On Self-Overlapping Curves, Interior Boundaries, and Minimum Area Homotopies

Parker Evans

Department of Mathematics, Tulane University
New Orleans, LA, USA. pevans2@tulane.edu

May 2018

Abstract

In this thesis, we study self-overlapping (SO) curves and minimum area homotopies in the plane. The latter is a recently introduced construct and is intimately connected with self-overlapping curves. Here, we study the *direct splits* of curves, the closed subcurves which follow the orientation of the curve from the first appearance of a vertex v to its second and final appearance. We use the direct splits as a tool to simplify the study of a curve. In aid to this approach, we introduce so-called *simple subcurve decompositions*, formed by iteratively removing loops from the curve to completely decompose the curve into simple subcurves. These simple subcurve decompositions naturally yield auxiliary curves, we call *elementary forms*, which are much easier to understand than the original. These elementary forms are particularly useful for studying the Whitney index, or turning index, of the direct splits of a curve. In fact, in the *elementary* setting, where no intersection points are linked, we show that whitney indices of the direct splits can be used to completely determine the self-overlapping-ness of a plane curve. More specifically, we show that the condition of being *top-heavy*, when $whit(\gamma_i) \leq 0$ for each proper direct split γ_i, is necessary and sufficient for an elementary curve γ with positive outer basepoint and $whit(\gamma) = +1$ to be self-overlapping. We then show that half of this holds in general. One of our main theorems tells us that a top-heavy curve γ with positive outer basepoint and $whit(\gamma) = +1$ is self-overlapping. Surprisingly, this follows as a corollary to a more general theorem, which says that any curve γ with positive outer basepoint, $whit(\gamma) = +1$ and no self-overlapping direct splits is SO. Additionally, we define a new operation 'wrap', $Wr_+(\gamma)$, which adds a positive Jordan curve around the entire curve, connected to the curve at its positive outer basepoint. Our final main result is that given any plane curve γ with positive outer basepoint and n vertices, there is an integer $k \leq n + 1$ so that $Wr_+^k(\gamma)$ is an interior boundary.

Acknowledgements

I would like to thank my advisor Carola Wenk for her enthusiastic encouragement and support over the past two years on this project. I would also like to thank Brittany Fasy, with whom partial collaboration of this work was done. Finally, I would like to thank my other thesis readers Rafal Komendarczyk and Anastasia Kurdia. The author of this thesis was partially supported by the National Science Foundation grant CCF-1618469 and by a Goldwater Scholarship from the Goldwater Foundation.

Contents

List of Figures **vi**

1 Introduction **1**
1.1 Minimum Homotopy Area 1
1.2 The History of Self-Overlapping Curves 3
1.3 New Results 4

2 Preliminaries **5**
2.1 Regular and Generic Curves 5
2.2 Combinatorial Relations Between Intersection Points 6
2.3 Homotopies and Homotopy Moves 7
2.4 Direct Splits 8
2.5 Decompositions 10
2.6 Minimum Homotopies 11
2.7 Interior Boundaries 12
2.8 Wraps 14

3 Properties of SO Curves **15**

4 Simple Subcurve Decompositions **19**
4.1 Definition and Existence 19
4.2 Algebraically Computing Simple Subcurve Decompositions 21
4.3 Relation to Seifert Decompositions 23

5 Classes of Curves **24**
5.1 Good and Ideal Curves 26
5.2 Basic k-Interior Boundaries 28

6 Elementary SO Curves **30**
6.1 G-R Trees 31
6.2 Top-Heavy and Irreducible Curves 35
6.3 Classification of Elementary SO Curves 36

7 Direct Splits of SO Curves **38**
7.1 Arbitrarily Large Interior Boundaries on SO Curves 38
7.2 Arbitrary Negative and Positive SO Direct Splits on SO Curves 40

8 Top Heavy Curves **42**
8.1 Top-Heaviness: Local vs. Global 42
8.2 Properties of Globally Top-Heavy Planar Images 43
8.3 The Main Theorems 48
8.4 Balanced Loop Insertion 53

9 Conclusion **56**

10 References **58**

A Appendix **59**

List of Figures

1	A map with a noisy input for a path.	2
2	A SO plane curve γ. Positive vertices shown in green, negative vertices in red, and the positive outer basepoint in black. The signed intersection sequence is shown below. The winding numbers of each face are enclosed by circles. The index of each vertex is shown as well.	6
3	All three Titus moves and their reversals. As shown, a-moves destroy vertices, b-moves create vertices, and III-moves do not create or destroy vertices. Figure from [10].	8
4	A curve γ and a direct split decomposition $\Omega = \{\gamma_u, \gamma_v, C\}$.	11
5	Figure from the program written by the author. A SO decomposition of a plane curve from a minimum homotopy H. The different SO subcurves are shown in different colors, with the anchor points shown in purple.	13
6	A curve γ and its positive wrap $Wr_{+}(\gamma)$. The (unsigned) intersection sequences of both curves are shown. Note that as $\gamma(0)$ becomes the first vertex on the wrap, we must smooth to make the curve regular and normal again.	15
7	A curve γ with positive outer point $\gamma(0)$ along with its transformation into $Wr_{-}(\gamma)$. First, we do a I_b move, then we wrap after forcing a negative outer basepoint. The signed intersection sequences of both $\gamma, Wr_{-}(\gamma)$ are shown. Positive vertices are in green, negative vertices in red, and basepoints in black.	16
8	A blank cut on a small SO curve, which simplifies the curve immediately. The simple curve γ_1 produced by the cut is shown as well as the region R removed from the interior by the cut.	17
9	Outwards and inwards loops on a curve γ. The edges adjacent to the basepoint are labeled in order as they appear on γ. Note that the edges e_1 and e_4 lie outside the interior of the loop in the case of an outwards loop and inside the interior in the case of an inwards loop.	19
10	A curve γ with a (standard) simple subcurve decomposition $\Omega = \{\gamma^{\tilde{a}}, \gamma^{\tilde{b}}, \gamma^{\tilde{c}}, \gamma^{\tilde{d}}, \gamma^{\tilde{e}}\}$ and its elementary form γ_Ω.	23
11	Figure from [9]. An uncrossing move shown above and two Seifert decompositions of a plane curve shown below.	24
12	A connected sum of γ and γ'. Sample paths P, P_L, P_R are shown along the original curves. A representative of $\gamma \# \gamma'$ is shown below.	25
13	A good curve.	27
14	An ideal curve, with its minimum homotopy decomposition. The anchor points are shown in purple, with each SO curve in the SO decomposition shown in a different color. The sign-changing vertices with indices $1, 2, 5, 8, 9, 12$ can all be seen to be anchor points.	28
15	Lattice of curve classes, with inclusions going upwards.	29
16	An elementary 5-boundary with its corresponding elementary G tree. Note that post-order traversal of T yields the (unsigned) intersection sequence of γ.	30

17 An irreducible G-R tree T with $k = 3$. The depths of the nodes are enclosed by rectangles. When nontrivial, the order preferences on the edges are labeled in blue. The elementary curve γ associated to T through the map ϕ is shown by its side. Note that post-order traversal of T yields the signed intersection sequence of γ. . . . 33
18 The tree T_3 and its corresponding elementary SO curve γ_3. The direct split of the 4^{th} vertex is an elementary -3 boundary. 39
19 The tree T_3 from Proposition 7.3 and its associated elementary SO curve γ_3. After application of three II_b moves to the right, the direct split $\tilde{\gamma}_+$, seen to be a 3-boundary, is created. 40
20 A small self-overlapping curve γ and its doubling $D(\gamma)$. 41
21 Two representatives γ, γ' of the same planar image so that γ is not top-heavy, but γ is. 43
22 A plane curve γ that is top-heavy, but whose equivalence class $E[\gamma]$ is neither locally nor globally top-heavy. 44
23 Curve C chosen in the proof of Proposition 8.5 part (i). The boundary $\mathscr{B}$ is shown in blue. 46
24 A globally top-heavy planar image along with the cuts P_1, P_2, P_3 from the proof of Corollary 8.6. It is geometrically clear here that $\gamma \cong Wr_+(\gamma_{w_1})\,\#Wr_+(\gamma_{w_2})\#Wr_+(\gamma_{w_3})$. 48
25 A top heavy plane curve with its simple path decomposition shown. The intersection sequence is also shown to the side, with the intersection sequence broken into the the simple path decomposition as well. 49
26 The combinatorial structure necessary to apply balanced loop deletion - a wrapped curve, with outer wrap α and a negative loop γ_- as first loop in the simple path decomposition. 51
27 A locally irreducible curve γ with positive outer basepoint and $whit(\gamma) = +1$ that is not top-heavy. The perpetrator here is the direct split γ_v which has $whit(\gamma_v) = +1$ but is not SO. Note that γ_v is locally top-heavy, but that its basepoint is not outer. 53
28 Balanced loop insertion on a SO curve γ with respect to an edge e. Note that the curve produced, γ'', is also SO. 54
29 Global balanced loop insertion applied to a non-SO curve γ, seen by its empty positive loop. 56

1 Introduction

In this thesis, we study self-overlapping curves, their generalizations, interior boundaries, and so called minimum homotopy area, as defined in [6]. We proceed with the goal of showing the intimate connections between these three mathematical objects. Let us begin by providing motivation for studying minimum homotopy area and self-overlapping curves.

1.1 Minimum Homotopy Area

Imagine you have a GPS trajectory that does not quite match the roads on a map and you want to know what path you should *actually* take. Or, you have a computer vision system and you want the system to interpret its visual data and know what it is *actually* looking at. Both of these problems can be solved by a robust method of curve comparison, a way to explicitly measure similarity between curves. Additionally, curve matching allows us to determine similarity between vector graphics images, which are represented as a collection of points and curves, and even to aid searches in medical imaging [1, 2]. The minimum homotopy area metric offers the most robust way to date compare curves, effectively aiding solutions to all the aforementioned problems.

A curve in the plane is a continuous map $\gamma : [0, 1] \to \mathbb{R}^2$. The Fréchet metric, a modern standard for comparing curves, can be computed using *homotopy*, a continuous deformation of one curve into another. Formally, a homotopy from a curve α to a curve β is a continuous map $H : [0, 1] \times [0, 1] \to \mathbb{R}^2$, such that $H(0, s) = \alpha(s)$ and $H(1, s) = \beta(s)$. We can view each curve $H(i, \cdot)$ for $0 < i < 1$ as an intermediate curve, in the process of continuously bending or morphing α into β.[1]

The Fréchet distance is the maximum 'width' of a homotopy between two curves, the maximum length of $H(\cdot, s)$ for $s \in [0, 1]$, after accounting for possible monotonic reparametrizations of the curves. There also exist methods to compute the perpendicular measure of the maximum 'height' of a homotopy, the maximum length of $H(t, \cdot)$ for $t \in [0, 1]$ [5]. While the Fréchet metric can be computed in polynomial time for polygonal curves, it remains highly sensitive to *noise*, unwanted perturbations. A noisy input curve, such as the one shown in Figure 1, may have a high Fréchet distance to a curve that it is actually quite similar to. Such scenarios are commonplace today, in cases when GPS signals are disturbed by large objects in one's surrounding area, like buildings in a city. Modern variations on the Fréchet distance exist, to quell such problems, but different continuous methods of comparison seem much more promising. Better than width or height, we can actually measure the minimum *area* swept out by a homotopy between two curves. Intuitively, the area is a more robust measure than both the width and the height of a homotopy, since perturbations have less impact on area. Surprisingly and conveniently, the area of the homotopy is often actually easier to compute than the width or height [6].

Suppose we wish to homotope a plane curve α to β, where $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$. An important thing to note is that instead of morphing α to β through a homotopy H, we can contract the closed concatenation $\alpha * \bar{\beta}$ to a point, where $\bar{\beta}(t) = \beta(1 - t)$ is the reversal of β. These two problems are equivalent. Thus, it suffices to work with contracting closed curves in the plane to their basepoints. We call homotopies that sweep the minimum possible area *optimal*.

A polynomial algorithm exists to compute the area of an optimal homotopy collapsing a concatenation of two *simple*, non self-intersecting, curves on a 2-manifold [5]. The restriction to curves

[1] It is standard to assume that the initial points $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$ remain fixed on all intermediate curves of the homotopy. There exists a generalization called a *free homotopy* where this assumption is abandoned.

Figure 1: A map with a noisy input for a path.

that are a concatenation of two simple curves is quite limiting. Most curves one encounters in practice will likely not meet this requirement. Thus, to generalize this work, here we study the problem of computing optimal homotopy area for closed curves with only the prescription that they have a finite number of self-intersections. This condition is much less stringent than being the concatenation of two simple curves.

In [10, 13], a worst-case exponential algorithm for computing the optimal homotopy area is presented. In this thesis, we will show a few classes of curves for which the minimum homotopy area reaches its minimum possible value. Structural insights were provided in [13] for computing minimum area homotopies, but the goal of finding a general polynomial time algorithm that handles all cases still remains.[2]

A polynomial time algorithm for computing minimum homotopy area for closed planar curves is consequently useful for both theory and applications – the proposed algorithm and its extensions would provide the most robust way to date to compare curves, including ones with distinct initial and terminal points, inspire further work on efficient untanglings, and would also be useful for computer vision systems, computer graphics, medical imaging, trajectory matching, and map matching. Curves are so ubiquitous in mathematics that measuring their similarity is not only a natural and longstanding problem, but also a fundamental one. Consequently, it is reasonable to expect other applications will eventually arise.

The main theorems of [10, 13] show that computing minimum homotopy area for a closed planar curve γ boils down to finding an optimal decomposition of γ into SO subcurves. Thus, the study of SO curves is immediately relevant to all the aforementioned applications of minimum homotopy area. Exemplified in [10, 13] is the fact that most insights into SO curves yield corresponding insights into computing the minimum homotopy area. Of course, self-overlapping curves are beautiful mathematical objects with a history of almost a hundred years and are interesting in their own right. While only a few papers exist that focus entirely on SO curves, there is a rather rich history

[2] In his unpublished paper [15], Nie provides a polynomial-time algorithm for computing the minimum homotopy, but missing details have left the work unverified.

to these interesting objects [4, 8, 11, 14, 18, 22]. Here, we provide a brief account of this history to set the stage for the findings of this thesis.

1.2 The History of Self-Overlapping Curves

The story begins with Titus, who provided the first solution to the problem of finding a finite algorithm to determine whether a plane curve γ is SO [22]. For our purposes, we define a plane curve γ as self-overlapping iff there is an immersion $F : D^2 \to \mathbb{R}^2$, a map of full rank, so that $\gamma = F|_{\partial D^2}$. Since any smooth curve C in the plane is the image of an immersion of S^1, a smooth plane curve is *self-overlapping* (SO) exactly when there is an extension of the immersion of S^1 to the whole two-disk D^2. In fact, Titus' algorithm applies to interior boundaries, generalizations of SO curves. Titus' algorithm is geometric. He defines three different possible ways to cut the curve into two smaller pieces, then shows in his main theorem that a curve is an interior boundary iff it can be cut into two smaller interior boundaries by one of these kinds of cuts [22]. His results also hold in the case of SO curves, too; a curve ζ is SO iff there is a Titus cut of type I, II′, or II″ so that the resulting two smaller curves, ζ^*, ζ^{**} are both SO. Thus, Titus' algorithm terminates in a finite number of steps, despite being cumbersome to actually compute. Titus also provides a completely algebraic way to compute his algorithm, based on the (signed) intersection sequence of the curve, an ordered list of the two appearances of the vertices, following the orientation of the curve. Next came Blank, who, in his dissertation, proved that a curve is SO iff there is a sequence of cuts, different than Titus cuts, that completely decompose the curve into simple pieces [4]. Blank additionally constructed a word $w(\gamma)$ for a plane curve γ and showed one can determine the existence of a cut decomposition by looking for algebraic decompositions of the word $w(\gamma)$. Again, we see the interplay of algebra and geometry through SO curves. Marx's work [14] is an extension to the work of Titus and Blank. He employs Blank's approach of assigning and then decomposing the words of curves, but does so instead to determine if a plane curve is an interior boundaries. Additionally, Titus, Blank, and Marx all solved the problem of computing the number of inequivalent immersions for a SO plane curve γ, where two immersions $F, F' : D^2 \to \mathbb{R}^2$ are *equivalent* iff $F' = F \circ H$ where $H : \mathbb{R}^2 \to \mathbb{R}^2$ is a sense-preserving homeomorphism of the plane. Blank's algorithm, unlike Titus', runs in polynomial time. Shor and Van Wyk then expedited Blank's algorithm to $O(n^3)$ time where n is the complexity of the curve [18]. Their algorithm is currently the fastest in existence for determining whether a plane curve is SO. The works of Titus, Blank, Marx, and Shor and Van Wyk all primarily focus on two things: formulating an algorithm to determine whether a curve is self-overlapping and in the case that a curve is SO, computing the number of inequivalent immersions. Eppstein and Mumford then examined a more difficult problem related to SO curves. They showed that it is NP-complete to determine whether a fixed SO curve γ is the 2D projection of an immersed surface $F : M \to \mathbb{R}^3$ where M is a compact two-manifold with boundary [8]. While interesting, their paper has no relevance to this thesis. Finally, there is the wonderful work of Graver and Cargo, who got the last word in on the problem of computing the number of inequivalent immersions for a SO curve. They show how to entirely reduce the problem of computing SO-ness to graph theory. In particular, they show how to construct a plane graph, what they call a *covering graph*, Φ for a SO curve γ so that each face of Φ is mapped homeomorphically onto a face on the planar multigraph $G(\gamma)$ of the curve γ [11]. They then show that a curve is SO iff it has a covering graph and that the number of covering graphs is equal to the number of inequivalent immersions. In a sense, covering graphs are like X-ray vision for SO curve γ, with respect to a specific immersion F: they allow us to explicitly 'see' what the immersion F

is doing to the interior of the disk.

