GEOMETRIC INTERSECTION OF CURVES ON SURFACES

DYLAN P. THURSTON

This is a pre-preprint. Please give comments!

Abstract. In order to better understand the geometric intersection number of curves, we introduce a new tool, the *smoothing lemma*. This lets us write the geometric intersection number $i(X, A)$, where X is a simple curve and A is any curve, canonically as a maximum of various $i(X, A')$, where the A' are also simple and independent of X.

We use this to get a new derivation of the change of Dehn-Thurston coordinates under an elementary move on the pair of pants decomposition.

Contents

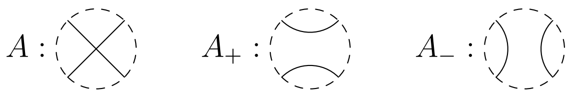
Key words and phrases. Dehn coordinates,curves on surfaces,tropical geometry.

1. Introduction

In 1922, Dehn [2] introduced what he called the *arithmetic field of a surface*, by which he meant coordinates for the space of simple curves on a surface, together with a description for how to perform a change of coordinates. In particular, this gives the action of the mapping class group of the surface. He presumably chose the name "arithmetic field" because in the simplest non-trivial cases, namely the torus, the once-punctured torus, and the 4-punctured sphere, the result is equivalent to continued fractions and Euler's method for finding the greatest common divisor. These coordinates were later rediscovered [1, 14], and now go under the name Dehn-Thurston coordinates. But the question of how they change under change of coordinates has been largely neglected. It was not until 1982 that Penner [11, 12] gave explicit formulas for the change of coordinates, but probably due to the intimidating complexity of his formulas, little further work has been done with this.

In this paper, we study geometric intersection numbers $i(\cdot, \cdot)$ of two curves. One principal tool is the following lemma:

Lemma 1 (Smoothing)**.** *Let X be a simple closed curve and A an arbitrary curve (with self-intersections), embedded so that $X \cup A$ is taut (has a minimal number of intersections). Pick an intersection of A, and let A_+ and A_- be the two different ways to smooth that intersection:*



Then

$$i(A, X) = \max(i(A_+, X), i(A_-, X)).$$

This lemma and extensions are proved in Section 3; some refinements are in Section 5. With the aid of this lemma we are able to get simple formulas for the change of coordinates on the space of simple curves in several cases, including parametrizations by triangulations or by pairs of pants. This has been considered before: the case of triangulations is well-known while the case of pairs of pants was previously considered by Penner [11,12]. In this paper we use somewhat different coordinates and new tools to get an end result that is simpler and more conceptual.

Among other consequences, we get an $O(n^2)$ algorithm for solving the word problem in the mapping class group, with constants independent of the surface. We are also able to easily find the stable and unstable measured laminations of a pseudo-Anosov element of the mapping class group. This might help in solving the conjugacy problem as well.

Another application of the Smoothing Lemma is *convexity*. By picking a triangulation (resp. a pants decomposition) we give coordinates on the space of simple curves. Given a simple curve X and another curve A, where A is not necessarily simple, we can think of $i(A, X)$ as a function on the space of coordinates. In Section 7 we show that this function is convex in terms of these coordinates.

Finally, we note that there is also an algebraic version of the Smoothing Lemma, related to hyperbolic geometry. The two lemmas are related, on the algebraic side, by tropicalization, and, on the geometric side, by passing to the boundary of Teichmüller space. This algebraic lemma in the case of triangulations gives examples of *cluster algebras* [3]. The cluster algebras constructed in this way from surfaces are the only known infinite class of examples of cluster algebras which are mutationally finite but not finite. The convexity mentioned above is also interesting in this case, where it translates into a proof of positivity of certain coefficients. For closed surfaces, our results should be compared with the work of Okai [9, 10], who solved the analogous problem for hyperbolic geometry and the corresponding Fenchel-Nielsen coordinates. Details will appear in a forthcoming paper.

1.1. **Acknowledgements.** Much of the work was done while I was at Harvard, first as a NSF Postdoctoral Fellow and later as a Benjamin Peirce Assistant Professor. I would like to thank Joel Hass, Feng Luo, Chung-chieh Shan, and William Thurston for helpful conversations.

2. Preliminaries on curves

In this paper, we will be considering compact oriented surfaces Σ, possibly with boundary, and several varieties of 1-manifolds on them.

Definition 2. A *curve* on Σ is a properly immersed 1-manifold, possibly with boundary. The set of all curves on Σ is denoted $\mathcal{C}(\Sigma)$. $\mathcal{C}_0(\Sigma)$ is the set of curves with no boundary, and for S a subset of the boundary components of Σ, $\mathcal{C}_S(\Sigma)$ is the set of curves with boundary only on S.

The circle components of $C \in \mathcal{C}(\Sigma)$ are called *loops* and the other components are called *arcs*.

Note that a curve is what other authors have called a multi-curve or sometimes a curve system.

Definition 3. A curve $C \in \mathcal{C}_S(\Sigma)$ is *taut* if it is immersed with the minimum number of self-intersections in its homotopy class, it has no trivial components, and it has no components which are homotopic into any of the boundary curves in S. If C is not taut, it is said to have *excess intersections*.

We exclude various trivial cases so that for any $C \in \mathcal{C}_S(\Sigma)$, there is a $D \in \mathcal{C}_{\bar{S}}(\Sigma)$ which interesects C essentially, where $\bar{S}$ is the set of connected components of $\partial\Sigma \setminus S$.

There is a nice geometric characterization of tautness.

Theorem 1 (Freedman-Hass-Scott [4], Neumann-Coto [8])**.** *Let C be a curve with no components which are proper powers of other curves and no two parallel components. Then C is taut if and only if it is length-minimizing in its homotopy class with respect to some Riemannian metric on Σ.*

(The reason for the restriction on C is that in the excluded cases, length-minimizers are covers of a single curve and do not intersect transversally.)

