** Consider the one-sided curve $v$ show in Figure 2. Note that $N - v$ is orientable and thus contains no one-sided curves. Hence, if $\alpha$ is in $C_1(N)$, then we must have that $i(v, \alpha) \neq 0$. It follows that $v$ is disjoint from every other face of $C_1(N)$. Consequently $C_1(N)$ is not connected. □

![Figure 2](image.png)  
**Figure 2.** Isolated curve used for Proposition 2.4. Note, the barred circle in the figure represents a cross cap in $N$. That is, antipodal points are identified on the circle in $N$.

To prove the former statement, we need to introduce some necessary machinery. The first result that we will need is a lemma due to Putman [8].
Lemma 2.5 (Putman). Consider a group $G$ acting upon a simplicial complex $X$. Fix a vertex $v \in X$ and a set $S$ of generators for $G$. Assume the following hold.

1. For all vertices $v' \in X$, the orbit $G \cdot v$ intersects the connected component of $X$ containing $v'$.
2. For all $s \in S$, there is some path $P_s$ in $X$ from $v$ to $s \cdot v$.

Then $X$ is connected.

For our purposes, our simplicial complex will be $C_1(N)$ and our group will be the mapping class group of $N$, $\text{Mod}(N)$.

1. With this in mind, we need a generating set for $\text{Mod}(N)$. Consider the collections of curves shown in Figure 4. Given $\alpha$ in a surface $F$, we denote the positive Dehn Twist about $\alpha$ by $T_\alpha$. Let $w_i$ denote the push map of $z_i$ along $\alpha_i$. Let $\sigma_{s+1}$ denote the elementary braid on $N$. Given a curve $\alpha$ that intersects $z_i$ followed by $z_n$ as shown in Figure 3, $\sigma_{s+1}(\alpha)$ is defined as shown in Figure 3.

![Figure 3. Mapping given by the braid element $\sigma_{s+1}$.](image)

Let $y$ denote the cross cap slide performed on the component of the compliment of $\xi$ that is a Klein bottle with boundary.

2. We have the following result due to Stukow [10].

Theorem 2.6 (Stukow). Let $N$ be a non-orientable surface with $c$ cross caps and $n$ punctures. If $c \geq 3$, $c$ is even, and $n \geq 2$, then the mapping class group $\text{Mod}(N^n_{g,s})$ is generated by

$$\{T_\ell, w_{s+1}, \sigma_{s+1}, \ldots, \sigma_{s+n-1}, y, T_\lambda \mid \ell \in D\}.$$
Figure 4. Curves used to generate $\text{Mod}(N)$ as in Theorem 2.6. Let the collection of curves on the top surface be denoted by $\mathcal{D}$.

**Proposition 2.7.** Let $N$ be a non-orientable surface with an even number of cross caps $c$ and $n$ punctures. If $c \geq 3$ and $n \geq 2$, then $\mathcal{C}_1(N)$ is connected.

**Proof.** To show this result, we use Lemma 2.5. Consider the basepoint $v$ indicated in Figure 5. Using a Classification of Surfaces argument, one can show that there exists a unique orbit of one-sided curves under the action of $\text{Mod}(N)$. Hence, condition (1) of Lemma 2.5 is clearly satisfied. Next, it's clear that the only generators that may act non-trivially on $v$ are $T_{c_r}$ and $T_{b_{r+1}}$. However, both of these Dehn Twists acting on $v$ produce curves that are disjoint from $v$. Hence, condition (2) of Lemma 2.5 is satisfied. By Lemma 2.5, we have that $\mathcal{C}_1(N)$ is connected. □
Proposition 2.4 motives the following definition.

**Definition 2.8.** Let $C'_1(F)$ denote the abstract simplicial complex with the same set of vertices as $C_1(F)$, but $(v_0, \ldots, v_k)$ is a $k$-simplex in $C'_1(F)$ if $i(v_i, v_j) \leq 1$ for all $i$ and $j$.

**Proposition 2.9.** Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. If $\frac{3c}{2} + s \geq 5$, then $C'_1(N)$ is connected.

*Proof.* Let $\alpha$ and $\beta$ be two one-sided curves in $N$ that are in minimal position and consider Figure 6. Inductively, suppose that $i(\alpha, \beta) = n > 1$ and any curves with intersection number fewer than $n$ are connected in $C'_1(N)$.