In this thesis, we do not consider the question of the number of inequivalent immersions, for we feel this problem has been sufficiently addressed by now. Instead, we look for combinatorial insights on the structure of SO curves. In particular, we study the *direct splits* of SO curves. The direct split of a curve γ with respect to a vertex v is the closed subcurve traveling along the orientation of γ from the first appearance of v to its second appearance. If $\gamma^{-1}(v) = \{t_v, t_v^*\}$ with $t_v < t_v^*$, then the direct split $\gamma_v = \gamma|_{[t_v, t_v^*]}$. The direct splits are soul of the curve, with respect to homotopy area. In this thesis, we will show, unsurprisingly, that they are also quite useful for studying SO curves.

1.3 New Results

The layout of the thesis is as follows. We begin in Section 2 by providing an unfortunately long but completely necessary set of preliminaries. In Section 3, we state and prove some crucial properties of SO curves. Section 4 focuses on defining and investigating a vital tool for our work, simple subcurve decompositions, which are actually a special case of the SO decompositions that are integral to [10, 13]. We step in a different diretion in Section 5, focusing mostly on minimum area homotopies. In particular, we define a few natural classes of curves through minimum homotopy area and completely characterize them. In a sense, this section is a bit of a digression, but we believe it belongs since many of the arguments of the section employ similar ideas and techniques used elsewhere in the thesis. In Section 6 we study *elementary* SO curves, ones where no intersection points are linked. As we show, these are, in a strong sense, the 'simplest' possible SO curves. The main theorem of this section shows that we can determine if an elementary curve C with $whit(C) = +1$ is SO just by checking the Whitney indices of the direct splits. We call this condition being *top-heavy*, when $whit(\gamma_i) \leq 0$ for all proper direct splits γ_i. In Section 7, we employ the main theorem of Section 6 to show a few surprising results on possible direct splits of SO curves. In Section 8, we begin by studying images of normal curves in the plane, which are top-heavy no matter the choice of positive outer basepoint. We show that these *globally top-heavy* planar images are particularly special, and are quite similar to the elementary SO curves studied in Section 6. In the latter portion of Section 8, we prove our main theorems. We show that half of the main theorem on elementary SO curves holds in general. Specifically, a top-heavy curve γ with positive outer basepoint $\gamma(0)$ and $whit(\gamma) = +1$ is SO. The results of Section 7 show that it is not necessary, in general, for a SO curve to be top-heavy. In fact, far from it. We show that given any SO curve C we can find another SO curve γ such that C is a direct split on γ. The main theorems of Section 8 show a few surprising results. The first is Theorem 8.10, which shows a curve γ with positive outer basepoint and $whit(\gamma) = +1$ is SO if it is top-heavy. This result shows us the surprising fact that we can find a sufficient condition for SO-ness just through the Whitney indices of subcurves, namely the direct splits. In fact, the result on top-heavy curves actually follows from studying a more general class of curves, we call *irreducibles*, curves with no SO direct splits. Most shocking is Theorem 8.8, from which both of these results follow. We prove directly that one can perform an operation called a *wrap*, in which we add a positive outer cycle surrounding the whole curve, which when applied enough times, always eventually turns any plane curve into a positive interior boundary. We conclude Section 8 by showing a similar transformation that always transforms plane curves γ with $whit(\gamma) = +1$ and postive outer point $\gamma(0)$ into a top-heavy, and hence, SO, curve.

Much of this research was enabled by the use of a computer program which can determine SO-ness of a given planar curve, compute the minimum homotopy area for planar curves, and display the SO decomposition associated to the optimal homotopy. The program, developed by the author,

can be found here:
http://www.cs.tulane.edu/~carola/research/code.html

2 Preliminaries

We now lay the necessary groundwork on planar curves, homotopies, self-overlapping curves, interior boundaries, and minimum area homotopies that will be needed for this paper.

2.1 Regular and Generic Curves

In this paper, we work with regular, generic, closed planar curves $\gamma : [0,1] \to \mathbb{R}^2$ with $\gamma(0) = \gamma(1)$. We denote the set of all such curves by $\mathscr{C}$. A curve is *regular* if $\gamma'(t)$ exists for all t and is non-zero; a curve is *generic* (or *normal*) if the embedding has only a finite number of intersection points, each of which are simple crossing points. A point $v \in \mathbb{R}^2$ is a simple crossing point when $\gamma^{-1}(v) = \{t_v, t_v^*\}$ and $\gamma'(t_v), \gamma'(t_v^*)$ are linearly independent. Being normal is a weak restriction, as normal curves are dense in the space of regular curves [23].

Viewing a generic curve γ by its image $[\gamma] \subset \mathbb{R}^2$, we can treat γ as a directed multigraph $G(\gamma) = (V, E)$ embedded in the plane. Here, the ordered list of vertices $V = \{p_0, p_1, \ldots, p_n\}$ correspond to the intersection points $V(\gamma)$ of γ, with $p_0 = \gamma(0)$. An edge (p_i, p_j) corresponds to a simple path along γ between intersection points p_i and p_j. We call a curve *simple* if it has no intersection points. Note that since we opt to include the basepoint as a vertex, $G(\gamma)$ is not 4-regular. We pay particular attention to the basepoint in this thesis, so it will be useful to include the basepoint in the planar embedded multigraph. We notate $|\gamma| = \left|V(\gamma)\backslash\{p_0\}\right|$ for the size of the curve, the number of vertices, not including the basepoint. Given a curve $\gamma : [0,1] \to \mathbb{R}^2$, we can reverse the curve to achieve a new curve $\overline{\gamma}$ defined by $\overline{\gamma}(t) = \gamma(1-t)$.

We call two generic curves γ_1 and γ_2 *combinatorially equivalent* when their induced planar multigraphs are isomorphic; in particular, any two generic curves γ_1 and γ_2 with the same images, $[\gamma_1] = [\gamma_2]$, the same orientations, and basepoints, $\gamma_1(0) = \gamma_2(0)$, are combinatorially equivalent. We write $\gamma \simeq \gamma'$ to indicate combinatorial equivalence. As standard practice, we will not distinguish between combinatorially equivalent curves. Hence, we may refer to a curve just by its image, orientation, and basepoint.

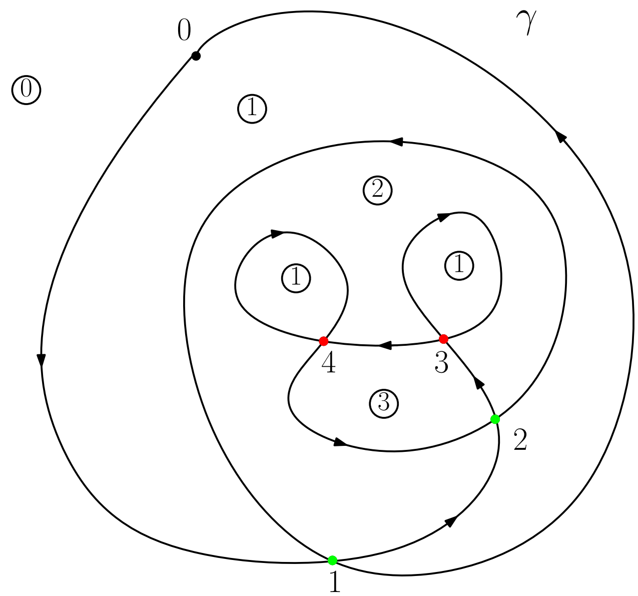
Given a plane curve γ with embedded planar multigraph $G(\gamma)$, we call a point $p \in \mathbb{R}^2$ *outer* to γ iff for any open ball $B_\varepsilon(p)$ we have $B_\varepsilon(p) \cap F_{ext} \neq \varnothing$ where F_{ext} is the exterior face on $G(\gamma)$. Here, we notate the open ball of radius ε about $p \in \mathbb{R}^2$ as $B_\varepsilon(p) = \{x \in \mathbb{R}^2 \,|\, d(x,p) < \varepsilon\}$. We call a point $p \in \mathbb{R}^2$ *inner* to γ otherwise. Note that these definitions also apply to any point p on the image of the curve, $[\gamma]$. If the curve under discussion is clear, we may simply call a point inner or outer. Given a curve γ with outer basepoint $p_0 = \gamma(0)$, the point p_0 will be adjacent to exactly two faces F, F' on $G(\gamma)$, where one of F, F' is the exterior face. Without loss of generality, suppose this is F. Then if $wn(F', \gamma) = +1$, we call p_0 a *positive outer basepoint*. Otherwise, $wn(F', \gamma) = -1$ and we call p_0 a *negative outer basepoint*.

Given an embedded planar multigraph $G(\gamma)$ associated with a generic curve γ, the boundary of each face (the connected components of $\mathbb{R}^2\backslash[\gamma]$) is a collection of edges from $G(\gamma)$. We define the *winding number of x with respect to* γ as $wn(x, \gamma)$ the number of signed crossings in any simple path from x to the exterior face on $G(\gamma)$ that avoids the intersection points of γ. The sign of a crossing between the path and an edge e of $G(\gamma)$ is positive if the path crosses e from left to right, negative otherwise. This number is independent of the path chosen and is constant over each face F of $\mathbb{R}^2\backslash[\gamma]$. Hence, we may write $wn(F,\gamma)$ when convenient. The *winding area* of a curve γ is given by

$$W(\gamma) = \int_{\mathbb{R}^2} \left|wn(x,\gamma)\right| dx = \sum_{F \in G(\gamma)} A(F)\left|wn(F,\gamma)\right|,$$

where $A(F)$ is the area of the face F and $wn(x,\gamma) = 0$ for $x \in [\gamma]$.

The *Whitney index* $whit(\gamma)$ of a curve γ is defined to be the winding number of the derivative γ' about the origin.



intersection sequence

$0, 1_+, 2_+, 3_-, 3_+, 4_-, 4_+, 2_-, 1_-, 0$

Figure 2: A SO plane curve γ. Positive vertices shown in green, negative vertices in red, and the positive outer basepoint in black. The signed intersection sequence is shown below. The winding numbers of each face are enclosed by circles. The index of each vertex is shown as well.

2.2 Combinatorial Relations Between Intersection Points

Following Titus, we now describe how the intersection points interact with each other:

Definition 2.1 (Combinatorial Relations)**.** *Take $\gamma \in \mathscr{C}$ with vertices p_i, p_j reached at $t_i < t_i^*$ and $t_j < t_j^*$, respectively. Whenever it is clear, we may call the i^{th} intersection point p_i just by its index i for brevity. As in [21], we define the three combinatorial relations between i and j as follows:*

(i) i links j: $i \; L \; j \iff t_i < t_j < t_i^* < t_j^*$ or $t_j < t_i < t_j^* < t_i^*$.

(ii) i is separate from j: $i \; S \; j \iff t_i < t_i^* < t_j < t_j^*$ or $t_j < t_j^* < t_i < t_i^*$.

(iii) i is contained in j: $i \subset j \iff t_j \le t_i < t_i^* \le t_j^*$.

Any two intersection points $p_i, p_j \in V(\gamma)$ necessarily have one of these three relations. We define the *sign* of each intersection point p_i, denoted $\text{sgn}(p_i) = \text{sgn}(p_i, \gamma)$, as positive if the vector γ' rotates clockwise from t_i to t_i^* and negative otherwise. As standard practice, we say the rotation is either clockwise or counterclockwise depending on which direction we can rotate $\gamma'(t_i)$ to $\gamma'(t_i^*)$ by an angle of less than π radians. Note that the sign of a vertex $\text{sgn}(p_i)$ is not absolute; it is completely dependent on the basepoint. Two curves γ, γ' with the same image and orientation may have a vertex v with $\text{sgn}(v, \gamma) = -\text{sgn}(v, \gamma')$. See Figure 21.

We now describe how to define the *signed intersection sequence* or *Titus intersection sequence* of a closed curve $\gamma : [0,1] \to \mathbb{R}^2$. Suppose we start at the basepoint $\gamma(0)$, and trace along the curve γ, following its orientation until we reach $\gamma(1) = \gamma(0)$. As we traverse along γ, for the first occurrence of every vertex v we label it with index $i+1$, where i other vertices have been discovered, not including the basepoint. After doing so, let us traverse γ a second time. We can then form the signed intersection sequence as follows. The first time we reach each vertex v, print its index, along with its sign $\text{sgn}(v)$. Then, at the second appearance of v, we add the index of v with opposite sign $-\text{sgn}(v)$ to the signed intersection sequence. Additionally, we prefer to begin and end the intersection sequence with 0, corresponding to the basepoint, so that each consecutive pair of indices i, j in the intersection sequence corresponds to a unique edge e from p_i to p_j. See Figure 2 for an example of a signed intersection sequence.

The signed intersection sequence carries a great deal of information about the curve. In fact, as shown by Theorem 3 in [22], two curves with positive outer basepoint are combinatorially equivalent if they have the same signed intersection sequence. We will write $\gamma \cong \gamma'$ to indicate that γ and γ' have the same signed intersection sequences. Moreover, from the set of all pairwise relations between intersection points, one can recover the (unsigned) intersection sequence [21]. Thus, the set of relations between the intersection points is equivalent to the (unsigned) intersection sequence.

While all of the relations $\{\subset, \text{ L }, \text{ S }\}$ are equally important in the sense that together they are unavoidable, i.e., any two intersection points must relate through one of the three relations, linkage is truly the most important relation. All of the complexity and variety of SO decompositions arise from linkages of SO subcurves, linkage of the basepoints of the SO subcurves. With this in mind, we construct a vital object for the rest of our studies. Many combinatorial conditions for a curve γ are easier to express naturally on its linkage graph.

Definition 2.2. *Given a curve γ, the* ***linkage graph*** $G_L(\gamma)$ *of γ is the graph with vertex set $V(G_L(\gamma)) = V(\gamma)$ and edge set $E(G_L(\gamma)) = \{ij \,|\, i \text{ L } j\}$.*

Similarly, we define the poset $P(\gamma)$ of a curve γ as the poset with vertex set $V(\gamma) \backslash \{p_0\}$ without the basepoint and relation $\subset$. It is not hard to verify that this is a bonafide partial order: it is reflexive, anti-symmetric, and transitive. Here, we do not allow the basepoint $p_0 \iff \gamma(0)$ to be in our poset, for it is trivially a maximal element and will prevent us from counting the number of non-trivial maximal elements.

2.3 Homotopies and Homotopy Moves

A homotopy H between two generic curves γ and γ' is a continuous function $H : [0,1]^2 \to \mathbb{R}^2$ such that $H(0, \cdot) = \gamma$ and $H(1, \cdot) = \gamma'$. In $\mathbb{R}^2$, any curve is null-homotopic, i.e., homotopic to a constant map. Given a sequence of homotopies $\{H_i\}_{i=1}^k$, we notate the concatenation of these homotopies in order as $\sum_{i=1}^k H_i$.

We will use the notation $\overline{H}$ for the reversal $\overline{H}(i,t) = H(1-i,t)$ of a homotopy. Additionally, if $H(0,\cdot) = \gamma$ and $H(1,\cdot) = \gamma'$, and we are only particularly interested in the source and target curves of the homotopy, we may write $\gamma \xrightarrow{H} \gamma'$.

The *Titus moves* are the set of basic local alterations to a curve defined by their action on $G(\gamma)$. Titus moves come as three pairs of moves [10]: The I moves destroy/create an empty loop, II moves destory/create a bigon, and III moves flip a triangle. For I and II moves, we denote the moves that remove intersection points as I_a and II_b, and moves that create intersection points as I_a and II_a. Notice that III moves do not create/add vertices; however, we do relabel the intersection points; see Figure 3.

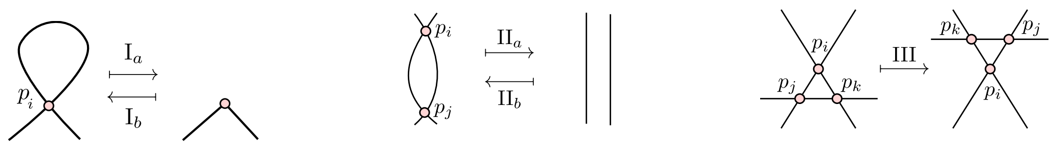


Figure 3: All three Titus moves and their reversals. As shown, a-moves destroy vertices, b-moves create vertices, and III-moves do not create or destroy vertices. Figure from [10].