Corollary 4. *If $C \in \mathcal{C}(\Sigma)$ is tautly embedded and $D \in \mathcal{C}(\Sigma)$ is arbitrary, then there is a taut embedding of $C \cup D$ which agrees with the given embedding of C.*

We will generally be interested in parametrizing simple closed (multi-)curves.

Definition 5. A *simple curve* on Σ is an taut, embedded closed curve C, not necessarily connected. The space of simple curves up to ambient isotopy (or equivalently up to homotopy) is denoted $\mathcal{SC}(\Sigma)$. Similarly for $\mathcal{SC}_0(\Sigma)$ and $\mathcal{SC}_S(\Sigma)$.

Our "simple curves" are also called "curve systems" or "integer measured laminations". The name "integer measured lamination" comes from the general theory of measured laminations; all results in this paper extend to the more general case.

To parametrize simple curves, we will use intersection numbers.

Definition 6. Given $C \in \mathcal{C}_S(\Sigma)$ and $D \in \mathcal{C}_T(\Sigma)$, where S and T are disjoint, the *geometric intersection number* $i(C, D)$ is the number of intersections between C and D in a taut embedding of $C \cup D$. We will also write this as $i_C(D)$ or $i_D(C)$.

3. Smoothing curves

The following lemma is a central tool in this paper.

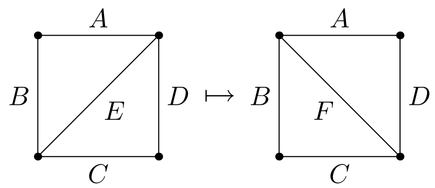
Lemma 7 (Smoothing)**.** *Let S and T be disjoint subsets of $\partial\Sigma$, and let $X \in \mathcal{SC}_S(\Sigma)$ and $A \in \mathcal{C}_T(\Sigma)$, with A tautly embedded. Let x be a crossing of A, and let A_+ and A_- be the two smoothings of A at x. Then*

$$i_X(A) = \max(i_X(A_+), i_X(A_-)).$$

For instance, suppose A, B, C, D, E, and F are the 4 sides and two diagonal of a quadrilateral in Σ. (The corners are four boundary components of Σ, possibly with repetitions.) We can use this theorem to compute $i_X(F)$ in terms of the number intersections of F with the other 5 sides: we consider the non-simple curve $E \cup F$. This curve has one self intersection; if we smooth out the intersection, we get two opposite sides of the square, so we have

$$\begin{aligned} i_X(E \cup F) &= \max(i_X(A \cup C), i_X(B \cup D)) \\ i_X E + i_X F &= \max(i_X A + i_X C, i_X B + i_X D). \end{aligned}$$

If we use the number of intersections with the edges of a triangulation to parametrize the different curves X, this equation lets us change coordinates from one triangulation to another:



This coordinate change was known before and is in any case easy to derive by a case analysis. We will be using this theorem in more sophisticated ways later.

Before we prove Lemma 7, let us study excess intersections in this context a little.

Definition 8. A *weak bigon* between two curves C_1 and C_2 (intersecting transversally) is a map $f : D^2 \to \Sigma$ so that ∂D^2 is the union of two intervals A_1 and A_2 so f agrees with C_1 on A_1 and f agrees with C_2 on A_2.

Hass and Scott [5] have considered various notions of "bigons" and whether they exist for various types of embeddings with excess intersections; unfortunately, none of their theorems are quite what we need. Our notion of "weak bigon" is even weaker than their weakest notion of bigon, since the boundary of the bigon may wrap more than once around C_1 and/or C_2.

Note that the existence of a weak bigon is equivalent to a lift $\tilde{C}_1$ of C_1 and a lift $\tilde{C}_2$ of C_2 intersecting more than once in the universal cover $\tilde{\Sigma}$ of Σ.

Lemma 9. *If X is embedded and C is a loop, then no weak bigon between A and X wraps more than once around C.*

Proof. Consider the annulus cover Σ_C of Σ corresponding to C. Then the lifts of C and X to Σ_C are embedded. An arc bounding a bigon that wraps more than once around the lift of C is necessarily not embedded. $\square$

Lemma 10. *Let $A \in \mathcal{C}_T(\Sigma)$ and $X \in \mathcal{SC}_S(\Sigma)$, with S and T disjoint, and suppose we are given a transverse immersion of $X \cup A$. Then this immersion has excess intersections between X and A if and only if there is a weak bigon between X and A.*

Proof. If there is a bigon between X and A, by Lemma 9 it involves an arc of A. We can push that arc of A across the disk defining the bigon until it lies on the other side of A. When we do this we remove two intersections of A with X (at the two corners of the bigon), and, since X is simple, we do not introduce any new intersections between A and X.

Conversely, suppose that there is excess intersection between A and X. By a result of Hass and Scott [6, Theorem 2.2], there is a homotopy from $A \cup X$ to an immersion with minimal intersection which never increases the number of self-intersections. If we put the homotopy in general position, this means that there is a sequence of $3 \leftrightarrow 3$, $2 \to 0$, $1 \to 0$, and $\partial 1 \to 0$ moves which makes $A \cup X$ taut. Since there are excess intersections between A and X, there will be at least one $2 \to 0$ move involving A and X. (This is the only possible move that reduces the number of AX intersections.) Consider the configuration just before the first of such move; by definition, there will be an embedded bigon between A and X. We can follow this bigon backwards through the sequence of moves that preceded it to find a weak bigon in our original configuration. $\square$

Question 11. Can excess intersections more generally be characterized by the existence of weak monogons, bigons, and boundary triangles? The argument using curve shortening generalizes to show that if there is excess intersection, then there is also a weak monogon, bigon, or boundary triangle. The other direction seems more difficult.