![Figure 6. Curve surgeries used in the proof of Proposition 2.9.](image)

Let’s first consider the configuration on the left of Figure 6. Notice that one of the two dotted curves is one-sided in the correct position with respect to $\alpha$. Call said curve $\gamma$. Notice that $i(\alpha, \gamma) = 1$ and $\gamma$ intersects $\beta$ fewer than $i(\alpha, \beta)$ times. After we note that $\gamma$ is essential, we are done by induction.

Now consider the configuration on the right of Figure 6. Again, one of the two dotted curves is one-sided. Notice that if $i(\alpha, \beta) = 2$, then $i(\beta, \gamma) = 1$ and consequently $\gamma$ is essential. If $i(\alpha, \beta) > 2$ and $\gamma$ is nonessential, then there must be a bigon formed between $\alpha$ and $\beta$, which contradicts our assumption of $\alpha$ and $\beta$ being in minimal position. Moreover, $\gamma$ intersects $\beta$ strictly fewer than $i(\alpha, \beta)$ times. Inductively, we are done. 

Now we turn our attention to the connectivity of $C_{NS}(N)$. In the spirit of the proof of Proposition 2.3, if two non-separating curves are connected by a path that contains a separating curve, then we push our path off of this separating curve.

**Proposition 2.10.** Let $N$ be a non-orientable surface with $c$ cross caps and $s \leq 1$ boundary components. If $\frac{3c}{2} + s \geq 5$, then $C_{NS}(N)$ is connected.

*Proof.* Let $\alpha, \beta \in C_{NS}(N)$. Suppose that $P = (\alpha = \gamma_0, \ldots, \gamma_n = \beta) \subset C(N)$
is a path from \( \alpha \) to \( \beta \). We show that \( C_{NS}(N) \) is connected by pushing this path off any separating curves. Suppose that \( \gamma_i \) is a separating curve. Then \( N - \gamma_i \) has two components, which we denote by \( X_1 \) and \( X_2 \). If \( \gamma_{i-1} \) is in \( X_1 \) and \( \gamma_{i+1} \) is in \( X_2 \), then \( i(\gamma_{i-1}, \gamma_{i+1}) = 0 \) and we may remove \( \gamma_i \) from \( \mathcal{P} \). If \( \gamma_{i-1} \) and \( \gamma_{i+1} \) are both in \( X_1 \), then by our assumption of the complexity of \( N \), we may find a nonseparating curve \( \gamma' \) in \( X_2 \). To obtain our modified path, we simply replace \( \gamma_i \) with \( \gamma' \). Repetitively applying this argument to each separating curve in \( \mathcal{P} \), we obtain the desired result. \( \square \)

A natural question to ask about these various sub-complexes is whether or not they are quasi-isometric to the full complex. Let’s recall the definition of a quasi-isometry.

**Definition 2.11.** A function \( f : Y \to X \) is a \((K,E)\) quasi-isometric embedding for \( K \geq 1, E \geq 0 \) if, for every \( x, y \in Y \), we have

\[
\frac{1}{K}(dy(x,y) - E) \leq d_X(f(x),f(y)) \leq K \cdot dy(x,y) + E.
\]

If \( f \) is \( E \)-dense (an \( E \) neighborhood of \( f(Y) \) equals all of \( X \)) then we say that \( f \) is a quasi-isometry and that \( X \) is quasi-isometric to \( Y \).

We begin by showing that \( C_{NS}(N) \) is quasi-isometric to \( C(N) \). To see this, we first need to notice the following corollary of Proposition 2.10.

**Corollary 2.12.** Let \( N \) be a non-orientable surface with \( c \) cross caps and \( s \leq 1 \) boundary components. If \( \frac{c}{2} + s \geq 5 \), then for all \( \alpha, \beta \in C_{NS}(N) \), we have

\[
d_{C_{NS}}(\alpha, \beta) = d_C(\alpha, \beta).
\]

**Proof.** This follows from the proof of Proposition 2.10, where we show that the shortest path between \( \alpha \) and \( \beta \) can always be realized in \( C_{NS}(N) \). \( \square \)

**Proposition 2.13.** Let \( N \) be a non-orientable surface with \( c \) cross caps and \( s \leq 1 \) boundary components. If \( \frac{c}{2} + s \geq 5 \), then \( C_{NS}(N) \) is quasi-isometric to \( C(N) \).