It is well-known that any homotopy such that each intermediate curve is piecewise regular and generic, or almost generic, can be achieved by a sequence of Titus moves. Thus, without loss of generality, we assume that each event in the homotopy, the times when the curve combinatorially changes, is a single Titus move.

On a directed planar multigraph $G(\gamma)$, we can naturally define the left and right face of any edge e using the right hand rule. As in [10], we call a homotopy *left sense-preserving* if $H(i+\epsilon,t)$ lies on or to the left of $H(i,t)$ for any $i,t \in [0,1]$ any $\epsilon > 0$. We define *right sense-preserving* homotopies in a completely analogous fashion.

2.4 Direct Splits

We now define the central object of this thesis.

Definition 2.3 (Splits)**.** *Let $\gamma \in \mathscr{C}$. Suppose p_i is an intersection point, with $\gamma(t_i) = \gamma(t_i^*) = p_i$ and $t_i < t_i^*$. The **direct split** of γ with respect to p_i is $\gamma_{[t_i,t_i^*]}$, and the **wrap split** of γ with respect to p_i is $\gamma_{[t_i^*,1]} * \gamma_{[0,t_i]}$. Here, $*$ is meant to indicate path concatenation. We call a subcurve **free** if is either a direct split or a wrap split. A direct split will be called **proper** as long as it is not the whole curve.*

Remark: If the index i of the vertex v is known, we may notate the direct and wrap splits on γ with respect to v as γ_i and γ_i^*, respectively. Otherwise, we will notate them as γ_v and γ_{v^*}.

It is worth mentioning that the direct split and wrap splits are not regular at the point v. This irregularity is no problem, though, as we can smooth the splits at the point through an arbitrarily small deformation. We will return to this point shortly in Section 3.

Given a free subcurve $\widetilde{\gamma}$ on a curve γ, we write $\gamma\backslash\widetilde{\gamma}$ for the curve with image $[\gamma]\backslash[\widetilde{\gamma}] \cup \{p\}$ and the same orientation as γ, where p is the basepoint of $\widetilde{\gamma}$. In the case that γ_i is a direct split, we then have $\gamma\backslash\gamma_i$ as combinatorially equivalent to γ_{i^*}.

Note that the splits of p_i are closed subcurves, traveling along the orientation of the curve, each with p_i as the basepoint. While there are many closed subcurves $\tilde{\gamma}$ of γ from a vertex v back to v while following edges of $G(\gamma)$, in a sense, many of these are degenerate. An alternate characterization of free subcurves is that these are the only subcurves on γ which we can contract to their basepoint, without altering the rest of the curve. Indeed, this motivates the name free, for they are 'free' to be plucked off the curve, without affecting the curve outside of the split. Let us elaborate on what we mean. Take a direct split γ_v and consider the set $S = \{u \in V(\gamma) \,|\, u \text{ L } v \text{ or } u \subset v\}$. Then, on the level of the intersection sequence, we know exactly what the deletion of γ_v from γ will do – the signed intersection sequence of $\gamma \backslash \gamma_v$ will simply be the signed intersection sequence of γ, with all vertices from S completely removed. This holds for any $u \subset v$ will not exist on the image $[\gamma \backslash \gamma_v]$, and any vertex u so that $u \text{ L } v$ will no longer be a vertex on $\gamma \backslash \gamma_v$, since we deleted one of its preimages under γ. A similar analysis can also be performed in the case of the wrap split. As the wrap split is the complement of the direct split on $[\gamma]$, we have $S' = \{u \in V(\gamma) \,|\, v \subset u \text{ or } v \text{ S } u \text{ or } u \text{ L } v\}$ as the set of vertices removed by deleting the wrap split γ_{v^*}.

Now, to show that the direct split carry a great deal of information about the curve, let us show how to determine the relations between intersection points completely in terms of the direct splits:

- $p_i \text{ L } p_j$ iff $p_i \in [\gamma_j]$ and $p_j \in [\gamma_i]$. In this case, p_i is not a vertex on p_j and p_i is not a vertex on γ_j.
- $p_i \text{ S } p_j$ iff $p_i \notin [\gamma_j]$ and $p_j \notin [\gamma_i]$.
- $p_i \subset p_j$ iff p_i is a vertex of γ_j.

Conversely, the relations also reflect information about the direct splits. In particular, the relations tell us exactly what the vertices of the direct and wrap splits will be. We note the following:

- $V(\gamma_v) = \{u \in V(\gamma) \,|\, v \neq u, u \subset v\}$
- $V(\gamma_{v^*}) = \{u \in V(\gamma) \,|\, v \neq u, v \subset u \text{ or } u \text{ S } v\}$

Thus, given any vertex $u \in V(\gamma)$, we have a natural partition of the vertex set $V(\gamma)$, employing the direct and wrap splits:

$$V(\gamma) = V(\gamma_v) \uplus V(\gamma_{v^*}) \uplus L(v) \uplus \{v\}$$

where $L(v) = \{u \in V(\gamma) \,|\, u \text{ L } v\}$ are all the vertices which occur as intersections of γ_v and γ_{v^*}. Here, we use $\uplus$ to indicate disjoint union.

Let us make another small, but key observation about the nesting of direct splits that will lead to many nice inductive proofs.

Observation 2.4. *If $\gamma_i \in \mathscr{C}$ is a direct split on γ and γ_j is a direct split on γ_i, then γ_j is a direct split on γ.*

Given a curve γ so that $\{\gamma_2, \gamma_1\}$ is a SO decomposition and γ_1 is a direct split, we write $\gamma = \gamma_2 \vee \gamma_1$. For example, a simple figure eight would naturally decompose in this fashion.

2.5 Decompositions

Let us now introduce a general class of decompositions of plane curves, which are of particular interest to us. Given any curve $\gamma \in \mathscr{C}$, we have exactly $2|\gamma|$ free subcurves on γ, namely the direct and wrap splits of each vertex. Suppose we iteratively remove free subcurves until the only piece of γ left is a simple curve. Then the set of subcurves $\Omega = \{\gamma_i\}_{i=1}^k$ is called a *free subcurve decomposition.* More precisely, we call the set $\Omega = \{\gamma_i\}_{i=1}^k$ a free subcurve decomposition of γ when

- γ_i is free on $\gamma \setminus \cup_{j=1}^{i-1} \gamma_j$.
- $[\gamma_i] \cap [\gamma_j] \subset V(\gamma)$
- $[\gamma] = \cup_{i=1}^k [\gamma_i]$.

In the case that each γ_i is either positive or negative SO, we call Ω a *self-overlapping decomposition* (SO decomposition). Here, we call a curve negative SO if its reversal is SO in the regular sense. The next section will clarify this definition. Additionally, if each γ_i is simple, then we call Ω a *simple subcurve decomposition* (SSD). Given a free subcurve decomposition $\Omega = \{\gamma_i\}_{i=1}^k$, we may write $V(\Omega)$ for the set of basepoints of each $\gamma_i \in \Omega$.

A vital fact about free subcurve decompositions is proved in [13].

Lemma 2.5. *Let $\gamma \in \mathscr{C}$ and take a subset of vertices $S \subset V(\gamma) \setminus \{p_0\}$ so that no two elements $u \neq v \in S$ are linked. Then there is a unique free subcurve decomposition Ω_S with $V(\Omega_S) = S \cup \{p_0\}$.*

Proof. We sketch the proof here. Let $P(S)$ be the (sub) poset of $P(\gamma)$ on the vertices in S. Then take v as a minimal element in $P(S)$. The core of the proof lies with the following claim: if a free subcurve decomposition Ω_S exists, then we must have the direct split $\gamma_v \in \Omega_S$. Repeatedly applying this claim and peeling off direct splits, we will end up with our unique free subcurve decomposition Ω. The basepoint p_0 of γ is necessarily the basepoint of the final subcurve, since p_0 is trivially maximal with respect to $\subset$. Additionally, we will never accidentally delete any vertices while removing direct splits along the way, since no two elements in S are linked. $\square$

Note that in the previous proof, we actually showed the following:

Corollary 2.6. *Let Ω be a free subcurve decomposition of γ. Then there is at least one direct split $\gamma_v \in \Omega$.*

Beyond SO decompositions and simple subcurve decompositions, there is an additional special class of decompositions we will repeatedly use.

Lemma 2.7 (Direct Split Decomposition)**.** *Let $\gamma \in \mathscr{C}$. Then we have a free subcurve decomposition $\Omega = \{\gamma_i\}_{i=1}^j \cup \{C\}$ such that*

1. *Each γ_i is a direct split.*
2. *C is a simple subcurve containing the basepoint $\gamma(0)$.*

Proof. This proof follows very naturally by induction. We will simply remove direct splits of maximal vertices until no vertices remain. Let us take a first maximal element p_1 in $P(\gamma)$. Then add the direct split γ_{p_1} to Ω. Now, all the vertices removed by this action are $S_1 = L(p_1) \cup S(p_1)$, those vertices linked with or contained in p_1. We now describe the general inductive step. Suppose we have built up our decomposition to $\Omega = \{\gamma_{p_i}\}_{i=1}^k$ so that each p_i is maximal in $P\left(\gamma \backslash (\cup_{l=1}^{i-1} \gamma_{p_l})\right)$ and p_i S p_l for $i \neq l$. Suppose that $\gamma_{cur} = \gamma \backslash \left(\cup_{i=1}^k \gamma_{p_i}\right)$ is not simple, or else we can set the current curve to be our final simple curve C in the decomposition. Let u be a vertex left on γ_{cur}. Then we claim that u is separate to each p_i. Take any vertex p_i. Then the only other possibility is $p_i \subset u$, since we have already precluded $u \subset p_i$ and u L p_i. As $p_i \subset u$ contradicts the maximality of p_i on poset $P\left(\gamma \backslash (\cup_{l=1}^{i-1} \gamma_{p_l})\right)$, we must have u S p_i. We then set p_{k+1} as a maximal element in $P\left(\gamma \backslash (\cup_{l=1}^{k} \gamma_{p_l})\right)$ and add $\gamma_{p_{k+1}}$ to Ω. In each stage of this process, we reduce the size of the current curve, so eventually we must reach a time so that $|\gamma_{cur}| = 0$. Once this occurs, we set $C = \gamma_{cur}$. Since $\gamma(0)$ is trivially maximal with respect to $\subset$, we note that it must lie on γ_{cur}. □

See Figure 4 for an example of a direct split decomposition.

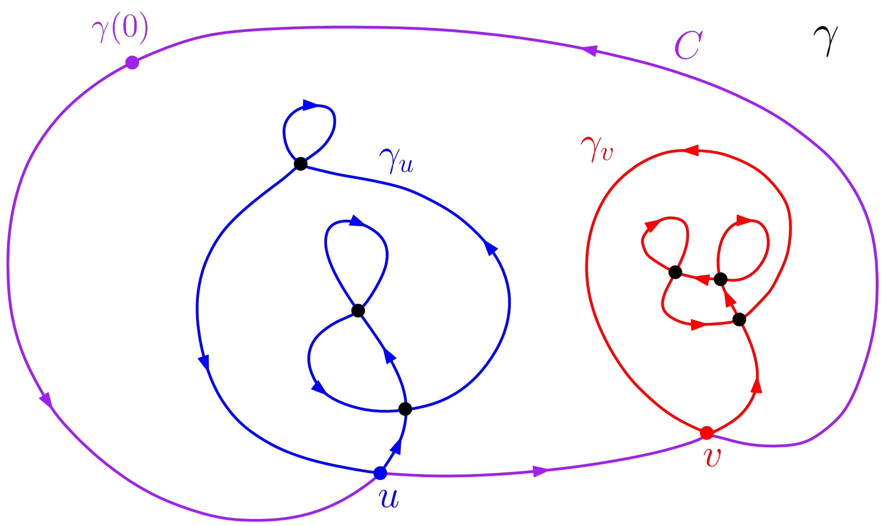


Figure 4: A curve γ and a direct split decomposition $\Omega = \{\gamma_u, \gamma_v, C\}$.

2.6 Minimum Homotopies

We are now in a position to define the minimum homotopy area of a curve, as in [10].

Definition 2.8 (Homotopy Area)**.** *Let $\gamma \in \mathscr{C}$ and H be a nullhomopty of γ. Define $E_H : \mathbb{R}^2 \to \mathbb{Z}$ by setting $E_H(x)$ as the number of connected components of $H^{-1}(x)$. Intuitively, this counts the number of times that H sweeps over x. The homotopy area* $\textsc{Area}(H)$ *of H is then defined as*

$$\textsc{Area}(H) = \int_{\mathbb{R}^2} E_H(x)\ dx$$

Finally, we define minimum homotopy area $\sigma(\gamma)$ as:

$$\sigma(\gamma) = \inf_H \{\textsc{Area}(H) : \text{ } H \text{ is a nullhomotopy of } \gamma\}$$

A basic, but essential, property of the minimum homotopy area is that it is bounded below by the winding area, as shown in [6, 10]:

Lemma 2.9. *Let $\gamma \in \mathscr{C}$. Then $\sigma(\gamma) \geq W(\gamma)$.*

Now, for the most important result yet, the link between minimum area homotopies and self-overlapping curves. The main theorem from [10] says that any curve has a minimum homotopy that consists of subhomtopies contracting free SO subcurves of γ. See Definition 2.3.

Theorem 2.10 (Minimum Homotopy Decompositions)**.** *Let $\gamma \in \mathscr{C}$. Then γ has a minimum homotopy H such that $H = \sum_{i=1}^{k} H_i$ and each H_i is a nullhomotopy for a free SO subcurve on $\widetilde{\gamma}_i$, where $\widetilde{\gamma}_i \xrightarrow{H_i} \widetilde{\gamma}_{i+1}$.*

In other words, we can view minimum homotopies H as being associated to SO decompositions. Theorem 2.10 tells us that in essence, minimum homotopies are just sequences of nullhomotopies of free self-overlapping subcurves. Thus, computing the minimum homotopy area of a curve γ can be reduced to finding optimal SO decompositions. See Figure 5 for an example SO decomposition.

If $\Omega = \{\gamma_i\}_{i=1}^{k}$ is a SO decomposition of a curve, we may refer to the vertices $V(\Omega)$ which are the basepoints of $\{\gamma_i\}_{i=1}^{k}$ as the *anchors* or *anchor points* of the homotopy H. It is worth commenting that there is a unique canonical order to contract the SO curves in a SO decomposition Ω. If $\{v_i\}_{i=1}^{k}$ are the anchors of the curves $\{\gamma_i\}_{i=1}^{k}$, then a canonical homotopy is a concatenation $H = \sum_{i=1}^{k} H_i$ where each subhomotopy H_k is a nullhomotopy of the subcurve whose anchor point v_j has the k^{th} largest index in the intersection sequence. In essence, we contract the SO curves from 'back' to 'front', since the back one is always free.

Often times we will be interested in finding a way to split the curve into smaller pieces, while respecting minimum homotopy area. When we are successful, we use the following terminology.

Definition 2.11. *Let γ_i be a direct split on a curve γ. If homotopy area is preserved by splitting the curve into the direct and wrap splits at the vertex p_i:*

$$\sigma(\gamma) = \sigma(\gamma_{i^*}) + \sigma(\gamma_i)$$

then we call γ_i ***ideal****.*

2.7 Interior Boundaries

Interior Boundaries are natural generalizations of SO curves. We prefer to think of an interior boundary as a generalization of SO curves from the perspective of minimum homotopy area. With this in mind, we define:

Definition 2.12. *We call a curve γ a* **+k*-interior boundary***, *or* **k*-boundary*** *for short, when $\sigma(\gamma) = W(\gamma)$, $whit(\gamma) = +k$ and γ is positive consistent, $wn(x, \gamma) \geq 0$ for each $x \in \mathbb{R}^2$. We call γ a $-k$-boundary when its reversal $\overline{\gamma}$ is a $+k$-boundary.*

While Titus' original formulation of interior boundaries was much different, the two coincide [13]. We will provide a sketch of this fact shortly.

Titus' original formulation of interior boundaries was as follows.