Lemma 7. By Corollary 4, we can suppose that X and A are simultaneously embedded so that they are both taut. Then $i_X(A_+) \leq i_X(A)$, since the concrete configuration of A_+ has $i_X(A)$ intersections with X, and a taut configuration can only have fewer intersections. Similarly $i_X(A_-) \leq i_X(A)$.

So we only need to show that equality is achieved at least once. Consider the lifts $\tilde{A}$ and $\tilde{X}$ of A and X to the universal cover of Σ, and a lift $\tilde{x}$ of x. Name the 4 rays of $\tilde{A}$ that meet at $\tilde{x}$ by p, q, r, and s, in cyclic order, so that $\tilde{A}_+$ joins p to q and r to s and $\tilde{A}_-$ joins p to s and q to r. If $\tilde{A}_+ \cup \tilde{X}$ is not taut, then there is a bigon between the two curves. (The other cases are ruled out, by assumption on A.) Since the original $\tilde{A}$ was taut, this bigon must not have appeared originally; that is, it must have come from a triangle involving X and two rays of $\tilde{A}$ from $\tilde{x}$; without loss of generality, these are p and q. Any strand of $\tilde{X}$ that crosses this triangle must meet both strands p and q: it cannot meet one strand twice, by tautness of $\tilde{A} \cup \tilde{X}$, and it cannot meet the $\tilde{X}$ side, since $\tilde{X}$ is embedded. Thus there is a unique innermost such triangle which involves the first intersections with $\tilde{X}$ along p and q. Similarly if $\tilde{A}_- \cup \tilde{X}$ is not taut, there is a triangle involving X and (say) the first intersections along q and r. But these two triangles together yield a X-p-r bigon, contradicting the assumption that $A \cup X$ was taut. □

By using Lemma 7 repeatedly, $i(A, X)$, for X simple and A arbitrary, can be turned into a maximum of various $i(A', X)$, where the A' are simple. There may a priori be more than one way to do this: at each stage we smooth one crossing, then homotop the resulting curves to make them taut, and iterate. There is no guarantee that we will get the same set of curves regardless of the order we choose to look at the crossings. The following lemma lets us do all the smoothings at once.

Lemma 12. *Let $A \in \mathcal{C}_S(\Sigma)$ and $X \in \mathcal{SC}_T(\Sigma)$, where S and T are disjoint and A is tautly embedded. Let $\mathrm{Sm}(A) \subset \mathcal{SC}_S(\Sigma)$ be the set of all curves obtained by smoothing all of the crossings in A, with any trivial components deleted. Then*

$$i(A, X) = \max_{A' \in \mathrm{Sm}(A)} i(A', X).$$

Proof. As before, for any smoothing A' of A, the concrete realization of A' has as many intersections with X as A does, and this number can only decrease when we make $A' \cup X$ taut, so one inequality is clear. For the other, we need to show that one curve in $\mathrm{Sm}(A)$ achieves an equality. By Lemma 7, we know that by smoothing one crossing to get a curve A', simplifying the configuration of the curve to A'', and repeating we can get to a curve that achieves equality. By [6, Theorem 2.2], the process of simplifying can be done with a sequence of $3 \leftrightarrow 3$, $2 \to 0$, $1 \to 0$, and $\partial 1 \to 0$ moves. By Lemma 13 below, $\mathrm{Sm}(A'') \subset \mathrm{Sm}(A')$. By induction, one of the elements of $\mathrm{Sm}(A'')$ achieves the desired equality with $i(A, X)$. □

Lemma 13. *Let A and A' be two configurations of curves on Σ differing by a single positive elementary move. Then* $\mathrm{Sm}(A') \subset \mathrm{Sm}(A)$. *In* *particular, if A and A' are two taut configurations of the same homotopy class, then* $\mathrm{Sm}(A') = \mathrm{Sm}(A)$*.*

Proof. Direct verification in cases. □

4. Dehn-Thurston coordinates

In this section we review Dehn-Thurston coordinates, adapted to the versions that will give us the nicest formulas, and extend them to related settings.

We start with the case of a triangulated surface, where we parametrize simple closed curves in the complement of the vertices of the triangulation.

Proposition 14 ([**?**])**.** *On a surface Σ equipped with a triangulation T, a simple closed curve in minimal position with respect to the triangulation is uniquely determined by the number of intersections with each edge of the triangulation. Any number of intersections is acheivable, provided it satisfies the triangle inequality inside each triangle and (parity condition).*

Proof. There is a unique way to fill in each triangle. □

To parametrize curves on an oriented surface Σ with no punctures, we start with a *pants decomposition*, a maximal set $\{P_i\}$ of non-parallel closed curves. The complement $\Sigma \setminus P_i$ is then a union of disks with three perforations, topologically like a pair of pants, whence the name pants decomposition.

This marking does not suffice to pin down the surface: there is still a large subgroup of the mapping class group that fixes these pants curves, namely Dehn twists around the pants curves. There are several equivalent ways to fix this subgroup:

- *Reversing map*: An orientation-reversing map $R : \Sigma \to \Sigma$ so that for each i, $R(P_i) = P_i$.
- *Hexagonal decomposition*: A simple curve H that meets each pants curve twice, decomposing each pair of pants into two hexagons.
- *Dual curves*: For each i, a curve D_i so that $i(D_i, P_j) = 2\delta_{ij}$.

These are all equivalent: the curves in the hexagonal decomposition are the fixed points of R (and conversely R switches the two hexagons in each pair of pants), and the dual curves are chosen to run from one side of each hexagon to the opposite. (Alternatively, the D_i are also fixed by R.)

However we thing about this marking, Dehn-Thurston coordinates for an integer measured lamination consist of

- *length paramters* $m_i = \{\#(C \cap P_i)\}$, the number of intersection with each pants curve; and
- *twist parameters* t_i for each pants curve.