**Proof.** By Corollary 2.12, we must have that the natural inclusion of \( C_{NS}(N) \) into \( C(N) \) is a quasi-isometric embedding. To see that this inclusion is a quasi-isometry, it remains to show that \( C_{NS} \) is dense in \( C(N) \). Let \( \gamma \) be in \( C(N) \). If \( \gamma \) is in \( C_{NS}(N) \), then we are done. If \( \gamma \) is a separating curve, then consider a component \( X \) of \( N - \gamma \). Notice that \( \pi_1(X) \) injects into \( \pi_1(N) \). Hence, selecting some element \( \alpha \) in \( C_{NS}(X) \) gives us an element of \( C_{NS}(N) \) such that \( d(\alpha, \gamma) = 1 \). It follows that \( C_{NS} \) is 1-dense in \( C(N) \) and the result follows. \( \square \)

Below we will show that neither \( C_1(N) \) nor \( C_2(N) \) is quasi-isometric to \( C(N) \). But before doing so, we need to introduce some terminology and develop some machinery.

**Definition 2.14.** A surface \( F \) is called sporadic if \( C(F) \) is not connected. Similarly, a surface \( F \) is called non-sporadic if \( C(F) \) is connected.

We will need the following result due to Masur and Minsky [7].

**Theorem 2.15** (Masur and Minsky). For a non-sporadic orientable surface \( S \) there exists \( c > 0 \) such that, for any pseudo-Anosov \( h \in \text{Mod}(S) \), any \( \gamma \in C(S) \) and any \( n \in \mathbb{Z} \),

\[
d(h^n(\gamma), \gamma) \geq c|n|.
\]
It follows immediately from Theorem 2.15 that $\mathcal{C}(S)$ has infinite diameter. Recall, that every non-orientable surface has an orientation double cover that is an orientable surface with the deck transformation given by the antipodal map. The following result is due to Masur and Schleimer [6].

**Theorem 2.16** (Masur and Schleimer). Let $N$ be a non-orientable surface with an orientation double cover $S$. Then $\mathcal{C}(N)$ is quasi-isometric to a subcomplex of $\mathcal{C}(S)$ that is invariant under the action of the deck transformation.

The final result that we will need is given by Farb and Margalit [1].

**Theorem 2.17** (Farb and Margalit). Let $A = \{\alpha_1, \ldots, \alpha_n\}$ and $B = \{\beta_1, \ldots, \beta_m\}$ be multicurves in an orientable surface $S$ that together fill $S$. Any product of positive powers of the $T_{\alpha_i}$ and negative powers of the $T_{\beta_i}$, where each $\alpha_i$ and each $\beta_i$ appear at least once, is pseudo-Anosov.

Using these three results, we can prove the following.

**Proposition 2.18.** Let $N$ be a non-orientable surface whose orientation double cover $S$ is non-sporadic. Then $\mathcal{C}(N)$ has an infinite diameter.

**Proof.** By Theorem 2.16, $i : \mathcal{C}(N) \hookrightarrow \mathcal{C}(S)$ is a quasi-isometric embedding with $i(\mathcal{C}(N))$ invariant under the deck transformation $f$ associated to the orientation double cover. The deck transformation is simply the restriction of the map $x \mapsto -x$ in $\mathbb{R}^3$ to double cover, embedded in the obvious way in $\mathbb{R}^3$.

![Figure 7](image1.png)  
**Figure 7.** Multicurves in the orientation double cover for Proposition 2.18 where $F \cong P\#S_{g,s}$. Let $A = \{a_i, d_i\}$ and $B = \{b_i, c_i, f_i, h_i\}$.

![Figure 8](image2.png)  
**Figure 8.** Multicurves in the orientation double cover for Proposition 2.18 where $F \cong K\#S_g$. Let $A = \{a_i, d_i\}$ and $B = \{b_i, c_i, f_i, h_i\}$.
Consider either the collection of curves in Figure 7 if $N \cong P \# S^n_3$ or the collection of curves in Figure 8 if $N \cong K \# S^n_g$. Notice that $A \cup B$ fill $S$ in both cases. Consider the homeomorphism

$$g := \prod_{\ell \in A} T_\ell \prod_{\ell \in B} T^{-1}_\ell.$$  

By Theorem 2.17, $g$ is pseudo-Anosov. Notice that $f$ takes $A$ to $A$ and $B$ to $B$. Moreover, if $x$ and $y$ are in $A$, then $T_x T_y = T_y T_x$. Similarly for $B$. It follows that $f \circ g = g \circ f$. Therefore, $g$ is a pseudo-Anosov homeomorphism that commutes with the deck transformation. Consequently, $g$ takes $i(C(F))$ into itself. The result follows from Theorem 2.15. 