Definition 2.13. *We call a curve $\zeta : [0, 1] \to \mathbb{R}^2$ a* ***Titus interior boundary*** *iff there exists a map $F : D^2 \to \mathbb{R}^2$ such that F is continuous, light, pre-images are totally disconnected, open, sense-preserving, and $F|_{\partial D^2} = \zeta$. The map F is called* ***properly interior****.*

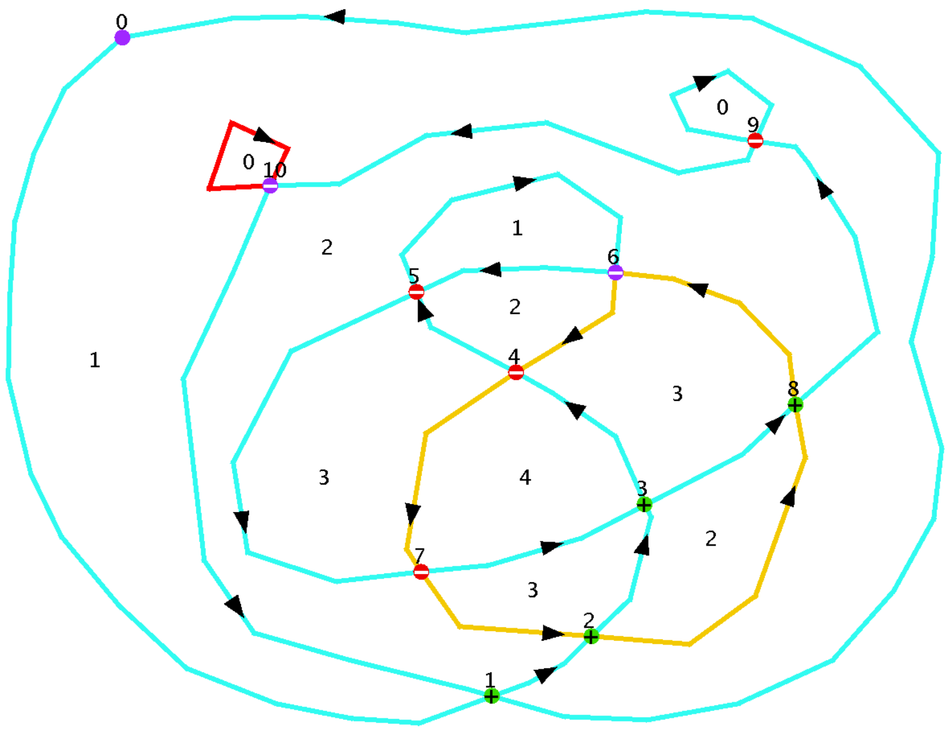


Figure 5: Figure from the program written by the author. A SO decomposition of a plane curve from a minimum homotopy H. The different SO subcurves are shown in different colors, with the anchor points shown in purple.

We begin with a lemma of Titus.

Lemma 2.14. *A linear retraction of the disk D^2 induces a minimum homotopy for a Titus interior boundary with $A(H) = W(\zeta)$.*

Proof. Let ζ be a Titus interior boundary. As mentioned in [21], the number of pre-images $\left|\zeta^{-1}(p)\right|$ of any point $p \in \mathbb{R}^2$ is equal winding number $wn(\zeta, p)$. This implies that the linear retraction of the disk to a point induces a minimum homotopy H achieving $W(\zeta)$. □

Now, our definition of interior boundary is actually slightly different from that of [13]. Nevertheless, all three definitions of interior boundaries are equivalent, as we now show. The condition (iii) was the definition of an interior boundary in [13].

Theorem 2.15 (Equivalence of Interior Boundaries)**.** *Let γ have a positive outer basepoint $\gamma(0)$ and $whit(\gamma) = +k$. Then the following are equivalent:*

(i) γ is an interior boundary.

(ii) γ is a Titus interior boundary.

(iii) γ has a SO decomposition $\Omega = \{\gamma_i\}_{i=1}^k$ where each γ_i is positive SO.

Proof. We first show (i) $\Leftrightarrow$ (ii). It follows by induction from Titus' work that if we can decompose a curve γ into curves $\{\gamma_i\}_{i=1}^k$ so that $\gamma_i \cap \gamma_j$ is either empty or is an intersection point, and $whit(\gamma_i) = 1$ for all i, that for any γ_i, γ_j which meet at a point, we can find an arc σ_{ij}, interior to both γ_i, γ_j. Doing this inductively, dividing the disk D^2 into k pieces, we can glue the interior mappings F_i together into a properly interior mapping F that extends each F_i and hence represents the curve γ. We conclude our interior boundaries are Titus interior boundaries as well.

Conversely, Lemma 2.11 shows that a Titus interior boundary ζ is ideal, $\sigma(\zeta) = W(\zeta)$. Lemma 1 in [21] states that Titus interior boundaries are consistent. Thus, ζ is an interior boundary.

We now show (i) $\Rightarrow$ (iii). Let γ be an interior boundary. By Theorem 2.10, we have a SO decomposition $\Omega = \{\gamma_i\}_{i=1}^j$ of an interior boundary. Suppose that some γ_k were negative SO. Then we must have $\text{int}(\gamma_k) \cap (\cup_{i \in [j], i \neq k} \text{int}(\gamma_i)) \neq \varnothing$, since γ is positive consistent. In [13] it is proven that for a minimum homotopy H, each face F has $E_H(x)$ constant for each $x \in F$. Thus, we must have some face F from the planar multigraph $G(\gamma)$ so that F lies in the interior of γ_k and in the interior of γ_i for some positive SO curve γ_i. We now have our contradiction, for it follows that if H_Ω is the canonical optimal homotopy with decomposition Ω, then F is swept more than $W(F)$ times, since the winding number will increase on F as we contract γ_k. Thus, no negative SO subcurve γ_k may exist in Ω. Since $whit(C) = +1$ for any positive SO subcurve and $whit(\gamma) = \sum_{i=1}^k whit(\gamma_i)$ for any SO decomposition $\{\gamma_i\}_{i=1}^k$, we must have $k = j$ and the claim is proven.

The converse (iii) $\Rightarrow$ (i) is easier. Suppose γ has a decomposition into k positive SO subcurves Ω. Then the canonical homotopy H_Ω associated to Ω will be left sense-preserving since each minimum nullhomotopy of the subcurves is left sense-preserving. By Lemma 3.5, Sense-Preserving Homotopies are Optimal, we have $\sigma(\gamma) = W(\gamma)$. For any SO decomposition $\Omega = \{\gamma_i\}_{i=1}^k$, by our characterization of winding numbers, it is not too hard to see that $wn(x,\gamma) = \sum_{i=1}^k wn(x,\gamma_i)$. In fact, this is true of any free subcurve decomposition. Thus, $wn(x,\gamma) = \sum_{i=1}^k wn(x,\gamma_i) \geq 0$ since γ_i is positive consistent, as a positive SO curve. $\square$

2.8 Wraps

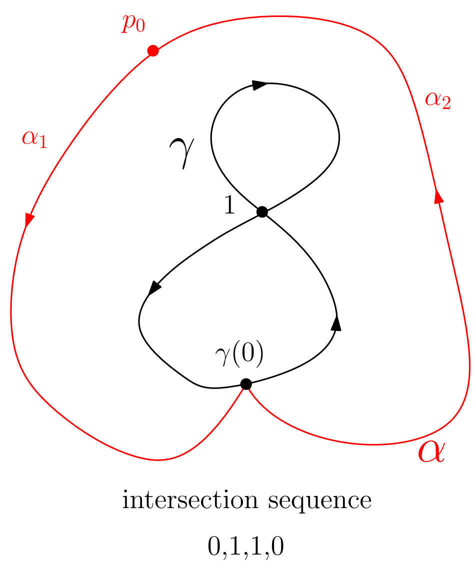
We now define a new operation called a *wrap*, which surrounds a given curve γ with a Jordan curve. As we will see in Theorem 8.8, this curious operation behaves very nicely with respect to minimum homotopy area.

Definition 2.16. *Given a curve γ with positive outer point $\gamma(0)$, let us choose a positive Jordan curve α with basepoint $\gamma(0)$ so that $\gamma \backslash \{\gamma(0)\} \in int(\alpha)$. Then choose a point $p_0 \in \alpha \backslash \{\gamma(0)\}$. Write α_1, α_2 as the subpaths along α from p_0 to $\gamma(0)$ and from $\gamma(0)$ to p_0. Let γ' be the curve $\alpha_1 * \gamma * \alpha_2$, smoothed at $\gamma(0)$ so that is normal. Then we call γ' the **positive wrap** of γ, written as $Wr_+(\gamma)$. The negative wrap is defined completely analogously if there exists a negative outer point, except we choose a negative Jordan curve.*

We may wish to wrap positively or negatively around a curve γ whose basepoint is of the opposite sign of our wrap. In this case, we need a different definition. We first begin by performing a I_b move to add a simple loop $\tilde{\gamma}$ of the opposite orientation tangent to the basepoint $\gamma(0)$. Let γ' be the curve after the I_b move. We can then shift our basepoint to a point P' to lie on $\tilde{\gamma}$ and then wrap as we did before on γ'. See Figure 6 for an example of these procedures.

Remark: The wrap does not define a new plane curve, but rather defines a curve up to signed intersection sequence. Algebraically, the wrap can be defined by its well-defined effect on the intersection sequence. If a curve γ has intersection sequence $\mathcal{I} = 0\, i_1 \cdots i_n\, 0$, with positive outer point

γ, then the new intersection sequence of $Wr_+(\gamma)$ will be $0\,p\,i_1 \cdots i_n\,p\,0$, with p positive the first time and negative the second time. If instead we consider $Wr_-(\gamma)$, it will have intersection sequence $0\,p\,q\,i_1 \cdots i_n\,q\,p\,0$ where both p and q are negative vertices. See Figures 6, 7.



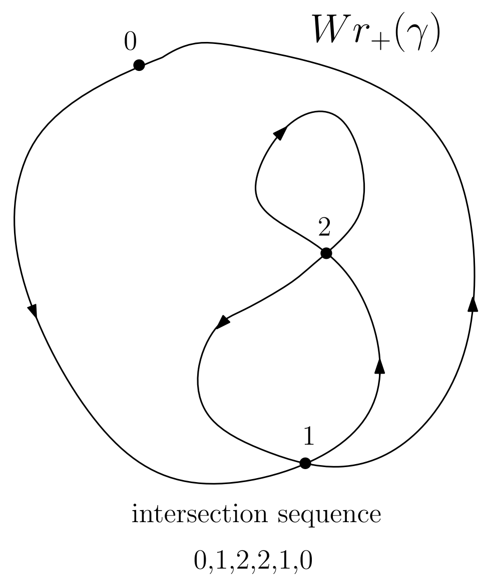


Figure 6: A curve γ and its positive wrap $Wr_+(\gamma)$. The (unsigned) intersection sequences of both curves are shown. Note that as $\gamma(0)$ becomes the first vertex on the wrap, we must smooth to make the curve regular and normal again.

3 Properties of SO Curves

We now provide some machinery to understand SO curves. Now, as SO curves are images of the boundary of immersed disks, they have a natural interior. In fact, if one walks along the curve, the interior is always to the left. It makes sense then that stretching individual regions of the curve would not rip the disk and hence should preserve SO-ness. Titus actually proved more than this by showing the claim holds for interior boundaries. [22].

Lemma 3.1 (SO-ness Invariant by Signed Intersection Sequence)**.** *Let $\gamma, \gamma' \in \mathscr{C}$ have positive outer basepoints and the same signed intersection sequences. Then γ is an interior boundary iff γ' is.*

The following lemma is stated in [10].

Lemma 3.2. *Let γ be a SO curve. Then $\sigma(\gamma) = W(\gamma)$. Moreover, the linear retraction of the disk D^2 induces a minimum homotopy.*

The following definition allows for quite a useful alternate formulation of self-overlapping-ness.

Definition 3.3. *Let $\gamma \in \mathscr{C}$ be SO. Suppose we find a simple path $P \subset \mathbb{R}^2$ so that $p = P(0)$ and $q = P(1)$ lie on $[\gamma]$ but are not vertices of γ. Here, we allow that P intersects γ, as long as P does not cross itself. Let us consider the path $\widetilde{P}$ from p to q along the orientation of γ. More precisely, if $t_p = \gamma^{-1}(p)$ and $t_q = \gamma^{-1}(q)$, we have $\widetilde{P} = \gamma|_{[t_p, t_q]}$. Suppose further that the following conditions are met:*

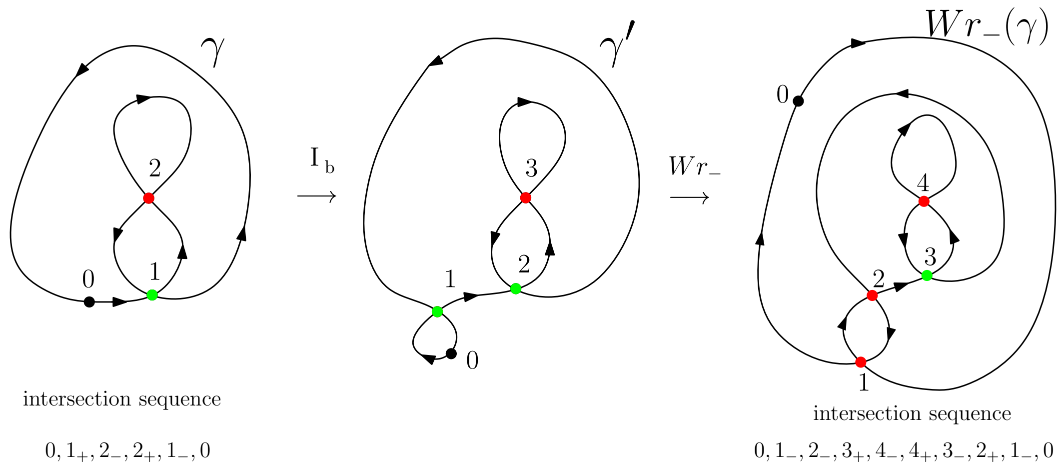


Figure 7: A curve γ with positive outer point $\gamma(0)$ along with its transformation into $Wr_-(\gamma)$. First, we do a I_b move, then we wrap after forcing a negative outer basepoint. The signed intersection sequences of both $\gamma, Wr_-(\gamma)$ are shown. Positive vertices are in green, negative vertices in red, and basepoints in black.

1. *P and $\widetilde{P}$ are simple paths (viewed separately from γ.) Additionally, $P \cap \widetilde{P} = \{p, q\}$. Equivalently, the path $\widetilde{P} * \overline{P}$ is a simple closed curve.*

2. *The interior of the region R bounded by $\widetilde{P} * \overline{P}$ lies to the left, i.e., the simple closed curve is positive, counterclockwise oriented.*

We can then perform a ***Blank cut*** *on γ along P. To do this, we view P as a path to cut along, like a perforated edge, and we replace the path $\widetilde{P}$ on γ with P, then smooth at p and q to make the curve regular again. We can interpret this action as* cutting *out the region R from the interior of γ. If we perform a sequence of Blank cuts until the curve becomes simple, we obtain a sequence of curves produced after each cut, $\{\gamma_i\}_{i=0}^k$, where $\gamma_0 = \gamma$ and γ_i is the curve after the i^{th} blank cut. We call such a decomposition of γ a* ***Blank cut decomposition****. We may refer to the decomposition as the sequence $\{\gamma_i\}_{i=0}^k$ of subcurves, or as the sequence $\{P_i\}_{i=1}^k$ of paths we cut along. See Figure 8.*

Much of the beauty of SO curves lies in the variety of equivalent definitions.

Theorem 3.4 (Equivalent Definitions of SO Curves, [4, 10, 14, 18])**.** *Let $\gamma \in \mathscr{C}$ have positive outer basepoint. Then the following are equivalent:*

1. *(Analysis) There is an immersion $F : D^2 \to \mathbb{R}^2$ so that $F|_{\partial D^2} = \gamma$.*

2. *(Geometry) There exists a Blank cut decomposition of γ.*

3. *(Algebra) $whit(\gamma) = 1$ and $\sigma(\gamma) = W(\gamma)$ and γ is positive consistent.*

4. *(Topology) For any minimum homotopy H of γ, H is left sense-preserving and has exactly one I_b move.*

If the reversal $\overline{\gamma}$ of a curve is SO, then we call γ *negative self-overlapping.* From the equivalent definitions of interior boundaries in the last section, we know that our SO curves are equivalently +1 boundaries and our negative SO curves are −1 boundaries.

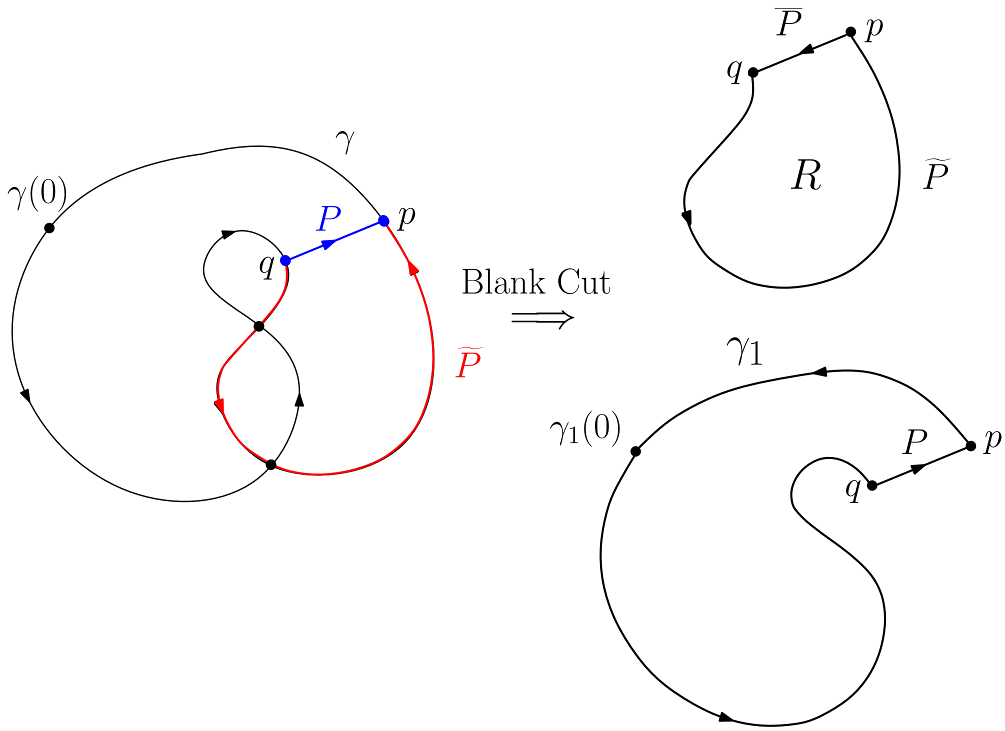


Figure 8: A blank cut on a small SO curve, which simplifies the curve immediately. The simple curve γ_1 produced by the cut is shown as well as the region R removed from the interior by the cut.