For the idea behind the twist parameters, note that once the m_i are fixed there is a unique way to fill them in and determine $C \cap Y$ for each pair of pants Y [?]. The m_i need only satisfy a parity condition, that the sum of the m_i for the three boundary components of a pair of pants is even; the triangle inequality is no longer necessary. For each P_i we have m_i points of intersection coming from each side, and we just have to decide how to glue them together: how far to rotate the "cuffs" before gluing. There are a $\mathbb{Z}$'s worth of ways to glue them together, giving the twist parameter t_i. The markings let us determine a canonical 0 twist for the t_i. The map R provides one convenient basepoint.

Definition 15. The twist parameter t_i is the unique parameter on $X \in \mathcal{SC}(\Sigma)$ so that

- t_i changes by two when X is twisted postively one step. Formally, $t_i(P_j X) = t_i(X) + 2\delta_{ij}$, where the product of curves is as defined below in Definition 18.
- t_i depends only on X restricted to the one or two pairs of pants that neighbor P_i.
- t_i is covariant with respect to orientation reversal: $t_i(R(X)) = -t_i(X)$.

We choose t_i to vary by two when the curve is twisted in order to guarantee that it will always be an integer. It is immediate from the definition how the twist parameters change by a Dehn twist around a pants curve:

Proposition 16. *Under a positive Dehn twist around the pants curve P_i, the corresponding twist parameter changes by*

$$t_i' = t_i + 2m_i.$$

Remark. There are different choices for the basepoint for the twist; for instance, Penner [**?**] has a different choice from the one above. See Appendix B for the relationship between the two sets of twist parameters.

More concretely, divide Σ up into annuli $\{A_i\}$ and pairs of pants Y_j. Arrange the dual curves D_i so that they are fixed by the orientation-reversing map R. (This is possible because of the relationship between the two ways of marking the twists.) Arrange the curve X so that in the pairs of pants it is fixed by the involution R and is transverse to the D_i. Also push any parallel copies of the pants curves into the corresponding annulus. Then to find the twist parameter for the pants curve P_i, orient both $X \cap A_i$ and $D_i \cap A_i$ to run consistently from one boundary component to the other. (The result is independent of this choice, as long as all components are oriented the same way.) Then the twist t_i is defined by

$$t_i = (X \cap A_i).(D_i \cap A_i)$$

where . denotes the *signed* intersection number. If the measure m_i is zero, then X may have several components parallel to the pants curve P_i, and it is ambiguous how to orient them. In this case, orient them all consistently to find a twist parameter, which has an ambiguous sign. (In this case, the twist parameter is plus or minus two times the number of parallel components.)

This definition is independent of the choice of arrangement of the the X to be fixed by R. It is immediate that all conditions in Definition 15 are satisfied.

We could also define the twists t_i by comparing $X \cap A_i$ to the curve H giving the hexagonal decomposition in the same fashion. The one additional complication is that it is then not always possible to pick $X \cap Y_j$ to be transverse to $H \cap Y_j$ while being fixed by R: if $X \cap Y_j$ has an odd number of strands running between two different boundary components, the middle one will necessarily agree with $H \cap Y_j$. When computing the intersection numbers in the annulus, treat two curves with the same endpoint as having half an intersection, with signs as appropriate.

4.1. **Generalized coordinates.** We will also consider more general oordinates, generalizing both triangulations and Dehn-Thurston coordinates. A *maximal decomposition* $\mathcal{D}$ of a surface Σ (possibly with punctures) is a maximal collection of non-parallel arcs or loops, with the endpoints of the arcs at punctures. Such a decomposition breaks the surface into pieces which are

- Triangles;
- Pairs of pants; or
- Cusped annuli: an annulus with a puncture on one boundary component.

As with pairs of pants decompositions, in addition to a decomposition we will need some more marking to completely specify the coordinates. A *marking* of a generalized decomposition $\mathcal{D}$ is a choice of a dual curve D_A for each loop $A \in \mathcal{D}$, so that $i(D_A, A) = 2$ and $i(D_A, B) = 0$ for every other arc or loop $B \in \mathcal{D}$. Concretely, when A is the boundary of two distinct pairs of pants in $\Sigma \setminus \mathcal{D}$, D_A is a is a minimally intersecting curve as before; when A appears twice on the boundary of a single pair of pants, D_A is either two parallel copies of a minimally intersecting curve or a single curve halfway in-between two such; when A is on the boundary of one pair of pants and one cusped annulus, D_A is an arc running from the puncture at the cusp of the annulus, over the pair of pants, and back to the cusp; and when A is on the boundary of two cusped annuli, D_A is a pair of disjoint arcs between the two cusps, each minimally intersecting A. These two arcs will either be parallel or at distance 1. In all of these cases, we can define the twist parameter associated to A by

$$tA = T(A; X, D_A)$$

where the twisting T is defined below.

5. Twisting of curves

In this section we will put the definition of the twist parameters above in a different and more universal light.

We first recall the notion of *multiplication of curves*, due to Luo [7].

Definition 17 (Luo)**.** Suppose A and B are two unoriented arcs intersecting transversally at a point x in an oriented surface. The *resolution of $A \cup B$ at p from a to b* is the same as $A \cup B$ away from a neighborhood of p and, in the neighborhood, turns left from A to B, as in Figure **??**.

Definition 18 (Luo)**.** Given $A, B \in \mathcal{SC}(\Sigma)$, the *product* AB is the multi-curve obtained by taking the taut embedding of $A \cup B$ and resolving all intersection points from A to B, as in Figure **??**.