Finally, we need the following definition given by Masur and Schleimer [6].

**Definition 2.19.** Let $X$ be a subsurface of a surface $F$. The *subsurface projection*, $\pi_X : C(N) \to C(X)$, is defined as follows. Fix $\alpha$ in $C(N)$ and isotope $\alpha$ to minimize the number of connected components of $\alpha \cap X$. Now,

- if $\alpha \subset X$, then set $\pi_X(\alpha) = \alpha$.
- if $i(\alpha, \partial X) > 0$, then pick any arc $\alpha' \subset \alpha \cap X$, set $Y$ equal to a closed neighborhood of $\alpha' \cap \partial X$, and set $\pi_X(\alpha)$ equal to any component of $\partial Y$ which is essential in $X$.
- if $\alpha \subset F - X$, then set $\pi_X(\alpha) = \emptyset$.

We can now prove that the natural inclusion of $C_2(N)$ into $C(N)$ is not a quasi-isometric embedding.

**Proposition 2.20.** Let $N$ be a non-orientable surface whose orientation double cover $S$ is non-sporadic. The natural inclusion of $C_2(N)$ into $C(N)$ is not a quasi-isometric embedding.

**Proof.** Let $\mu \in N$ be a one-sided curve. Consider the subsurface $X = N - \mu$. By Proposition 2.18, we know that for any $m$ there exists $\alpha$ and $\beta$ in $C_2(X)$ such that $d_{C(X)}(\alpha, \beta) \geq m$. By the injectivity of $\pi_1(X)$ into $\pi_1(N)$, we have that both $\alpha$ and $\beta$ are in $C_2(N)$. Next, notice that if $(\alpha = \gamma_0, \ldots, \gamma_n = \beta)$ is a path from $\alpha$ to $\beta$ in $C(N)$, then $d_{C(X)}(\alpha, \beta) \leq 6n$. This follows from the fact that if $i(\gamma, \delta) = 0$, then $i(\pi_X(\gamma), \pi_X(\delta)) \leq 4$ and by Proposition 2.1 we have $d_{C(X)}(\pi_X(\gamma), \pi_X(\delta)) \leq 6$. Chaining together the above inequalities we have

$$m \leq d_{C(X)}(\alpha, \beta) \leq 6 \cdot d_{C(N)}(\alpha, \beta) \leq 6 \cdot d_{C_2(N)}(\alpha, \beta).$$

It follows that we may select $\alpha$ and $\beta$ in $C_2(N)$ that are arbitrarily far apart in $C_2(N)$. Moreover, we may choose both $\alpha$ and $\beta$ to be disjoint from $\mu$ and consequently $d_{C(N)}(\alpha, \beta) = 2$. It follows that the natural inclusion takes vertices at arbitrarily large distances to distance 2. Therefore, the natural inclusion is not a quasi-isometric embedding. 

We may use a similar argument to show that the natural inclusion of $C_1(N)$ into $C(N)$ is not a quasi-isometric embedding.

**Proposition 2.21.** Let $N$ be a non-orientable surface whose orientation double cover $S$ is non-sporadic. The natural inclusion of $C_1(N)$ into $C(N)$ is not a quasi-isometric embedding.
Proof. Let $\gamma \in N$ be a non-separating curve that is not one-sided. Consider the non-orientable subsurface $X = N - \gamma$. Applying the same argument as in the proof of Proposition 2.20 with this different $X$ and instead $\alpha$ and $\beta$ being in $C_1(X)$, we have that we may select $\alpha$ and $\beta$ in $C_1(N)$ that are arbitrarily far apart in $C_1(N)$ but that are distance 2 apart in $C(N)$. It follows that the natural inclusion takes vertices at arbitrarily large distances to distance 2. Therefore, the natural inclusion is not a quasi-isometric embedding. \qed

We conclude this section by studying the dimensions and maximal simplices of $C(N)$ and its natural sub-complexes $C_1(N)$ and $C_2(N)$.

**Proposition 2.22.** Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. Then a pants decomposition of $N$ is composed of $c + s - 2$ pairs of pants.