Now, if we have a curve γ that is piecewise regular, we can smooth γ at its irregularities to obtain a curve $\tilde{\gamma}$ that is arbitrarily close to γ. In this case, we say γ is SO iff its smoothed version $\tilde{\gamma}$ is SO. In particular, this action will be performed when we consider direct splits and wrap splits, which are necessarily irregular at their basepoints.

As mentioned earlier, self-overlapping curves γ have a natural interior. If F is an immersion of γ, then $F(\text{int}(D^2))$ is the image of the interior of the disk, and can be regarded as the interior of the curve. The interior of the curve always lies locally to the left of the curve, following its orientation. Intuitively, this partially explains why a cut decomposition should exist for a SO curve: we are simply cutting our immersed disk back into simple pieces. Moreover, through the cut decomposition characterization of self-overlapping-ness, one can talk about whether a planar image $[\gamma]$, with orientation is self-overlapping or not – this simply amounts to the existence of a cut decomposition.

In fact, with our new and improved understanding of SO curves, we can now see a few more equivalent definitions of interior boundaries. In particular, note the similarity between (1-4) here and (1-4) in Theorem 3.4.

Theorem 3.5 (Equivalent Definitions of Interior Boundaries (Revisited), [4, 10, 14, 18])**.** *Let $\gamma \in \mathscr{C}$ have positive outer basepoint and whit$(\gamma) = +k$. Then the following are equivalent:*

1. *(Analysis) There is a properly interior mapping $F : D^2 \to \mathbb{R}^2$ so that $F|_{\partial D^2} = \gamma$.*

2. *(Geometry) There exists a sequence of Blank cuts and positive I_b moves simplifying γ.*

3. *(Algebra) $\sigma(\gamma) = W(\gamma)$ and γ is positive consistent.*

4. *(Topology) For any minimum homotopy H of γ, H is left sense-preserving and has exactly k I_b moves.*

5. *(Geometry/Topology) There is a SO decomposition $\Omega = \{\gamma_i\}_{i=1}^{k}$ of γ such that each γ_i is positive SO.*

The only reason (5) could not be stated for SO curves is that it is trivially true for any SO curve γ: $\Omega = \{\gamma\}$ is a SO decomposition, but this is no help to us.

We now state a few useful lemmas.

Lemma 3.6 (Sense-Preserving Homotopies Are Optimal, [10, 13])**.** *Suppose H is a sense-preserving nullhomotopy for a curve γ. Then, H is minimum: $\sigma(C) = Area(H) = W(C)$.*

Lemma 3.7 (Right Sense-Preserving Homotopies Preserve SO-ness)**.** *Let H be a regular right sense-preserving homotopy so that $\gamma \xrightarrow{H} \gamma'$. If γ is SO, then γ' is SO.*

Proof. Since γ is SO, it has a left sense-preserving nullhomotopy H'' by Theorem 3.4. Let us reverse our given homotopy H to obtain H' by $H'(s,t) = H(1-s,t)$. Then we note that the concatenation $H' + H''$ is a left sense-preserving nullhomotopy for γ. Since sense-preserving homotopies are optimal, $\sigma(\gamma') = W(\gamma')$. Also, as H is regular, $W(\gamma') = W(\gamma) = 1$. Thus, by Theorem 3.4, γ' is SO. □

We think of SO curves as being smooth in the sense that they have a nice interior; the immersion forming the curve stretches, but does not rip, the disk. For the following discussion, let H be a regular homotopy between curves γ, γ' with unit Whitney index. We think of H as preserving the smoothness of the immersion F of γ if it is right sense-preserving. With this intuition, the previous lemma makes perfect sense: if we homotope a SO curve with a right sense-preserving homotopy, then our target curve is *at least as smooth* as the source curve, so our target should be SO as well.

In fact, with Lemma 3.5 in hand, specifically condition (4), almost the exact same proof tells us immediately that Lemma 3.7 generalizes to the following:

Lemma 3.8. *Suppose H is a regular right sense-preserving homotopy so that $\gamma \xrightarrow{H} \gamma'$. If γ is $+k$-boundary, then γ' is a $+k$-boundary as well.*

We conclude this section with a nice link between SO curves and two-boundaries.

Lemma 3.9 (2-boundaries from SO Curves)**.** *Let γ be a SO curve. Then for any negative loop γ_- on γ, $\gamma\backslash\gamma_-$ is a 2-boundary.*

Proof. Take γ and a negative loop γ_- and shrink γ_- with a regular right sense-preserving homotopy H so that $\gamma \xrightarrow{H} \gamma'$ and the loop corresponding to γ_- on γ' is simple. Then γ' is also SO by Lemma 3.7. We note that $\gamma\backslash\gamma_- = \gamma'\backslash\gamma'_-$ where γ'_- corresponds to γ_-. Thus, the claim is true for γ iff it is true for γ'. Now, from the proof of Theorem 5 in [22] we can immediately see that $\gamma'\backslash\gamma'_-$ is a Titus interior boundary, since γ_- is empty. As $whit(\gamma'\backslash\gamma'_-) = 2$, by Theorem 3.5, γ' is a 2-boundary. □

4 Simple Subcurve Decompositions

Here, we present a new combinatorial idea to simplify a planar curve to its bare bones structure, or as we will call it, an *elementary form*, which retains much of the structure of the original curve. The simple subcurve decompositions introduced in this section bare a strong resemblance to the so-called *Seifert decompositions* of a plane curve, as we show later in this section.

4.1 Definition and Existence

We define a *loop* as a direct split γ_v that is simple, when viewed separately from the rest of the curve. Thus, intersection points of γ may lie on γ_v, but none occur as intersections of γ_v with itself. Now, every non-simple plane curve has a loop. In fact, the direct split γ_w, where w is the last intersection point on γ, is necessarily a loop. Thus, we can iteratively remove loops until γ becomes simple. This procedure yields a decomposition of γ into simple subcurves. Note that these subcurves are simple only in the sense that they do not intersect themselves; they may intersect each other on γ. We will call such a decomposition a *simple subcurve decomposition* (SSD).

To clarify the above definition, some additional definitions are in order.

Definition 4.1. *(Outwards and Inwards Loops) Let $\gamma \in \mathscr{C}$ and let $\tilde{\gamma}$ be a loop on γ based at a vertex v. Then since $\tilde{\gamma}$ is a simple, closed curve, it has a natural interior int$(\tilde{\gamma})$. Let us consider the two edges e_1, e_4 connected to v on $\gamma\backslash\tilde{\gamma}$. If e_1, e_4 lie inside int$(\tilde{\gamma})$, then we call $\tilde{\gamma}$ an* ***inwards loop****. Otherwise, both edges must lie outside int$(\tilde{\gamma})$ and hence we call $\tilde{\gamma}$ an* ***outwards loop****.*

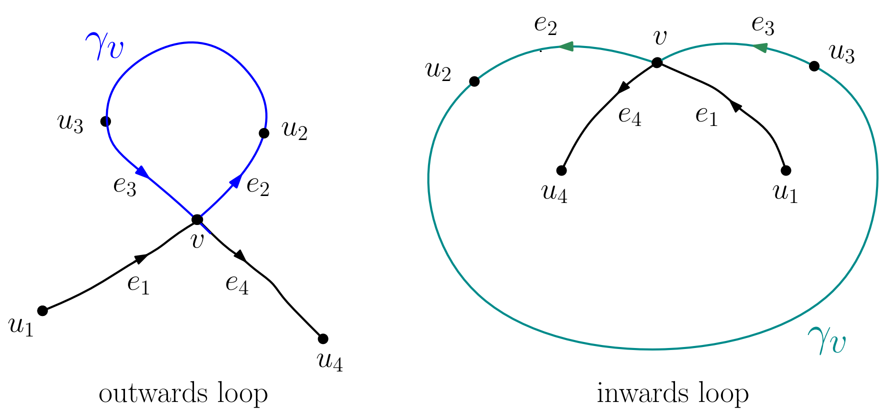


Figure 9: Outwards and inwards loops on a curve γ. The edges adjacent to the basepoint are labeled in order as they appear on γ. Note that the edges e_1 and e_4 lie outside the interior of the loop in the case of an outwards loop and inside the interior in the case of an inwards loop.

To construct a *standard* SSD, we require that all loops are outwards at the time of their addition to the decomposition. More precisely, if $\Omega = \{\gamma^i\}_{i=1}^j$ are the simple subcurves in our decomposition, so that γ^i was the i^{th} simple subcurve added, then for Ω to be a standard SSD, we require that γ^j is an outwards loop on the curve $\gamma\backslash\left(\cup_{i=1}^{j-1}\gamma^i\right)$. If we allow the addition of inwards loops, we call such a SSD a *generalized simple subcurve decomposition* (generalized SSD). If we have a curve γ with free subcurve $\tilde{\gamma}$ and a SSD Ω of γ so that Ω has a subset $\Psi \subset \Omega$ so that Ψ is a SSD of $\tilde{\gamma}$, then we may refer to Ψ as the *sub-SSD* of Ω with respect to $\tilde{\gamma}$.

We now observe that one can always form a standard SSD for a plane curve $\gamma \in \mathscr{C}$.

Lemma 4.2 (Existence of an Outwards Loop)**.** *Let $\gamma \in \mathscr{C}$ have an outer basepoint. Then if γ is non-simple, it has an outwards loop.*

Proof. We show that the first self-intersection of the curve is the basepoint of an outwards loop. Let us set $t_0 = \sup\{t \in [0,1] \,|\gamma_{[0,t]}$ is injective$\}$. That is, we trace along γ from the basepoint, following the orientation, until the curve intersects itself. It is apparent that $v = \gamma(t_0)$ is a vertex of γ. We call v the first self-intersection of γ. Now, we show that the direct split γ_v is a loop. Note that we cannot have any $u \subset v$, for this would imply that we have a subsequence $vuuv$ in the intersection sequence, and hence γ intersects itself at u before v. Equivalently, we could define v as the first vertex to appear twice in the intersection sequence. But since the vertices $u \in V(\gamma)$ so that $u \subset v$ are the vertices of the direct split γ_v, it follows that γ_v is simple, once detached from γ. By definition, then γ_v is a loop. It is not hard to see that γ_v cannot be inwards by construction of v. Since $\gamma(0)$ lies outside of $\text{int}(\gamma_v)$, as an outer basepoint, we note that if γ_v were inwards, the path $P = \gamma_{[0,t_0]}$ would cross $[\gamma_v]$ to get from outside the simple curve to inside it. This is then a contradiction, for if the crossing occurred at a point q on $[\gamma_v]$, then q would be the first self-intersection of γ, as we would reach q a second time before we reach v a second time. Thus, γ_v is outwards and we are done. □

Let us note that a nearly identical proof to the one given in the lemma proves that the last self-intersection point of the curve must be the basepoint of an outwards loop as well, if we have an outer basepoint.

Beyond just being a nice decomposition of our curve, each standard SSD Ω enables us to define a new curve γ_Ω, which we call an *elementary form* of the original curve γ. The key property of γ_Ω is that we have a natural way to embed the vertices of γ_Ω into $V(\gamma)$. In fact, we have an injection $\phi : V(\gamma_\Omega) \hookrightarrow V(\gamma)$ that preserves relations. Before getting to elementary forms, we need to introduce an important concept.

Definition 4.3. *Let $\gamma \in \mathscr{C}$. We call γ* ***elementary*** *iff no vertices are linked.*

Let us now show the existence of an elementary form γ_Ω associated to standard SSD of a curve γ. Note that our elementary forms are indeed elementary, since any pair of vertices u, v in a SSD are necessarily not linked.

Proposition 4.4 (Existence of an Elementary Form)**.** *Let Ω be a standard SSD of a curve γ. Then according to $\subset$, the lattice T formed from basepoints $V(\Omega)$ of Ω is a tree. Moreover, there is a curve γ_Ω so that the intersection points $V(\gamma_\Omega)$ have the same combinatorial relations as $V(\Omega)$.*

Proof. This proof is very straightforward, but since the result is extremely important, we present it anyway. The tree nature of T follows by induction rather easily on the size of Ω. With $|\Omega| = 1$, this is trivial. Since Ω induces a nullhomotopy H_Ω of γ, we must have some loop γ_i in Ω as free on γ. But then we look at $\Omega \backslash \{\gamma_i\}$, which is a tree T'. Adding γ_i back as a leaf to T', by connecting it to the smallest element γ_j which has $\gamma_i \subset \gamma_j$, shows that T is also a tree. We can label each node in the tree with the index of the basepoint p_i of the curve γ_i associated to the node. Then, let us do a post-order traversal of T. That is, for each node, we visit the node, then visit its children in order of their index, then return to the node. Let us begin our post-order traversal at a maximal element of T. This will ensure us an outer basepoint. Now, if we print each node's index both times we touch the node, we will achieve the unsigned intersection sequence of the curve. We can then add an appropriate sign to each visit depending on the orientation of the simple subcurve γ_i with basepoint $v_i \in V(\Omega)$. A curve γ_Ω with this signed intersection sequence has vertices $V(\gamma_\Omega)$ with the same combinatorial relations as $V(\Omega)$, and is normal and regular by construction of how we pasted together the simple subcurves. ☐

We now take a moment to impart an important intuition upon the reader. One should view the elementary form as a purified form of the original curve, whittled down to its essential structure. Let us explain this point in detail. For any simple subcurve decomposition Ω of a curve γ, we can partition our vertex set $V(\gamma)$ into

$$V(\gamma) = V(\Omega) \cup L(\Omega)$$

where $L(\Omega) := \{v \in V(\gamma) \mid v \text{ L } u \text{ for some } u \in V(\Omega)\}$, the set of points on $V(\gamma)$ that link a basepoint in $V(\Omega)$. This fact holds because each subcurve $\tilde{\gamma} \in \Omega$ is simple. Thus, for any vertex $v \in V(\Omega)$ we either have v L u for v S u for any $u \in V(\gamma)\backslash V(\Omega)$. In other words, every vertex of our curve γ is either the basepoint of a simple curve on Ω, or it is an intersection of two simple subcurve $\gamma^{\tilde{a}}, \gamma^{\tilde{b}} \in \Omega$. With this in mind, it is clear that the elementary form is achieved from the original curve γ by removing all such intersection points of the latter kind. In essence, γ_Ω is formed from γ by elementarify-ing γ, by trimming the fat of excess intersection points. We preclude inwards loops from our standard SSD's because this process is impossible to complete with an inwards loop $\tilde{\gamma}$ in our decomposition. $\tilde{\gamma}$ will always have at least two intersections with the simple subcurve that is tangent to it at it its basepoint $\tilde{b}$.

Note that the elementary form is only defined up to signed intersection sequence. As we will only be concerned with self-overlapping-ness of the elementary form, this information is sufficient by Lemma 3.1. See Figure 10 for an example SSD and associated elementary form.

4.2 Algebraically Computing Simple Subcurve Decompositions

We now formulate the explicitly the relationship between SSD's geometrically and algebraically:

Proposition 4.5 (Algebraic Formulation of SSD's)**.** *Let γ be a plane curve and $S \subset V(\gamma)$ a subset of vertices. Then we have a unique (generalized) SSD Ω with $V(\Omega) = S$ iff*

1. *No two vertices from S are linked.*
2. *For each $v \in V(\gamma)\backslash S$, $\exists u \in S$ so that v L u.*
3. *$\gamma(0) \in V(\Omega)$.*

Proof. The direction ⇒ was just shown.

To see the converse, we take such a set of vertices S and we construct a SSD Ω with $V(\Omega) = S$. Uniqueness will follow easily. The core of the proof lies in the following argument: if we take a minimal vertex $v \in S$ with respect to $\subset$, then the direct split γ_v must be a loop. This same argument was presented in Lemma 4.2. Essentially, the same statement holds if we keep deleting minimal vertices. In the general inductive step, it is not too hard to see that if we take a direct split γ_v so that any vertex u with $u \subset v$ has already been added to $V(\Omega)$, then the rest of the direct split is just a simple subcurve. Again, this holds because vertices on γ_v correspond bijectively with vertices $u \in V(\gamma)$ with $u \subset v$. Hence, we have a unique subcurve $\gamma^v = \gamma_v \backslash (\cup_{u \subset v} \gamma^u)$ to add to Ω, with basepoint v, where γ^u is the previously defined simple subcurve in Ω with basepoint u. As $\gamma(0)$ is necessarily maximal in the poset $(V(\gamma), \subset)$, it necessarily appears in every SSD vertex set. $\square$

Remark: The above conditions can be reformulated in terms of the linkage graph $G_L(\gamma)$ of the curve: Let $S \subset V(\gamma)$. Then we have a unique (generalized) SSD Ω with $V(\Omega) = S$ iff

1. S is an independent set in $G_L(\gamma)$.
2. The union of the neighborhoods, $N(v) = \{u \in V(\gamma) \,|\, uv \in E(G_L(\gamma))\}$, of each vertex, inside $G_L(\gamma)$, is the whole vertex set:

$$V(\gamma) \backslash \{0\} = \cup_{v \in S} N(v) \cup S$$

   In other words, S is a *maximally* independent set in $G_L(\gamma)$.