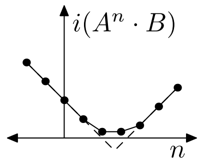
Among the elementary properties of the product, Luo showed

- AB is again a simple curve (with no forbidden components).
- If $A \cap B \neq \emptyset$, then $(AB)A = A(BA) = B$.
- Although the product is not associative in general, the expression $A^k B$ is independent of the order of multiplication.
- If A is a loop and $k = i(A, B)$, than $A^k B = D_A(B)$, where $D_A : \Sigma \to \Sigma$ is the Dehn twist around A.
- If $A, B \in \mathcal{SC}_0(\Sigma)$, then $i(AB, C) \leq i(A, C) + i(B, C)$.

In light of these facts, it makes sense to define $A^k B$ for $k < 0$ to be BA^{-k} if $i(A, B) \neq 0$. If $i(A, B) = 0$, we define $A^k B$ to be $A^{|k+l|}B'$, where $B = B' \cup A^l$ and B' has no components parallel to A.

A more substantial fact is

Theorem 2 ([7])**.** *Given $A, B \in \mathcal{C}_0(\Sigma)$ and $C \in \mathcal{C}(\Sigma)$, the intersection number $i(A^n B, C)$ is a convex function of $n \in \mathbb{Z}$.*



Thus $i(A^nB, C)$ as a function of n is a convex, integer-valued function, with slope bounded in absolute value by $i(A, C)$. In fact for $n \gg 0$ the slope is equal to $i(A, C)$, and for $n \ll 0$ the slope is equal to $-i(A, C)$. (This is proved in Proposition 22.) As a result, $i(A^nB, C)$ looks like the function at right as a function of n. The idea behind the twist parameter is to find the position of the extrapolated minimum.

Definition 19. If A is a simple closed curve and $B, C \in \mathcal{C}_0(\Sigma)$, with $i(A, C) \neq 0$, the *twisting of B and C around A* is

$$T(A; B, C) = \lim_{n\to\infty} \frac{i(A^{-n}B, C) - i(A^nB, C)}{2}.$$

If $i(A, C) = 0$, let k be the number of components of C parallel to A, and set $T(A; B, C) = \pm k i(A, B)$. (The sign is defined to be indeterminate if $i(A, B) = 0$ or $i(A, C) = 0$.)

Proposition 20. *For A a loop and $B, C \in \mathcal{SC}_0(\Sigma)$,*

$$T(A; B, C) = -T(A; C, B).$$

Proof. This is clear from the definition if $i(A, B) = 0$ or $i(A, C) = 0$. (For instance, if B has l components parallel to A, then $T(A; B, C) = -li(A, C)$.) Otherwise, let $k = i(A, B)$ and $l = i(A, C)$. Then we can restrict n to be a multiple of k in the limit defining the twisting to get Dehn twists:

$$\begin{aligned} T(A; B, C) &= \lim_{n\to\infty} i(A^{-nk}B, C) - i(A^{nk}B, C) \\ &= \lim_{n\to\infty} i(D_A^{-n}(B), C) - i(D_A^{n}(B), C) \\ &= \lim_{n\to\infty} i(B, D_A^{n}(C)) - i(B, D_A^{-n}(C)) \\ &= \lim_{n\to\infty} i(A^{nl}C, B) - i(A^{-nl}C, B) \\ &= -T(A; C, B). \end{aligned}$$

□

The twisting gives the position of the extrapolated minimum of the function $i(A^nB, C)$. To completely determine the function, we also need to know the height of that extrapolated minimum, which is given by cutting open the surface at A.

Definition 21. For $A \in \mathcal{SC}_0(\Sigma)$, NA is a regular neighborhood of A and $\Sigma \mid A$ is Σ minus the interior of NA. For $B \in \mathcal{SC}(\Sigma)$, $B \mid A \in \mathcal{SC}(\Sigma \mid A)$ is the restriction of the taut embedding of $B \cup A$ to $\Sigma \mid A$, with any components parallel to A removed.

Proposition 22. *For any A a simple closed curve, $B, C \in \mathcal{C}_0(\Sigma)$ and any n with $|n|$ sufficiently large,*

$$i(A^n B, C) = |i(A, C)n - T(A; B, C)| + i(B \mid A, C \mid A).$$

Proof. The proposition is immediate if $A \cap B = \emptyset$ or $A \cap C = \emptyset$. Otherwise, embed $\Sigma \mid A$ into Σ as the complement of a neighborhood NA of$A$. Arrange $B, C \subset \Sigma$ so that $i(B \mid A, C \mid A)$ is achieved in $\Sigma \mid A$. Orient both $B \cap NA$ and $C \cap NA$ from one boundary component to the other. (Note that this will usually not agree with an orientation of all of B or C.) Let $s_A(B, C)$ be the *signed* number of intersections between B and C in NA. By the definitions,

$$s_A(AB, C) = s_A(B, C) + i(A, C).$$

For $k \gg 0$ (resp. $k \ll 0$), all of the intersections between $A^k B$ and C in NA are positive (resp. negative). In either case $A^k B \cup C$ is taut by Lemma 23 below, and

$$i(A^k B, C) = |s_A(A^k B, C)| + i(B \mid A, C \mid A),$$

from which we see that $t(A; B, C) = s_A(B, C)$ and get the statement of the lemma. □

Lemma 23. *Let A be a loop on Σ, and let $B \in \mathcal{SC}_S(\Sigma)$ and $C \in \mathcal{SC}_T(\Sigma)$, with S and T disjoint. If B and C are arranged so that the intersections between $B \mid A$ and $C \mid A$ are minimal and every intersection between B and C in NA is positive (resp. negative), then the number of intersections between B and C on all of Σ is also minimal.*

Proof. If the number of intersections between B and C is not minimal, then there is necessarily an embedded bigon between the two. This bigon cannot be entirely in $\Sigma \mid A$, since B and C are minimal on this subsurface. If either corner of the bigon is in $\Sigma \mid A$, then we see a boundary triangle between $B \mid A$ and $C \mid A$, which we could compress and reduce the number of intersections, contradicting minimality. If both corners of the bigon are in NA, then the intersections at the two corners have opposite signs, again a contradiction. □

5.1. **Comparing twisting.** This section is not necessary for the remainder of the paper. It may help understand the twisting parameters or lead to better proofs below. In some sense, the twisting $T(A; B, C)$ behaves like a homological intersection number.