**Proof.** Suppose that there are $m$ pairs of pants, then an Euler characteristic argument shows that

$$-m = \chi(N) = 2 - c - s.$$

The result follows. \qed

**Proposition 2.23.** Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. Let $\Delta$ be a maximal simplex of $C(N)$.

1. If $c$ is odd, then

$$\frac{3c}{2} + s - 2 \leq \dim(\Delta) \leq 2c + s - 3.$$

2. If $c$ is even, then

$$\frac{3c - 2}{2} + s - 2 \leq \dim(\Delta) \leq 2c + s - 3.$$

**Proof.** By Proposition 2.22, we have that $m = c + s - 2$. Counting the boundary curves and the curves we cut along to obtain our pant’s decomposition, we must have that

$$3m = 2t + \ell + s$$

where $t$ is the number of two-sided curves that we cut along and $\ell$ is the number of one-sided curves that we cut along. An easy computation shows that

$$t = \frac{3c - \ell + 2s - 6}{2}.$$

If $c$ is odd, then $1 \leq \ell \leq c$. It follows that

$$\frac{3c + 1}{2} + s - 3 \leq t + \ell \leq 2c + s - 3.$$

If $c$ is even, then $0 \leq \ell \leq c$. It follows that

$$\frac{3c}{2} + s - 3 \leq t + \ell \leq 2c + s - 3.$$

The result follows by noticing that any maximal simplex in $C(N)$ is given by $t$ two-sided curves and $\ell$ one-sided curves. \qed

**Corollary 2.24.** Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. If $\Delta$ is a top-dimensional maximal simplex of $C_1(N)$, then $\Delta$ has dimension $2c + s - 3$. 
Corollary 2.25. Let $N$ be a non-orientable surface with $c$ cross caps and $s$ boundary components. Let $\Delta$ be a top-dimensional maximal simplex of $C_2(N)$.

1. If $c$ is odd, then
   \[ \dim(\Delta) = \frac{3c - 1}{2} + s - 2. \]
2. If $c$ is even, then
   \[ \dim(\Delta) = \frac{3c - 2}{2} + s - 2. \]

3. Homotopy Type of the Curve Complex

In this section, we show that the argument used by Harer [2] to compute the homotopy type of the curve complex of an orientable surface can be generalized for non-orientable surfaces. Harer’s result is as follows.

Theorem 3.1 (Harer). For $g > 0$, the complex $C(S_g^0)$ is homotopy equivalent to a wedge of spheres of dimension $2g - 2$, and the complex $C(S_g^n)$ is homotopy equivalent to a wedge of spheres of dimension $2g + n - 3$ for $n > 0$. For $g = 0$, $C(S_g^0)$ is homotopy equivalent to a wedge of spheres of dimension $n - 4$.

We show the following generalization for non-orientable surfaces.

Theorem 3.2. Let $N$ be a non-orientable surface with $c > 0$ cross caps and $n$ punctures. The curve complex of $N$ is homotopy equivalent to a wedge of spheres of dimensions fewer than or equal to $c + n - 2$ if $n = 0$ and $c + n - 3$ if $n > 0$.

To prove Theorem 3.2, we start with the following observation.

Remark 3.3. Consider $C(N_c^n)$ and let $\{p_1, \ldots, p_n\}$ denote the collection of punctures in $N_c^n$. Let $\hat{C}(N_c^n)$ denote the sub-complex of $C(N_c^n)$ spanned by simplices $\langle \gamma_0, \ldots, \gamma_k \rangle$ such that $\gamma_i$ does not bound a disk containing $p_1$ and some other $p_j$. Notice that forgetting the point $p_1$ gives rise to a map
   \[ \varphi : \hat{C}(N_c^n) \to C(N_c^{n-1}) \]
when $n > 0$. When $n = 1$, we have that $\hat{C}(N_c^1) = C(N_c^1)$.

The following lemma is due to Harer [2]. We give it here for completeness and to show that it holds for non-orientable surfaces.

Lemma 3.4. $\varphi$ is a homotopy equivalence for all $n$.