Thus, we have now established the following equivalent conditions for a curve γ being elementary:

1. γ has exactly one simple subcurve decomposition.
2. There is a standard SSD Ω of γ so that γ and γ_Ω are combinatorially equivalent.
3. No two vertices in $V(\gamma)$ are linked.

From here on out we will notate $\mathscr{D}(\gamma)$ as the set of (generalized) simple subcurve decompositions of γ. By Proposition 4.5, we have a natural map $\sigma : \mathscr{D}(\gamma) \to \mathcal{P}(V(\gamma))$ by way of $\sigma(\Omega) = V(\Omega)$. Moreover, this map is an injection by Proposition 4.5. Due to this, we may refer to a simple subcurve decomposition Ω of a curve γ either by its set of simple subcurves $\{\gamma^v\}_{v \in V(\Omega)}$ or simply by its vertex set $V(\Omega)$. Unless otherwise specified, the reader is encouraged to assume γ^v is a simple subcurve from a SSD with basepoint v and γ_v as the direct split on γ with basepoint v.

With all this combinatorial information floating around, it is quite easy to leave the reader confused. We now adopt some standard notation to avoid this problem. When referring to an element $a \in V(\gamma_\Omega)$, there are exactly four different possible interpretations. We now provide standard notation for each of these objects, given a fixed element $a \in V(\gamma_\Omega)$.

1. $a \equiv$ the vertex from $V(\gamma_\Omega)$.
2. $\gamma_\Omega^a \equiv$ the unique simple subcurve with basepoint a the unique SSD of γ_Ω.
3. $\tilde{a} = i(a) \equiv$ the vertex from $V(\gamma)$ associated to a by the inclusion map $i : V(\gamma_\Omega) \hookrightarrow V(\gamma)$.
4. $\gamma^{\tilde{a}} \equiv$ the unique simple subcurve in Ω with basepoint $\tilde{a} \in V(\gamma)$.

Additionally, we may write $V(\Omega) \subset V(\gamma)$ for the set of vertices that are basepoints of subcurves $\tilde{\gamma} \in \Omega$. It follows that the inclusion map viewed as a map $i : V(\gamma_\Omega) \to V(\Omega)$ is a bijection.

A nice fact about SSD's is that they preserve maximal elements in the poset $P(\gamma)$.

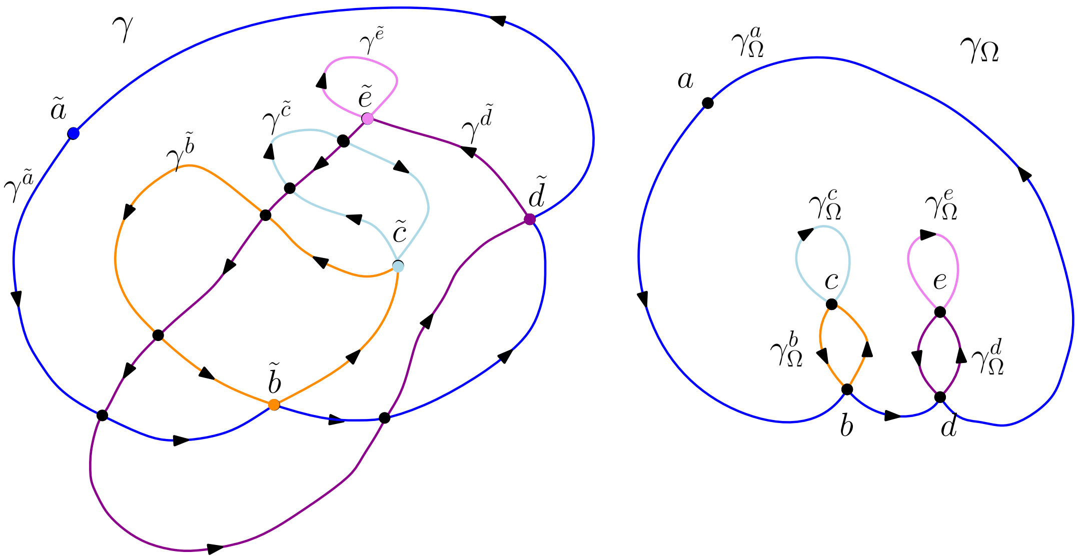


Figure 10: A curve γ with a (standard) simple subcurve decomposition $\Omega = \{\gamma^{\tilde{a}}, \gamma^{\tilde{b}}, \gamma^{\tilde{c}}, \gamma^{\tilde{d}}, \gamma^{\tilde{e}}\}$ and its elementary form γ_Ω.

Proposition 4.6 (Pullbacks of Maximal Points are Maximal)**.** *Let $\gamma \in \mathscr{C}$ and take Ω as a simple subcurve decomposition of γ. Then any maximal element in $P(\gamma_\Omega)$ pulls back to a maximal element in $P(\gamma)$.*

Proof. Let us partition $V(\gamma) = A \cup B$ where $B = V(\Omega) = i(V(\gamma_\Omega))$, the vertices from the decomposition Ω. Then we know that each element $a \in A$ has either $a \subset b$ or $a \mathrel{\mathrm{L}} b$ or some $b \in B$. Let us take b as a maximal point in $P(\gamma_\Omega)$ and set $\tilde{b} = i(b)$. Then $\tilde{b}$ is maximal in $(B, \subset)$, since the posets $(B, \subset)$ and $P(\gamma_\Omega = (V(\gamma_\Omega), \subset)$ are equivalent. Thus, if $\tilde{b}$ were not maximal, we would have some $a \in A$ so that $\tilde{b} \subset a$. But then by maximality, we have $p \subset \tilde{b} \subset a$ for all elements $p \in B$, which is a contradiction since then a neither links nor is contained in any element in B. Thus, $\tilde{b}$ is maximal. □

4.3 Relation to Seifert Decompositions

Of course, the elementary forms we associate to a curve γ are not only simpler and easier to understand, but they also bear a nice resemblance to their originator. In fact, the essential piece of information preserved by the elementary form is the Whitney index of direct splits. Let us be more explicit with this important point. Suppose we had a curve γ and a standard SSD Ω. Let a be a vertex on γ_Ω and $\tilde{a}$ be the associated vertex on γ. Then we claim

$$whit(\gamma_{\tilde{a}}) = whit(\gamma_{\Omega_a})$$

In fact, this follows from the work of Seifert [17]. He introduced the so-called *Seifert Decompositions* of a curve, which are very similar to the simple subcurve decompositions defined here. In his decomposition, though, an uncrossing move is applied to each vertex which is a basepoint of a simple subcurve. Thus, the resulting decomposition is of the curve into a set of Jordan curves. See Figure 11. It follows that for any plane curve γ, the set of (generalized) SSD's has a natural bijection with the set of Seifert decompositions of γ.

Now, it was proved by both Gauss and Kauffman that the whitney index $whit(\gamma) = p - n$ where p, n are the number of positive and negative Jordan curves, respectively, in any Seifert decomposition of γ [9]. So, to prove our claim we need only note that any SSD including a vertex v necessarily contains a sub-SSD of γ_v. Hence, the value $p - n$ for the sub-SSD Ψ of Ω will be the same for γ_{Ω_v} and for $\gamma_{\tilde{v}}$ for any vertex $v \in V(\gamma_\Omega)$. A similar argument shows the same result holds for wrap splits as well.

Applying the fact that one can compute Whitney index through the signed number of simple subcurves in any simple subcurve decomposition, we can now easily prove a valuable lemma.

Lemma 4.7 (Whitney Index Through Decompositions)**.** *Let $\gamma \in \mathscr{C}$ and Ω be a free subcurve decomposition of γ. Then*

$$whit(\gamma) = \sum_{C \in \Omega} whit(C)$$

Proof. All we need to do is piece together simple subcurve decompositions of each subcurve $C \in \Omega$. Let us first rewrite $\Omega = \{C_i\}_{i=1}^k$ and take Ψ_i to be a simple subcurve decomposition of C_i. Then $\Psi = \cup_{i=1}^k \Psi_i$ is a simple subcurve decomposition of γ. Since $whit(C_i) = p_i - n_i$ where p_i is the number of positive subcurves in Ψ_i and n_i the number of negative subcurves, we conclude that

$$whit(\gamma) = \sum_{i=1}^k p_i - \sum_{i=1}^k n_i = \sum_{i=1}^k (p_i - n_i) = \sum_{i=1}^k whit(C_i)$$

□

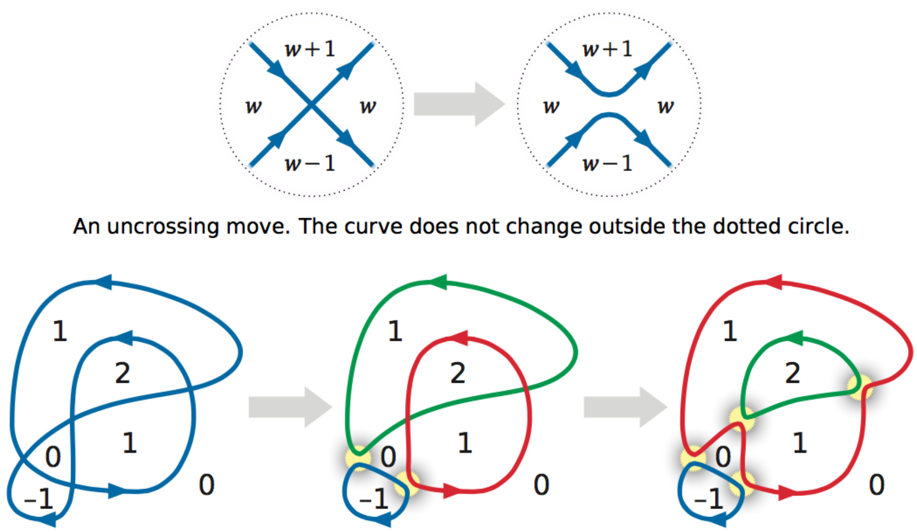


Figure 11: Figure from [9]. An uncrossing move shown above and two Seifert decompositions of a plane curve shown below.

5 Classes of Curves

We now define formally how we will glue two curves together, forming a connected sum, and how we will cut one curve into two pieces. The cuts should be thought of as a generalization of Blank cuts, which occur when γ_R or γ_L is a simple positive Jordan curve. Note that just like for the wrap of a curve, the connected sum of a curve only defines the new curve $\gamma'' = \gamma \# \gamma'$ up to signed intersection sequence.

Definition 5.1 (Connected Sum and Cuts)**.** *Suppose we have curves γ, γ' with disjoint images $[\gamma] \cap [\gamma'] = \varnothing$ and positive outer basepoints. Let us take a simple path P from $p = \gamma(0)$ to $q = \gamma'(0)$. Then, choose left and right ϵ neighbors P_L and P_R of P. Suppose that P_L connects p_L to q_L and P_R connects p_R to q_R. Choose $0 < t_L, t'_L, t_R, t'_R < 1$ so that $\gamma(t_L) = p_L, \gamma'(t'_L) = q_L, \gamma(t_R) = p_R, \gamma'(t'_R) = q_R$. We then define the **connected sum**, $\gamma'' = \gamma \# \gamma'$, of γ and γ' as the curve with positive outer basepoint p_L and and oriented image $\gamma|_{[t_L, t_R]} * P_R * \gamma'|_{[t'_R, t'_L]} * \overline{P_L}$. See Figure 12. We need to smooth the resulting curve at p_L, q_L, p_R, q_R, but this is no problem. Note that by construction, the intersection sequence of γ'' is the concatenation of the intersection sequences of γ and γ', with the indices of the vertices of γ' shifted up by $|\gamma|$.*

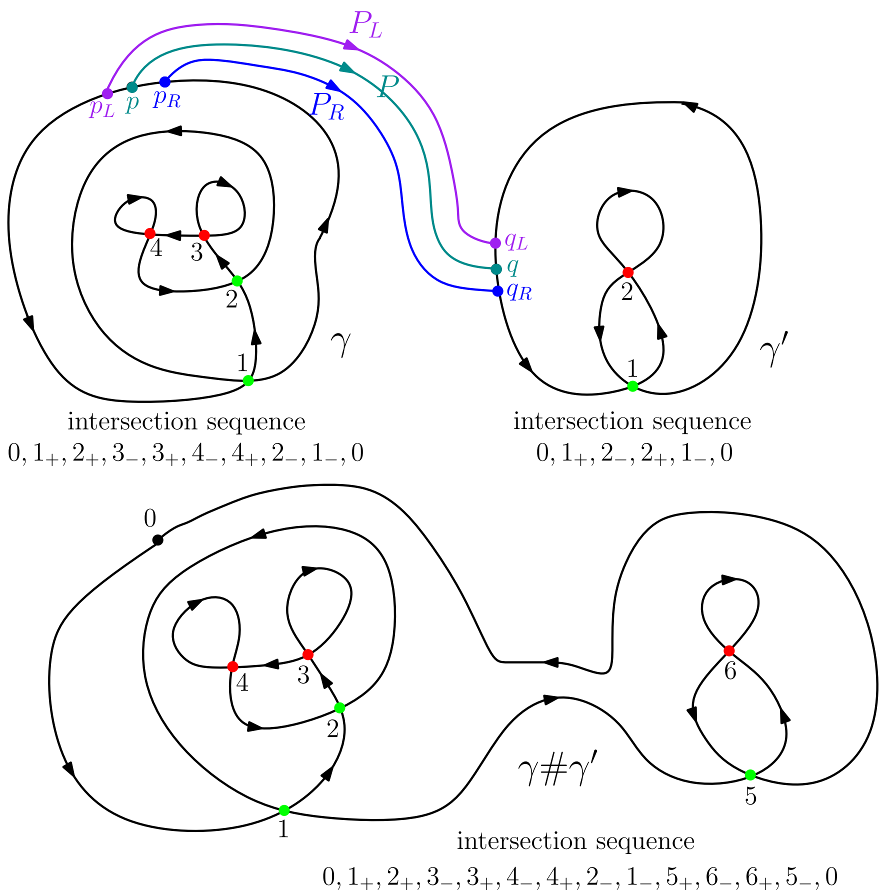


Figure 12: A connected sum of γ and γ'. Sample paths P, P_L, P_R are shown along the original curves. A representative of $\gamma \# \gamma'$ is shown below.

*Conversely, if we have a curve γ we find a simple path P with $P(0) \neq P(1)$ both exterior points on γ that are not vertices, then we may cut γ into two pieces along P. We call P a **valid** cut of γ. Write $P(0) = \gamma(t_0), P(1) = \gamma(t_1)$. Then we define the closed curves $\gamma_L = \gamma|_{[0,t_0]} * P * \gamma|_{[t_1,1]}$ and $\gamma_R = \gamma|_{[t_0,t_1]} * \bar{P}$. We call γ_L and γ_R the left and right cuts of γ with respect to P and write $\gamma = \gamma_L |_{\mathcal{C}}\, \gamma_R$.*

Note that if γ_R is a positive Jordan curve, then we have performed a Blank Cut. We now note a very nice property about these cuts, namely that the homotopy area is preserved, provided our cut does not intersect the curve away from its endpoints.

Theorem 5.2. *Suppose $C : [0,1] \to \mathbb{R}^2$ is a valid cut with $\gamma = \gamma_L \,|_C\, \gamma_R$. If $C \cap [\gamma] = \{C(0), C(1)\}$, then*

$$\sigma(\gamma) = \sigma(\gamma_L) + \sigma(\gamma_R)$$

Proof. It suffices to prove that any self-overlapping subcurve γ' of γ is either entirely contained in$[\gamma_L]$ or $[\gamma_R]$ or is split into two parts $[\gamma'_L] \subset [\gamma_L]$ and $[\gamma'_R] \subset [\gamma_R]$ so that $\gamma' = \gamma'_L \,|_C\, \gamma'_R$. Clearly winding numbers are preserved by such a cut since the resulting curves are disjoint, excluding $\mathcal{C}$. Suppose that γ' is split into two parts by γ. Then note that for some $\epsilon > 0$, we have the ϵ-neighborhood $N_\epsilon(C) = \{P \in \mathbb{R}^2 : P \notin F_{ext}, d(P, C) < \epsilon\}$ contained in the interior $int(\gamma')$. It follows then that we cut our immersed disk into two pieces by cutting along $\mathcal{C}$ and hence that we now have two immersed disks, $[\gamma_L]$ and $[\gamma_R]$. Each curve is the boundary of the immersed disk, they are self-overlapping, and we conclude

$$\sigma(\gamma) = W(\gamma) = W(\gamma_L) + W(\gamma_R) = \sigma(\gamma_L) + \sigma(\gamma_R)$$

□

5.1 Good and Ideal Curves

We are now ready for some easy classes of curves to understand. Both of these classes of curves respect the minimum homotopy area very nicely.