Coboundary is not zero, but is bounded in some sense.

Define $T(A; B, C, D)$; measures "clockwiseness" of triple $(B \mid A, C \mid A, D \mid A)$ around the boundary A.

5.2. **Twisting around multiple curves.** In Section 7 we will need a more general notion of twisting, where the curve we twist around has more than one component.

Definition 24. Let $\mathcal{A}$ be a collection of signed loops: a set of disjoint loops $A \subset \Sigma$, each with an associated sign $\epsilon(A) \in \{\pm 1\}$. Let

$$\mathcal{A} \cdot B := (\prod_{A \in \mathcal{A}} A^{\epsilon(A)}) \cdot B$$

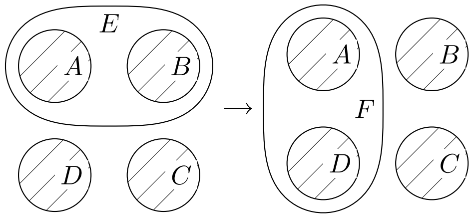
The *(multi-)twisting of B and C around A* is

$$T((A); B, C) = \lim_{n \to \infty} i(\mathcal{A}^n B, C) - i(\mathcal{A}^{-n} B, C).$$

6. Fundamental pants move

We will now use the theory we have developed for one of the principal goals of this paper: to find the change in Dehn-Thurston coordinates under an elementary pants move. There are two elementary pants moves, one taking place on a four-punctured sphere and the other one a once-punctured torus; we will consider the first case here and defer the other to Appendix A.

Consider a four-punctured sphere, with boundary components labelled A, B, C, and D. We wish to compare Dehn-Thurston coordinates with respect to two different pants decompositions, given by additional curves E and F that intersect twice (the minimum possible), as shown below.



This should be thought of as part of a larger pants decomposition of Σ. We need to find the new measure of F (Lemma 25), the new twist of F (Lemma 27), and how the twists of the boundary curves change (Lemma 28).

We will write mE and tE, etc., for the measure and twist coordinates of X with respect to E.

Lemma 25. *With the setup as above, mF is given by*

$$mF = \max(mA + mC - mE, mB + mD - mF,$$
$$|tE| + g(mA, mB; mE) + g(mC, mD; mE))$$

where

$$g(x, y; z) = \max(0, x - z, y - z, (x + y - z)/2).$$

Proof. Consider the curve $E \cup F$. This curve has two intersections and a total of four smoothings. By the Smoothing Lemma 7, $i_X(E \cup F) = mE + mF$ is the maximum of the intersections of these four smoothings. Two of these smoothings are the products FE and EF. If either of these two smoothings gives the maximum, then at $k = 0$ we are in the stable range of the function $i_X(E^k F)$ as in Proposition 22, and by that proposition we have

$$i_X(F) = |T(E; F, X)| + i(F \mid E, X \mid E).$$

The first term is $|tE|$. The second term is itself a sum of two terms, one from each of the two pairs of pants that we get by splitting along E. Let P be the upper pair of pants. There are four different topological possibilities for the arcs in $X \cap P$. The intersection $i(F \cap P, X \cap P)$ is in any case a function of mA, mB, and mE; let us call that function $g(mA, mB; mE)$. It's an easy computation that

$$g(x, y; z) = \begin{cases} 0 & z \geq x + y \\ x - z & x \geq y + z \\ y - z & y \geq x + z \\ \frac{x+y-z}{2} & \text{all triangle inequalities satisfied.} \end{cases}$$

Furthermore, $g(x, y; z)$ is in fact a maximum of these four terms. Similarly for the lower pair of pants.

If $i_X(E \cup F)$ is not equal to either of these two smoothings (giving EF and FE, respectively), then it is equal to the maximum of the other two smoothings. These other two smoothings each give us two components, parallel to the bounary curves $A \cup C$ and $B \cup D$, respectively. The formula in the statement follows. □

We can generalize the method used in the proof of Lemma 25 to get another version of the Smoothing Lemma:

Lemma 26. *Let A be a loop in Σ, $B \in \mathcal{C}_S(\Sigma)$, and $X \in \mathcal{SC}_T(\Sigma)$, where S and T are disjoint. Let $\mathrm{Sm}'(A \cup B)$ be the set of all smoothings of $A \cup B$ except for the two products AB and BA. Then*

$$\begin{aligned} i(B, X) = &(-i(A, X) + \max_{B' \in \mathrm{Sm}'(A \cup B)} i(B', X)) \\ &\vee (|T(A; B, X)| + i(B \mid A, X \mid A). \end{aligned}$$

Returning to the elementary move on pants decompositions, we next need to compute tF. To specify tF, we need to fix the orientation-reversing map after the elementary move; fortunately, the original map fixes both E and F, so we can use the same map.

Lemma 27. *With the setup as above, the twist of F is given by*

$$tF = -tE$$

if $mE + mF = (mA + mC) \vee (mB + mD)$, and

$$tF = -\operatorname{sign}(tE)(mF - g(mA, mD; mF) - g(mB, mC; mF))$$

otherwise, where g is as in Lemma 25.

Proof. If we repeat the same elementary transformation again, we switch back from the pants decomposition involving F to the pants decomposition involving E. In particular, if EF or FE is the dominant term in the maximum for $i(E \cup F, X)$, then there is an expression for mE in terms of mF, tF, and the measures of the boundary curves; we can use this to solve for tF. This yields the second case in the statement of the theorem. Note that the sign of tF is opposite to the sign of tE, since $E^1 F = F^{-1} E$.