Proof. Let $N' = N_c^0 - \{p_2, \ldots, p_n\}$ and choose a complete hyperbolic metric of finite area on $N'$. Each isotopy class of simple closed curves is given by a unique geodesic. Since there exists a point in $N'$ that does not meet a geodesic, we may select such a point, say $p'$. Then $\psi : C(N_c^{n-1}) \to \hat{C}(N_c^n)$ gives a mapping by setting $p_1 = p'$. Notice that $\psi$ can be viewed as an inclusion. Furthermore, since its clear that $\varphi \circ \psi$ is the identity, $Id$, on $C(N_c^{n-1})$, we must have that $\varphi$ is a retraction of $\hat{C}(N_c^n)$ onto the image of $\psi$ and consequently the naturally included copy of $C(N_c^{n-1})$. Applying the functor $\pi_n(-)$, we have that $\varphi_* \circ \psi_* = Id_*$ and consequently $\varphi_* : \pi_n(\hat{C}(N_c^n)) \to \pi_n(C(N_c^{n-1}))$ is surjective for all homotopy groups. To complete the proof we will show that $\varphi_*$ is injective for all homotopy groups. Applying Whitehead’s Theorem in the case of simplicial complexes, it follows that $\varphi$ is a homotopy equivalence.
Let \( f : \Sigma^n \to \hat{C}(N^n) \) be a simplicial map of a piecewise linear \( n \)-sphere and set \( f' = \psi \circ \varphi \circ f \). We show that \( f' \) is homotopic to \( f \), completing the proof. Let \( N'' = N' - p_1 \). Given vertices \( v_1, \ldots, v_k \) of \( \Sigma^n \), set \( c_i = f(v_i) \) and \( c'_i = f'(v_i) \) such that \( c_i \) and \( c'_i \) are the geodesic representatives in \( N'' \), on which we may choose a complete hyperbolic metric of finite area. Notice that \( c_i \) is isotopic to \( c'_i \) in \( N' \).

Let \( A_1, \ldots, A_k \subset N' \times I \) denote the imbedded annuli which give isotopies from \( c_i \) to \( c'_i \) such that \( A_i \cap (N' \times \{0\}) = c_i \) and \( A_i \cap (N' \times \{1\}) = c'_i \). Notice that the product structure on each \( A_i \) is compatible with the product structure on \( N' \times I \). We may arrange the \( A_i \) to be pairwise transverse and each \( A_i \) to be transverse to \( p_1 \times I \). We will arrange the \( A_i \) such that if \( c_i \cap c_j = \emptyset \) then either \( A_i \cap A_j = \emptyset \) or \( c'_i = c'_j \) and \( A_i \cap A_j = c'_i \). Suppose that \( i(c_i, c_j) = 0 \) for some \( i \) and \( j \). Notice that either \( (A_i \cap A_j) \cap (N' \times 1) = \emptyset \) or \( c'_i \) is equal to \( c'_j \). It follows that any unwanted intersections that ruin the desired arrangement mentioned above between \( A_i \) and \( A_j \) are circles.

Let \( c \subset A_i \cap A_j \) be such a circle. Suppose that \( c \) is isotopic to the boundary of \( A_i \). Since \( c \) is simple, it must be a generator in \( \pi_1(A_i) \). But \( \pi_1(N' \times I) \) being equal to \( \pi_1(N') \) implies that \( c \) must be a generator in \( \pi_1(A_j) \) otherwise \( c_j \) would be non-essential in \( N' \). It also follows that \( c'_i \) equals \( c'_j \). Now we may simply interchange the part of \( A_i \) which contains \( c'_i \) with the part of \( A_j \) that contains \( c'_j \) and push the modified \( A_i \) and \( A_j \) off each other in order to eliminate \( c \) without changing the product structure. We may repeatedly apply this argument to remove all circles that are generators in \( \pi_1(A_i) \) for all \( i \).

Now suppose that \( c \) is not homotopic to the boundaries \( A_i \) and \( A_j \). Then \( c \) bounds a disk \( D_i \) in \( A_i \) whose interior is disjoint from \( A_j \) and a disk \( D_j \) in \( A_j \) whose interior is disjoint from \( A_i \). Replace \( D_j \) with \( D_i \) and push out slightly to remove \( c \) from \( A_i \cap A_j \). This procedure maintains the product structure and no new intersections are created. We may repeatedly apply these processes until all the remaining circles of intersection between the \( A_i \) are removed.