Definition 5.3. *We say a face F is **good** when its depth, the minimal number of edges crossed by a path from F to the exterior face, is equal to the winding number. If all faces on a curve γ are good and γ is consistent, we call the curve good.*

Now, good curves are particularly nice because their self-overlapping curves are all just loops. This holds since any non-simple self-overlapping curve C must have a negative loop and hence cannot be good. Additionally, any direct split on a good curve is also necessarily good (an easy proof can be constructed by induction). The following proposition follows very quickly from Theorem 4 of [14], which shows that good curves are Titus interior boundaries, and hence $\sigma(\gamma) = W(\gamma)$ through our equivalence of interior boundaries.

Proposition 5.4 (Good Curves are Ideal)**.** *Let γ be a good curve. Then γ is an interior boundary.*

We now consider another nice class of curves.

Definition 5.5. *We call a curve γ **ideal** if $\sigma(\gamma) = W(\gamma)$.*

Note immediately that this is a natural class of curves containing both the class of self-overlapping curves and the class of interior boundaries. In fact, there is quite a nice relation between ideal curves and the latter, as we now show. First, we need a definition.

Definition 5.6 (Sign-Changing Vertices)**.** *Let $\gamma \in \mathscr{C}$ and let v be a vertex of γ. If $\{F_1, F_2, F_3, F_4\}$ is the set of faces on $G(\gamma)$ indicent to v and $w(F_1) = w(F_2) = 0$ and $w(F_3) = -1, w(F_4) = 1$, then we call v a **sign-changing** vertex. See Figure 14.*

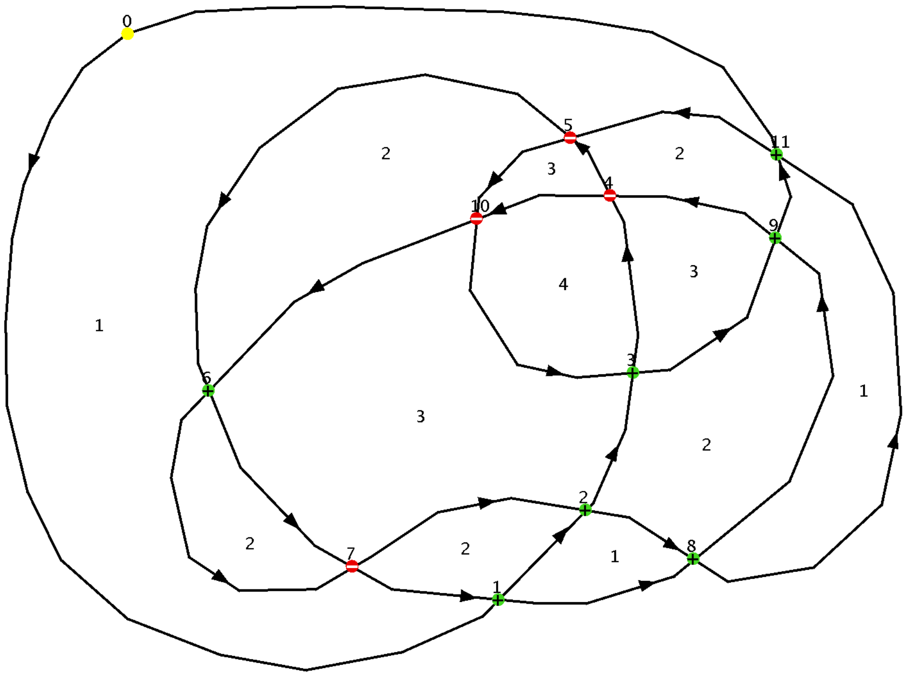


Figure 13: A good curve.

If a curve is an interior boundary, then we have a homotopy H which sweeps each face $W(F)$ times. This implies that we sweep only leftwards on our positive consistent regions and rightwards on our negative consistent regions. Intuitively, then, our sign-changing vertices on an ideal curve should be anchor points, otherwise we will sweep the wrong direction on F_3 or F_4. We rigidify this idea with the following theorem, clarifying the structure of ideal curves.

Theorem 5.7 (Classification of Ideal Curves)**.** *Let $\gamma \in \mathscr{C}$ and let $\{\gamma_i\}_{i=1}^k$ be the free subcurve decomposition of γ with vertex set $S = \{v_i\}_{i=1}^k \cup \{p_0\}$ where $\{v_i\}_{i=1}^k$ are all of the sign-changing vertices. Then γ is ideal iff each γ_i is an interior boundary.*

Proof. We first note that any ideal curve with consistent winding numbers must be a k-interior boundary, for if it had an optimal SO decomposition including curves of opposite orientations, then we would necessarily have a face F that is swept more than $W(F)$ times, a contradiction.

Next, we claim any sign-changing vertex v is an anchor of any optimal SO decomposition $\Gamma(\gamma)$. Suppose otherwise. In a tedious argument by cases, one can show that if v were not an anchor, we would have SO curves γ_i, γ_j of opposite orientations, one with F_3 in its interior, the other with F_4 in its interior, and γ_i and γ_j sharing a face adjacent to v in their interiors. This is a contradiction, for we will then sweep this face F more than $W(F)$. Thus, it is necessary that the sign-changing vertices are anchor points of a minimum homotopy. To conclude the proof, we observe that once we split at all such vertices v, we will have k curves $\gamma_1, \ldots, \gamma_k$ with consistent winding numbers. By the previous observation, for γ to be an interior boundary, it is necessary for each such γ_j, $j \in [k]$, to be ideal. By Lemma 3.5, each γ_j is an interior boundary. Conversely, if each γ_j is an interior boundary, then by construction, each face F with $|W(F)| \neq 0$ lies in the interior of exactly one interior boundary γ_j. Consequently, it will be swept $W(F)$ times when γ_j is contracted. It follows that γ is ideal if each γ_j is ideal. $\square$

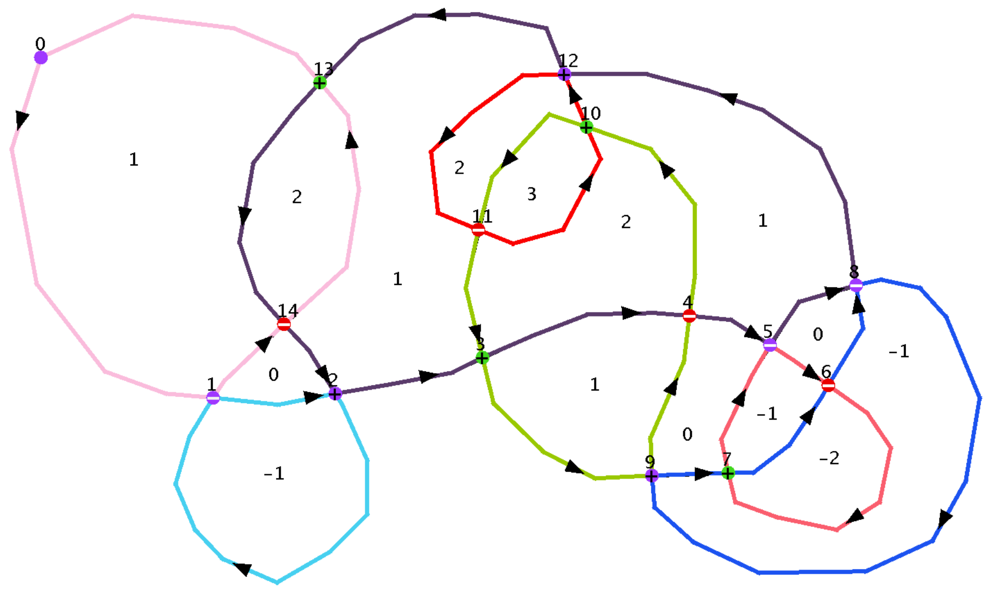


Figure 14: An ideal curve, with its minimum homotopy decomposition. The anchor points are shown in purple, with each SO curve in the SO decomposition shown in a different color. The sign-changing vertices with indices $1, 2, 5, 8, 9, 12$ *can all be seen to be anchor points.*

s

We now make a key observation. With Theorem 5.7, we can produce a naive polynomial time algorithm to compute whether a curve γ is ideal. First, split the curve at all sign-changing vertices to achieve a (potential) interior boundary decomposition $\Gamma = \{\gamma_i\}_{i=1}^{k}$. This takes only $O(n)$ time. We then need only compute if every γ_j is an interior boundary. Using the algorithm of Blank/Marx [4, 14], we can compute whether a curve is an interior boundary in polynomial time. Suppose this runtime is $O(f(n))$. We then note that in $O(nf(n))$ time we can compute whether a curve is an interior boundary, where $n = |\gamma|$.

Before we summarize our knowledge of some nice classes of curves, let us consider a final class of curves.

Definition 5.8. *Let us call a curve* ***basic*** *if all of its SO decompositions are SSD's.*

The lattice shown in Figure 15 summarizes the findings of this section. This lattice is actually even more informative than it initially appears. The three occurrences of dual inclusion are actually if and only ifs: a curve γ is a k-boundary iff it is consistent and ideal, γ is good iff it is basic ideal and consistent iff it is basic ideal and a k-boundary, and a curve is simple iff it is SO and good.

5.2 Basic k-Interior Boundaries

Interior boundaries were defined earlier, and serve as a nice generalization of SO curves, through their optimality with respect to minimum homotopy area. We now introduce a restrictive class of interior boundaries, which are, in a sense, the easiest to understand.

In terms of putting a k boundary γ into some sort of reduced form, we would like to simplify all k SO curves. This will give us a reduction of γ into a simpler k-boundary $\bar{\gamma}$. Now, there might

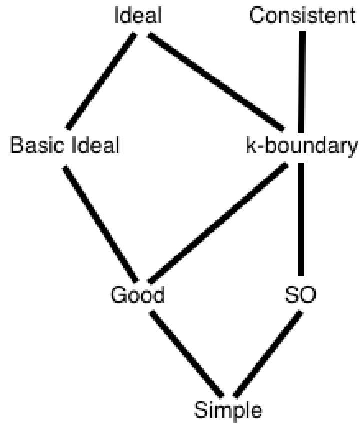


Figure 15: Lattice of curve classes, with inclusions going upwards.

be many ways to view γ as a k-boundary. Associated to each set of k anchor points $P = \{p_i\}_{i=1}^k$ of an optimal SO decomposition is an elementary form, γ_P, which is also a k-boundary. To define such an elementary form, we will essentially replace each SO curve γ_i based at p_i with a simple positive Jordan curve $\tilde{\gamma}_i$ which does not intersect any other $\tilde{\gamma}_i$. It turns out it is quite useful to look at an intermediary combinatorial object which will allow us to easily define this curve γ_P. We start with a simple observation: no two vertices $p_i, p_j \in P$ may be linked. This is true of any SO decomposition. Thus, we can form a tree T associated to P by setting $V(T) = P$ and forming edges between $p_i \neq p_j$ so that $p_i \subset p_j$ and there is no other $p_k \in P$ so that $p_i \subset p_k \subset p_j$. It turns out this tree, with some extra information will allow us to recover the (signed) intersection sequence of the desired curve γ_P. We now formalize this process.

Definition 5.9. *Let us say a k-boundary is **basic** if it has a SSDinto k positive simple subcurves. We denote the set of (positive) basic k boundaries with positive outer basepoint by E_k. Let us then define an equivalence relation $\sim$ on E_k by $\gamma \sim \gamma'$ iff γ and γ' have the same signed intersection sequences. We are particularly interested in $\mathcal{E}_k = E_k/\sim$, with all its members differing by signed intersection sequence.*

By an earlier observation in Section 4, since our basic k-boundaries are elementary, they have a unique simple subcurve decomposition Ω of γ into k positive simple subcurves. In other words, we have a unique SO decomposition, which is also a the unique (standard) SSD of γ. Now, our elementary k boundaries will be in bijection with a certain set of trees:

Definition 5.10. *Let T be a connected, planar rooted tree with exactly k vertices, and order pref-erences on the children of each node. Then we call T an elementary G tree. We denote the set of all trees by $\mathcal{S}_k$.*

The following result is equivalent to the Theorem on PG 27 of [3]. We note here that our elementary k boundaries are equivalent to Arnold's immersed circles of index $k+1$ with k double points. We present this proof for clarity of our new bijection. In fact, this approach of forming a tree based on relations of intersection points that are not linked, then traversing the tree to recover an intersection sequence will be appealed to later.

Proposition 5.11. *We have a bijection* $\phi : \mathcal{E}_k \to \mathcal{S}_k$.

Proof. We recall the Lemma "Existence of an Elementary Form", in which we showed we can form a tree T based off of any set of vertices which are the basepoints of a standard S.S.D. Take $\gamma \in \mathcal{E}_k$ and Ω as its unique S.S.D. Then we know there is a unique tree T formed by the process earlier, forming edges between nodes $u \subset v$ so that no other vertex w lies between the two in the poset $P(\gamma)$. We can then form a unique labeling on the children of each node based on the indices of the children in the intersection sequence of γ. This then yields a connected, rooted planar tree of the desired kind, with the root corresponding to the maximal element $\gamma(0)$. Any two different curves $\gamma, \gamma' \in \mathcal{E}_k$ will yield different trees, so this map is injective. To see that it is surjective, take any tree $T \in \mathcal{S}_k$ and we can do a post-order traversal of it to yield the signed intersection sequence of a curve $\gamma \in \mathcal{E}_k$ so that $\phi(\gamma) = T$. Let us first do a post-order traversal of T, labeling each node the first time we see it. Then, let us do a post-order traversal again, printing out the index of each node as we visit it, with a positive sign the first time and a negative sign the second time. By nature of the tree, this will yield the signed intersection sequence of a curve $\gamma \in \mathcal{E}_k$. Indeed, the curve will have exactly k positive intersection points. Moreover, it is an interior boundary, for given any vertex $v \in V(\gamma)$ we delete the direct splits of all nodes $u \subset v$, then $\gamma_v \backslash (\cup_{u \subset v} \gamma_u)$ will be a simple positive Jordan curve. Hence γ is a k boundary, for it has a S.S.D. into k positive simple subcurves, which is also a SO decomposition. By construction, $\phi(\gamma) = T$, so we are done. $\square$

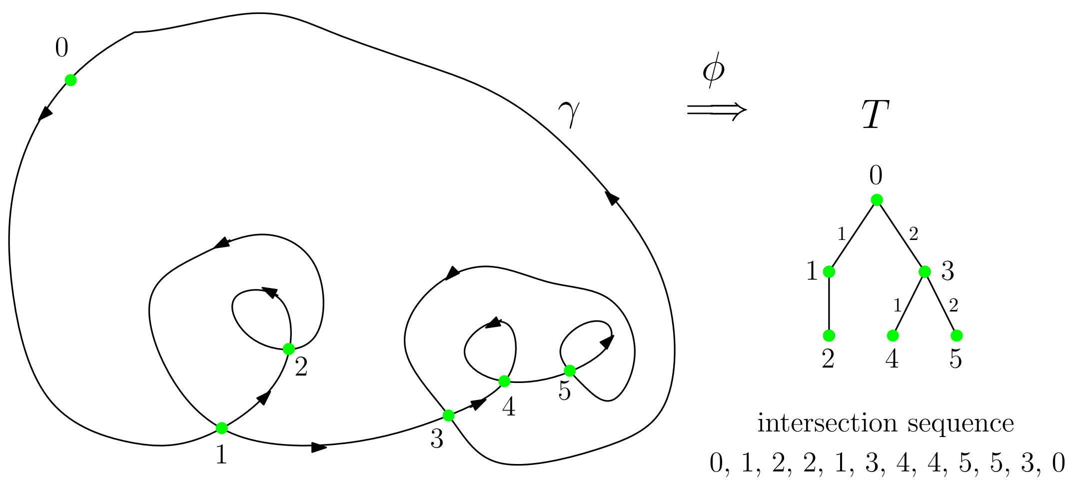


Figure 16: An elementary 5-boundary with its corresponding elementary G tree. Note that post-order traversal of T yields the (unsigned) intersection sequence of γ.

See Figure 16 for an example of the map ϕ from the previous proposition. The reader may note that the elementary G tree we assigned to a basic interior boundary γ is actually the dual-graph of $G(\gamma)$, with order preferences on the edges. This is indeed the case, but we prefer not to argue in terms of the dual graph.

6 Elementary SO Curves

We now study *elementary* SO curves. Here, we curves that are SO and elementary, following Definition 4.3. These curves serve as the 'simplest' examples of SO curves, and as the main theorem of this section will show, they are quite easy to classify. We will form a bijection between elementary SO curves and plane rooted trees with colored nodes, in a highly similar fashion to the bijection formed between basic k-boundaries and the elementary G trees. Indeed, the previous bijection employed many of the same techniques that will be relevant here.