In the other case, mE does not depend on tF, so we are unable to solve for tF; instead, we will argue indirectly. All the measure and twist parameters we use multiply by k if we replace a curve by k parallel copies of itself. We can use this to extend our formulas to include cases when the measure and twist parameters are rational (by requiring them to behave well under scaling) and then, by continuity, to cases when they are real. As a transformation on this real parameter space, the map from the new to old coordinates must preserve the integral points, those points that come from actual curves, as well as the points whose k-fold multiple is integral. In particular, it must be volume-preserving. In this case, we have $mF(mE) = -mE + (mA + mC) \vee (mB + mD)$; this map itself is volume-preserving. This implies in turn that the twist parameter tF must be $\pm tE$ plus a function of mE and the boundary measures. The action of the orientation-reversing map defining the twist parameters negates both tE and tF and commutes with the transformation from the E parameters to the F parameters; thus $tE = 0$ must correspond to $tF = 0$ regardless of mE and the other measures, and we have $tF = \pm tE$. Finally, if we increase tE enough while keeping the other parameters constant, we eventually arrive in the other case, in which tF has the opposite sign to tE. By continuity, we must therefore have $tF = -tE$. ☐

Finally, we need to find how the measure and twist parameters for the other curves in the pants decomposition change.

Lemma 28. *With the setup as above, the change in the twists of the boundary curves are given by*

$$\begin{aligned}\Delta tA = tA' - tA &= \operatorname{sign} tE \times l(l_A, |tE|; l_D)\\ \Delta tB &= \operatorname{sign} tE \times l(l_B, |tE|; l_C)\\ \Delta tC &= \operatorname{sign} tE \times l(l_C, |tE|; l_B)\\ \Delta tD &= \operatorname{sign} tE \times l(l_D, |tE|; l_A)\end{aligned}$$

where

$$\begin{aligned}l(x, y; z) &= 0 \vee \left(\frac{x + y - z}{2} \wedge x \wedge y\right)\\ l_A &= l(mA, mE; mB)\\ l_B &= l(mB, mE; mA)\\ l_C &= l(mC, mE; mD)\\ l_D &= l(mD, mE; mC).\end{aligned}$$

If two of these boundary curves happen to be the same curve in the pants decomposition of Σ, there will be two contributions to the change in its twisting parameter.

Note that l_A, …, l_D are the numbers of strands of X running from the corresponding boundary circle to E in the inital pants decomposition.

Proof. Use the definition of the twist number that involves counting intersection in the annulus with the hexagonal-decomposition curve H, where both X and H are arranged to be fixed by the orientation-reversing map R on the pairs of pants. For the curve A, this amounts to specifying two opposite basepoints on the boundary of the annulus around A: the two places where H meets this boundary. One of these basepoints is initially in the middle of the possibly empty bundle of strands running from A to E. After the pants move, the new basepoint is in the middle of the bundle of strands from A to D; we need to find how many strands apart these basepoints are on the boundary of the annulus. Note that this can be an integer or half-integer. The change in the twist is twice this distance, since H intersects A twice. There are $l_A = l(mA, mE; mB)$ strands from A to E. The strands from A to D must first intersect E, so will be a subset of these strands, as well as a subset of the l_D strands from E to D. These two bundles of strands are shifted relative to each other by $tE/2$ positions. There are four cases to consider:

- $l_D > l_A + |tE|$: In this case every strand from A to D continues to E, and $\Delta tA = 0$.
- $l_A > l_D + |tE|$: In this case the strands from A to D are the ones in the bundle from E to D. The center of this bundle is offset by $tE/2$ from the center of the AE bundle, so $\Delta tA = tE$.
- $|tE| > l_A + l_D$: In this case there are no strands from A to D, and the new location of H is at the far left or right of the original AE bundle, which is $l_A/2$ positions from the center, so $\Delta tA = \operatorname{sign} tA \times l_A$.
- Otherwise all triangle inequalities are satisfied, and the AE bundle and the ED bundle partially overlap on both ends. A short computation shows $\Delta tA = \operatorname{sign} tE \times (l_A + |tE| - l_D)/2$.

These four cases correspond to the four cases in the definition of $l(x, y; z)$ as in the statement of the lemma. □

Remark. The second appearance of the function $l(x, y; z)$ is not a coincidence: the hexagons of the hexagonal decomposition behave similarly to the "hexagon" that travels from A to E, along E for a distance of $tE/2$, from E to D, along D for a distance of $-\Delta tD/2$, from D back to A, and along A for a distance of $-\Delta tA/2$. Okai [10] works this out in the hyperbolic setting.

Remark. An alternate, more symmetric, formula for ΔtA (etc.) is

$$\Delta tA = \operatorname{sign} tE \times \Big(0 \vee \big(|tE| \wedge |tF| \wedge \frac{mE + mA - mB}{2} \wedge \frac{mF + mA - mD}{2} \wedge \frac{mE + mF - mB - mD}{2}\big)\Big).$$

Perhaps this formula points to some more structure yet to be discovered.

7. Convexity of length functions

Theorem 3. *Let Σ be a surface with a triangulation (resp. a pair of pants decomposition), and let Y be a curve on Σ, with intersections, multiple components, and ends on punctures. Then for $X \in \mathcal{SC}(\Sigma)$, $i(Y, X)$ is a convex, piecewise-linear function in the coordinates on $\mathcal{SC}(\Sigma)$.*

The fact that $i(Y, \cdot)$ is a piecewise-linear function is well-known; the new ingredient in this theorem is the convexity. Note that a piecewise-linear function is convex if it is the maximum of (at most countably many) linear functions.

This theorem gives some additional structure on $\mathcal{SC}(\Sigma)$, beyond its piecewise-linear structure: there is a distinguished class of convex functions.