Now we may move the \( A_i \) such that \( p_1 \times I \) intersects only one annulus at a time. Let \( 0 < t_1 < \cdots < t_m < 1 \) be the times that these intersections occur. Suppose that \( A_i \) meets \( p_1 \) at \( t_1 \). Let \( d_1 = A_i \cap (N' \times \{t_1 + \varepsilon\}) \) viewed in \( N' \) such that \( t_1 + \varepsilon < t_2 \) and \( \varepsilon > 0 \). Notice that \( d_1 \) may be isotope to a curve that is disjoint from \( c_i \). By the conditions imposed on the \( A_i \) above, we have that if \( i(c_i, c_j) = 0 \), then \( d_1 \) may be isotope to a curve that is disjoint from \( c_j \). If \( \tau = \langle v_1, v_{j_1}, \ldots, v_{j_k} \rangle \) is a \( j \)-dimensional simplex of \( \Sigma^n \) in the star of \( v_1 \), then add \( \langle v_1, w_1, v_{j_1}, \ldots, v_{j_k} \rangle \) to \( \Sigma^n \) along \( \tau \), where \( w_1 \) is a newly introduced vertex. We may extend \( f \) in the obvious manner by sending \( w_1 \) to \( d_1 \). Continuing this process for all \( t_i \), the result is a homotopy of \( f \) to \( f' \). This completes the proof.

We now turn to the proof of Theorem 3.2. Our argument below is in the spirit of Harer’s original argument for the orientable case.

**Proof.** Suppose by way of induction that the result holds for surfaces with any number of cross caps fewer than \( c \) and with an arbitrary number of punctures. The base case of a sphere with an arbitrary number of punctures is given by Theorem 3.1. Let \( \mathcal{C}(N^n)^\circ \) denote the barycentric subdivision of \( \mathcal{C}(N^n) \). Notice that the vertices of \( \mathcal{C}(N^n)^\circ \) correspond to multicurves and faces are added for chains of proper inclusions among these multicurves. More precisely, the collection of multicurves \( \langle \alpha_0^j, \ldots, \alpha_j^j \rangle \) with \( 0 \leq j \leq k \) span a \( k \)-simplex if \( \langle \alpha_0^j, \ldots, \alpha_j^j \rangle \subset \langle \alpha_0^{j+1}, \ldots, \alpha_{j+1}^{j+1} \rangle \) for all \( j < k \).
Given \( v \in \mathcal{C}(N_\circ^0) \), we define the weight of \( v \), denoted \( w(v) \), to be the number of curves contained in the multicurve \( v \) minus one. Equivalently, \( w(v) \) is the dimension of the simplex for which \( v \) was a barycenter.

Let \( X_k \) denote the full sub-complex of \( \mathcal{C}(N_\circ^0) \) composed of vertices with weight greater than or equal to \( k \). We will construct \( \mathcal{C}(N_\circ^0) \) by consecutively adding vertices of lower weight and show for each \( k \) that \( X_k \) is a wedge of spheres with dimension fewer than or equal \( c-2 \). Let \( d \) denote the dimension of the top dimensional maximal simplex in \( \mathcal{C}(N_\circ^0) \). Notice that \( X_d \) is a discrete set of points and is thus a wedge of zero spheres and of dimension fewer than \( c-2 \). Inductively, suppose that \( X_{k+1} \) has been shown to be a wedge of spheres of all of dimensions fewer than \( c-2 \). Consider some \( v = \langle \gamma_0, \ldots, \gamma_k \rangle \in X_k - X_{k+1} \). Assuming that \( v \) has a non-empty link, we let \( \{N^i\}_i \) be the set of connected components of \( N_\circ^0 - \cup_j \gamma_j \). Notice that the link of \( v \), \( L(v) \), in \( X_{k+1} \) is simply the join of \( \mathcal{C}(N^1), \ldots, \mathcal{C}(N^t) \). Without loss of generality, suppose \( N^1, \ldots, N^t \) are orientable surfaces with \( g_1, \ldots, g_t \) genera and \( r_1, \ldots, r_t \) boundary components respectively. Suppose that \( N^{t+1}, \ldots, N^t \) are non-orientable surfaces with \( c_{t+1}, \ldots, c_t \) cross caps and \( r_{t+1}, \ldots, r_t \) boundary components respectively. Via an Euler characteristic argument, we have

\[
2 - c = \chi(N_\circ^0) = \sum_i \chi(N^i) = \left( \sum_{i=1}^t 2 - 2g_i - r_i \right) + \left( \sum_{i=t+1}^t 2 - c_i - r_i \right) = 2t - \left( \sum_{i=1}^t 2g_i + r_i \right) - \left( \sum_{i=t+1}^t c_i + r_i \right).
\]

It follows that

\[
c = \left( \sum_{i=1}^t 2g_i + r_i + \sum_{i=t+1}^t c_i + r_i \right) - 2t + 2.
\]

By Theorem 3.1, we have that for \( i < \ell + 1 \) each \( N^i \) is a wedge of spheres of dimension fewer than \( c-2 \). Inductively, we have that for \( i > \ell \) each \( N^i \) is a wedge of spheres of dimension fewer than \( c-2 \). More explicitly, we have that \( \dim(\mathcal{C}(N^i)) = 2g_i + r_i - 3 < c-2 \) for \( i < \ell + 1 \) and \( \dim(\mathcal{C}(N^i)) \leq c_i + r_i - 3 \) for \( i > \ell \). Recalling that \( L(v) \) is the join of \( \mathcal{C}(N^1), \ldots, \mathcal{C}(N^t) \), we have that

\[
\dim(L(v)) \leq \left( \sum_{i=1}^t 2g_i + r_i - 3 \right) + \left( \sum_{i=t+1}^t c_i + r_i - 3 \right) + (t-1) = \left( \sum_{i=1}^t 2g_i + r_i \right) + \left( \sum_{i=t+1}^t c_i + r_i \right) - 3t + t - 1 = c - 3.
\]

Notice that \( X_k \) is obtained from \( X_{k+1} \) by forming the join of \( v \) with its link in \( X_{k+1} \) for all \( v \) in \( X_k - X_{k+1} \). It remains to show that \( X_k \) is not contractible. Consider \( v \) and \( w \) in \( X_k - X_{k+1} \) such that \( v = \langle \gamma_0, \ldots, \gamma_k \rangle \) and \( w = \langle \delta_0, \ldots, \delta_k \rangle \). If \( i(\gamma_i, \delta_j) \neq 0 \) for some \( i \) and \( j \), then there does not exist \( x \) in \( X_{k+1} \) such that \( v, w \subset x \). It follows that \( L(v) \cap L(w) = \emptyset \). If \( i(\gamma_i, \delta_j) = 0 \) for all \( i \) and \( j \), then set
\( x = (\gamma_0, \ldots, \gamma_k, \delta_0, \ldots, \delta_k). \) Removing redundancies let us write \( x = (\varepsilon_0, \ldots, \varepsilon_m). \) Notice that \( v, w \subset x \) and \( L(v) \cap L(w) \) is the link of \( x \) in \( X_{m+1}. \) But inductively, \( L(x) \) with its subspace topology in \( X_{m+1} \) is not contractible and thus forming the joins of \( v \) and \( w \) to their respective links suspends \( L(x), \) producing a non-contractible space. This completes the induction on \( c \) and the proof for the case of \( n = 0. \)

To handle the case of \( n = 1, \) recall that \( \hat{\mathcal{C}}(N_{1c}) = \mathcal{C}(N_0^c). \) By Lemma 3.4, we have that \( \hat{\mathcal{C}}(N_0^c) \) is homotopic to \( \mathcal{C}(N_0^c). \) It follows that \( \hat{\mathcal{C}}(N_0^c) \) is homotopic to a wedge of spheres each of dimension fewer than or equal to \( c + (1) - 3. \) To handle the remaining cases, suppose that

\[
\mathcal{C}(N_{n-1}^c) \simeq \bigvee_i S^{n_i}
\]

such that \( 0 \leq n_i \leq c + (n - 1) - 3 \) with \( n > 1. \) By Lemma 3.4, we have \( \hat{\mathcal{C}}(N_n^c) \) is homotopic to \( \hat{\mathcal{C}}(N_{n-1}^c). \) But \( \hat{\mathcal{C}}(N_n^c) \) is obtained from \( \hat{\mathcal{C}}(N_{n-1}^c) \) by forming the join with each ommitted vertex. All such vertices correspond to curves that encircle \( p_1 \) and one other \( p_i. \) If two such curves are disjoint then they must be isotopic. It follows that \( \varphi \) sends the links of such vertices homeomorphically onto \( \hat{\mathcal{C}}(N_{n-1}^c). \) Hence, \( \mathcal{C}(N_n^c) \) is homotopic to the join of \( \hat{\mathcal{C}}(N_{n-1}^c) \) with a wedge of zero spheres. This completes the proof. \( \square \)

**Acknowledgments.** I would like to thank my mentors Nicholas Vlamis, Matthew Durham, and Richard Canary for all of their time, expertise, and guidance throughout this REU program. I would like to specifically thank Nicholas Vlamis for all his comments, critiques, and suggestions on this paper. I would like to thank John Harer for an enlightening discussion about the homotopy type of the curve complex of an orientable surface. Finally, I would like to thank the University of Michigan for running this REU program and the NSF for providing research funding.

**References**