Question 29. Let f be a piecewise-linear function on $\mathcal{SC}(\Sigma)$ which is convex with respect to *every* triangulation (resp. pair of pants decomposition) of Σ. Is f necessarily the intersection number with a measured lamination on Σ? If we further assume that f has only a finite number of parts, is f the intersection number with a weighted collection of curves?

Sherman and Zelevinsky [13, Theorems 1.6–1.7] show that a similar phenomenon holds for the cluster algebras related to some simple surfaces.

The proof of Theorem 3, as is common for surfaces, is substantially simpler in the triangulation case as opposed to the case with no boundary.

Theorem 3, triangulation case. Put Y in taut position. By Lemma 12, we can write $i(Y, X)$ as the maximum of $i(Y', X)$, where $Y' \in \mathrm{Sm}(Y)$. Thus we may assume that $Y \in \mathcal{SC}(\Sigma)$.

Let $\mathcal{T}$ be the set of edges of the triangulation.

We will argue by induction on the total weight $w(Y)$ of Y, defined by

$$w(Y) = \sum_{E \in T} i(Y, E).$$

If $w(Y) = 0$, then $i(Y, E) = 0$ for all edges E and Y is a a union of parallel copies of edges; $i(Y, X)$ is a linear function of the coordinates, and so is convex.

Otherwise, let E be an edge so that $i(Y, E) \neq 0$. Then $Y \cup E$ has at least one intersection and

$$i(Y, X) = -i(E, X) + \max_{Y' \in \mathrm{Sm}(Y \cup E)} i(Y', X)$$

Since each smoothing Y' of $Y \cup E$ has fewer intersections with E than Y does (and no more intersections with other edges), each Y' has lower weight than Y, we are done by induction. □

Theorem 3, general case. As in the proof in the triangulation case, we can assume that Y is a simple curve and that Y does not intersect any arcs. This implies that Y does not intersect any cusped annuli and is contained in the pairs of pants in the decomposition of Σ.

Also as before, we may assume that Y intersects at least one ... □

Appendix A. Elementary moves

In this section we give a variety of elementary moves, for triangulations, pants decompositions, and a mixture of the two.

Appendix B. Comparison to Penner's coordinates

Penner [11, 12] also gave formulas for the change of Dehn-Thurston coordinates under an elementary pants move, but he used a somewhat different choice for the twisting parameter: his twisting parameter is always an integer with no restriction on the parity, but as a result it is not contravariant under orientation reversal. For concreteness, we give here (without proof) the relation between the two sets of coordinates.

Let t_P be Penner's version of the coordinates, and let t_B be the balanced coordinates as given in this paper (i.e., coordinates so that orientation reversal negates the twist). Suppose that two pairs of pants have respective bounding curves (A, B, E) and (C, D, E) in clockwise order. (Some of these may be repeated.) Then the twists on E are related by

$$t_P = \frac{t_b + l(mB, mE; mA) + l(mD, mE; mC) - mE}{2} \tag{1}$$

where $l(mA, mB; mC)$ is the number of strands running from A to C on a pair of pants with boundary (A, B, C):

$$l(x, y; z) = 0 \vee \left(\frac{x + y - z}{2} \wedge x \wedge y \right).$$

These formulas also encode the parity restrictions on t_b: t_b is allowable iff the division in (1) comes out evenly.

References

[1] *Travaux de Thurston sur les surfaces*, Astérisque, vol. 66, Société Mathématique de France, Paris, 1979, Séminaire Orsay, With an English summary.
[2] Max Dehn, Lecture notes from Breslau, 1922, the archives of the University of Texas at Austin, translated by John Stillwell.
[3] Sergey Fomin and Andrei Zelevinsky, *Cluster algebras I: Foundations*, J. Amer. Math. Soc **15** (2002), no. 2, 497–526, arXiv:math.RT/0104151.
[4] Michael Freedman, Joel Hass, and Peter Scott, *Closed geodesics on surfaces*, Bull. London Math. Soc **14** (1982), no. 5, 385–391.

[5] Joel Hass and Peter Scott, *Intersections of curves on surfaces*, Israel J. Math. **51** (1985), no. 1–2, 90–120.
[6] ———, *Shortening curves on surfaces*, Topology **33** (1994), no. 1, 25–43.
[7] Feng Luo, *Some applications of a multiplicative structure on simple loops in surfaces*, Knots, braids, and mapping class groups—papers dedicated to Joan S. Birman (New York, 1998), AMS/IP Stud. Adv. Math., vol. 24, Amer. Math. Soc., 2001, pp. 123–129.
[8] Max Neumann-Coto, *A characterization of shortest geodesics on surfaces*, Algebraic and Geometric Topology **1** (2001), no. 17, 349–368, arXiv:math.GT/0106200.
[9] Takayuki Okai, *An explicit description of the Teichmüller space as holonomy representation and applications*, Hiroshima Mathematical Journal **22** (1992), no. 2, 259–271.
[10] ———, *Effects of a change of pants decompositions on their Fenchel-Nielsen coordinates*, Kobe J. Math. **10** (1993), no. 2, 215–223.
[11] Robert C. Penner, *A computation of the action of the mapping class group on isotopy classes of curves and arcs in surfaces*, Ph.D. thesis, Massachusetts Institute of Technology, June 1982.
[12] ———, *The action of the mapping class group on curves in surfaces*, L'enseignement Mathématique **30** (1984), 39–55.
[13] Paul Sherman and Andrei Zelevinsky, *Positivity and canonical bases in rank 2 cluster algebras of finite and affine types*, Moscow Mathematics Journal **4** (2004), no. 4, 947–974.
[14] William P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bulletin of the American Mathematical Society **19** (1988), no. 2, 417–431, Circulated as a preprint in 1976.

Department of Mathematics, Barnard College, Columbia University, MC 4436 New York, NY 10027, USA
E-mail address: dthurston@barnard.edu