6.1 G-R Trees

Scanning the conventions for the trees below, one should have in mind that the nodes in these trees will correspond to simple subcurves, or equivalently, with the basepoints of simple subcurves, of an appropriate curve. The green nodes correspond to positive vertices and the red nodes to negative vertices. The technical conditions imposed on these trees will allow us to swiftly recover the signed intersection sequence of an associated curve. The reader may notice that the following trees very closely resemble the dual graph of the corresponding elementary curves. Indeed, this is the case. We will return to this point in more detail shortly. While the following definition is quite lengthy, we believe it is a natural one, as demonstrated by the Theorem 6.11. See Figure 17 for an example of a GR tree.

Definition 6.1. *Let T be an undirected plane graph that is connected and rooted, with each node either green or red, and order preferences on the neighbors of each node. The order preferences are equivalent to a well-ordering of the neighbors $N(v) = \{u \in V(T) \mid \{u,v\} \in E(T)\}$ for each $v \in V(T)$. Since T is a plane graph, we impose a few additional conditions on the drawing of the trees, which will allow for simplification of later proofs. The idea behind these conditions is geometric in nature – we use vertical edges to indicate a simple subcurve lying inside another, and horizontal edges to indicate tangency of simple subcurves.*

1. *We call the distance d of a node v to the root the **depth** of v in T. For each node v, there is at most one vertical edge going from v to a lower depth. If such an edge $e = uv$ exists, we call u the **parent** of v.*
2. *Edges between same colored nodes are vertical, and go between nodes differing in depth by one.*
3. *Edges between different colored nodes are horizontal. That is, they may only connect nodes at the same depth.*

Additionally, we require the following conditions be satisfied:

- *(i) The root is green.*
- *(ii) For some integer $k \in \mathbb{Z}_{\geq 0}$, there are $k+1$ green nodes and k red nodes.*
- *(iii) Let P_v be a path from a node v to the root r. Viewing P_v as a string, let $p \subset P_v$ be the substring consisting of all nodes reached by vertical edges. Then the number of green and red nodes p_{green}, p_{red} on p satisfy $p_{red} \leq p_{green}$.*

For technical reasons, we need not have an order preference on the edge from a node to its parent.

*A graph meeting these conditions will be called **valid** G-R tree. Let the set of all such trees be denoted by $\mathscr{T}$. Suppose that T is valid and none of the induced subtrees rooted at a green node* *are valid. Then we call T* ***irreducible****. When drawing such trees, we will omit the trivial order preference of a node with one neighbor, excluding its parent.*

The other central object of this section, is the set of elementary curves to which these trees will correspond.

Definition 6.2. *Let $\gamma \in \mathscr{C}$ satisfy the following:*

(i) γ has exactly $2k$ intersection points, k of which are positive, the other k negative.

(ii) γ is positive consistent.

(iii) No intersection points are linked.

(iv) γ has positive outer basepoint $\gamma(0)$.

We call such a curve a ***balanced elementary curve****. Note that since these curves have no linked vertices, they are elementary, by Definition 4.3. We denote the set of all such curves as S. We are only interested in curves differing by signed intersection sequence. Thus, we again mod out again by signed intersection sequence and consider $\mathscr{S} = S/\cong$.*

Before we formalize the correspondence between our G-R trees on $2k+1$ nodes and balanced elementary curves with $2k$ vertices, let us first sketch some parallels between these two objects. If we consider each node of a G-R tree T as a simple subcurve of a curve γ, then we can see how the conditions imposed on the trees and the curves are equivalent: the size conditions on each are clearly equivalent, the lack of linkage on γ the curves corresponds to the simple subcurves not intersecting each other, except at points of tangency, and the strange condition (iii) ensures that γ is positive consistent. The rooted condition of the tree, which ensures that the associated curve has a consistent counterclockwise outer edge, is guaranteed by the positive outer basepoint, along with the absence of linkages on γ.

One may naturally think of the dual graph and wonder why we construct these strange G-R trees instead of just taking the dual graph of a balanced elementary curve γ and working with that tree T. Indeed, our trees are intimately related to the dual graph, but there are a few reasons to work with our trees rather than the dual graph. The main reason is that the nodes in the dual graph correspond to *faces* of the embedded planar multigraph, when we are actually interested in the *subcurves*.

Additionally, given a balanced elementary curve γ, one can recover its dual graph T' from the G-R tree T we will associate to γ. In fact, the dual graph of the curve γ can be achieved by a few edge flips from T. If we take each node u in T with a horizontal neighbor v and trade the edge from u to v for an edge between u and the unique parent w of v, we will recover the dual graph. Conversely, given the dual graph of γ, it is not possible, in general, to recover the G-R tree T associated to γ. This fact is due to the extra information carried by the edge preferences and colors of the nodes. Without these extra conditions, the dual graph and the G-R trees are equivalent.

We now formally show the mapping between the G-R trees and balanced elementary curves.

Lemma 6.3. *We have a map $\phi : \mathscr{T} \to \mathscr{S}$.*

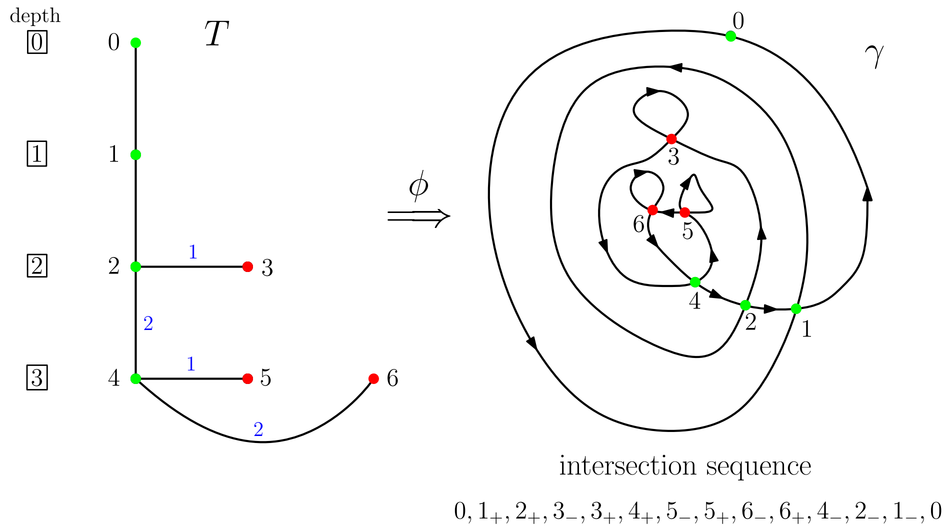


Figure 17: An irreducible G-R tree T with $k = 3$. The depths of the nodes are enclosed by rectangles. When nontrivial, the order preferences on the edges are labeled in blue. The elementary curve γ associated to T through the map ϕ is shown by its side. Note that post-order traversal of T yields the signed intersection sequence of γ.

Proof. Let us construct a curve $\gamma(T)$ given a tree $T \in \mathscr{T}$. We will do this by creating the signed intersection sequence of $\gamma(T)$. To do this, we will complete a post-order traversal of T. More precisely, we will recursively traverse the tree as follows: when visiting a node, go to each child in order, then return to the node. Let us do an initial post-order traversal of the tree and label each node as we discover it with label i, where we have labeled $i-1$ node already. We begin by giving the root the label 0. Afterwards, we do another post-order traversal to construct the signed intersection sequence. This is done as follows: when we first visit a node, we print its label once. Then, when we return, we print the label again, but with the opposite sign. We will give green nodes positive signs and red nodes negative signs for their first visit. Through this convention, we satisfy condition (i) of Definition 6.2.

We now check whether we have no linkages. Take any two intersection points p_i, p_j and let their associated nodes be v_i, v_j. Suppose one node lies in the subtree rooted at the other node. Without loss of generality, suppose v_2 lies in the subtree of v_1. We claim that we must have $p_2 \subset p_1$. First, let $S(v)$ be the induced subtree of T consisting of v, all its horizontal neighbors $u_1, \cdots, u_r$ and all the subtrees rooted at $v, u_1, \ldots, u_r$. Now, for the crucial observation: the direct split of v on this curve $\gamma(T)$ corresponds to the subtree $S(v)$. Since v_2 lies in $S(v)$, it is indeed the case that $p_2 \subset p_1$, since the intersection sequence will have a subsequence $p_1 p_2 p_2 p_1$. Otherwise, if neither node lies in each other's subtree, in which case neither p_1 S p_2.

Property (ii) of Definition 6.2 is the most difficult to verify. We note that $\gamma(T)$ is positive consistent iff the $wn(F) > 0$ for any face F inside a negative subcurve. As each red node r corresponds uniquely with a simple negative subcurve γ_r and there is a unique face F, which is the interior of γ_r, we have a bijection between the red nodes and all such faces F. To compute the winding number inside such a face, we simply look at the loops which we must cross to get to the exterior. One such path is easily given by our tree T. We observe this as follows. Traverse a path p_v from a red node v to the root r along edges in T. Each time we go up in the tree, we cross a loop. Thus, setting our count initially at -1 for the simple negative loop we begin in, we can increment our count by ±1, for each node we reach by going up along the path p_v. By condition (iii) of Definition 6.1, we can immediately see that $\gamma(T)$ is consistent. Since our basepoint $\gamma(0)$ corresponds to the root r of the tree T, which has only green children, it follows that the outer boundary of $G(\gamma(T))$ is a positive cycle, and hence our outer basepoint is positive. Consequently, our curve $\gamma(T) \in \mathscr{S}$ as desired. □

Let $\tau : V(T) \to V(\gamma)$ be the natural bijection between the vertices of a G-R tree and its associated curve $\gamma = \phi(T)$. The association between the subtree rooted at a node $S(v)$ and the direct split $\gamma_{\tau(v)}$ corresponding to that node, is quite useful and will be employed throughout this section. We now show that our map ϕ is actually a bijection. The proof is rather straightforward, following the ideas of the previous proof.

Lemma 6.4. *The map $\phi : \mathscr{S} \to \mathscr{T}$ is a bijection.*

Proof. It is clear by the construction that ϕ is injective, for different trees will give different traversals, which will yield different signed intersection sequences. To prove that ϕ is surjective, choose a balanced elementary curve $\gamma \in \mathscr{S}$. We construct a tree $T \in \mathscr{T}$ so that $\phi(T) = \gamma$. Examining our given curve γ, we note that since no intersection point links any other, we always have a loop which is simple. Thus, we can inductively decompose γ into simple subcurves. In fact, this gives us an intermediary bijection between the intersection points, including the basepoint $\gamma(0)$, and simple subcurves of γ. Let this bijection be ψ, so that $p_i \overset{\psi}{\Longleftrightarrow} \gamma_i$. Since no point links any other, the basepoint $\gamma(0)$ must correspond to a positive loop which contains all other loops. We will now describe a map τ between the subcurves γ_i and nodes of a graph T, so that $\phi = \tau \circ \psi$ is our desired bijection. We will check that all of the conditions of Definition 6.1 are satisfied. First of all, we can clearly see that we should assign γ_0, the outer simple subcurve, as the green root of our tree. Now, if any curve γ_i is tangent to γ_0, it must be completely contained inside and also positive. It follows that each such γ_i will be a green child of γ_0. We may then take any negative simple subcurves γ_- so that γ_- is tangent to a first level positive subcurve γ_+, they intersect at the basepoint p of γ_-, and add a red node r for γ_-, connected to the node v associated to γ_+. Completing this procedure yields all nodes of depth zero and one in our graph T. We may then iterate this process for each node in the first level. Order preference for each node's neighbors will be given naturally by the ordering on the curve. The base case is when an empty loop is reached. Such a loop will correspond to a leaf in the graph. We claim that once this procedure finishes, we will have a connected graph T that is a valid G-R tree. The graph G is connected because it remains connected at every phase of node addition to the graph. Since γ is normal and no intersection points are linked, any curve γ_i tangent to another γ_j, intersecting at just one vertex of the curve, will either be completely contained inside and hence turning the same way, or will be of the opposite orientation. In either case, we have a legal edge on T as described by criteria (2) and (3) of Definition 6.1. By construction, positive subcurves become green nodes and negative subcurves red nodes, so criterion (ii) is satisfied as well. Criterion (iii) follows because γ is consistent. Now, we have a vertical edge on T from a vertex v_i to another v_j only in the case that v_j is in correspondence with a loop γ_j completely contained inside γ_i, and γ_j is tangent to γ_i. Clearly, such a subcurve γ_i is unique, if it exists, with respect to another simple subcurve γ_j. Thus, T satisfies (1) of Definition 6.1, for each node in T has unique parents, if any. Hence, T is a valid G-R tree. □

Reviewing the proof of Lemma 6.4, we actually have bijections between vertices of γ, nodes of T, and simple subcurves decomposing γ. These bijections will be useful from here on out.

Unfortunately, the proof that these G-R trees are indeed trees is rather tedious, so we postpone it to the appendix. It is immediately clear that no cycles involving nodes of different depths may occur, by condition (2) of our G-R trees. Hence, the core of the proof is a technical argument precluding cycles on the same level.

Lemma 6.5 (G-R Trees are Trees)**.** *Let $T \in \mathscr{T}$. Then T is a tree.*

We now prove a collection of lemmas that will enable us to prove the main theorem of the section. First, we need a useful fact about SO curves: no empty positive loops may appear. This was proved by Titus in [22].

Lemma 6.6 (No Simple Positive Loops on SO Curves)**.** *Let γ be SO. Then there cannot be a direct split γ_+ on γ that is an empty positive loop, i.e., no edges of γ intersect γ_+.*

The following lemma will be crucial for the proof of the main theorem of this section.

Lemma 6.7 (+− Balanced Deletion on Elementary SO Curves)**.** *Let $\tilde{\gamma}$ be a free subcurve on a positive elementary SO subcurve γ so that $\tilde{\gamma} = \gamma_+ \vee \gamma_-$ where γ_- is a simple negative loop, and γ_+ is positive SO. Then γ is SO only if $\gamma' = \gamma \backslash \tilde{\gamma}$ is SO.*

Proof. Suppose γ were SO. Then $\gamma \backslash \gamma_-$ is a 2-boundary by Lemma 3.9. Applying Lemma 3.5, we know that we have a SO decomposition $\Omega = \{\gamma_1, \gamma_2\}$ so that both γ_1, γ_2 are positive SO. We claim that γ_+ is inevitable in such a decomposition, so that it is ideal. Since γ_+ is free on $\gamma \backslash \gamma_-$, no positive SO curve C on $\gamma \backslash \gamma_-$ can contain it since no SO curve can have an empty positive SO subcurve. We sketch a proof of this fact here. The key ingredients are the fact that SO-ness is invariant with respect to signed intersection sequences and the equivalence SO iff there is a Blank cut decomposition. The main idea is that since the subcurve is empty, we can redraw γ, respecting the signed intersection sequence, so the empty SO subcurve γ_+ lies within an ϵ ball $B_\epsilon(p)$ of its basepoint p. Then, to get a cut decomposition of this new plane curve γ', which has the same Titus intersection sequence, we would inevitably have to cut apart the SO subcurve γ_+ on γ' until it becomes a positive loop, a contradiction to Lemma 6.6. Thus, no positive SO subcurve C can contain γ_+ and hence it is ideal on $\gamma \backslash \gamma_-$. By Theorem 3.5, we know $W(\gamma \backslash \gamma_-) = \sigma(\gamma \backslash \gamma_-)$. Now, $W(\tilde{C}) + W(C \backslash \tilde{C}) = W(C)$ for any curve C and any free subcurve $\tilde{C}$. With this fact and that $\gamma \backslash \{\gamma_-, \gamma_+\} = \gamma'$, we have

$$\sigma(\gamma') + \sigma(\gamma_+) = \sigma(\gamma \backslash \gamma_-) = W(\gamma \backslash \gamma_-) = W(\gamma') + W(\gamma_+)$$

As γ_+ is SO, we have $\sigma(\gamma_+) = W(\gamma_+)$ and hence $\sigma(\gamma') = W(\gamma')$. By Theorem 3.4, γ' is SO since $whit(\gamma') = +1$. □

6.2 Top-Heavy and Irreducible Curves

To elucidate the connection between our G-R trees and our elementary SO curves, we need to study a couple of classes of nice curves.

Definition 6.8. *Let $\gamma \in \mathscr{C}$. Then if γ has no positive SO direct splits, we call γ* ***irreducible****.*

As we will soon see in Theorem 6.11, our definitions were aligned so that a G-R tree T is irreducible iff $\gamma = \phi(T)$ is irreducible. A special case of irreducibility is of particular interest to us:

Definition 6.9. *Let $\gamma \in \mathscr{C}$. If $whit(\gamma_v) \le 0$ for all proper direct splits, we call γ* ***top-heavy****.*

If we have a curve of Whitney index 1 that is top-heavy, then intuitively the negative subcurves are near the 'top', or 'back' of the curve and the positive subcurves densest near the 'bottom', or 'front' of the curve. See Figures 2, 4, and 17 for some examples of top-heavy curves. Note that a top-heavy curve is irreducible since any positive SO curve C has $whit(C) = +1$. In fact, in the elementary case the two are equivalent, as we are about to show.

Before our major results of this section, we need a useful lemma.

Lemma 6.10 (Direct Splits with $whit = +1$)**.** *Let γ be a curve with $whit(\gamma) \ge 2$. Then γ has a proper direct split γ_v with $whit(\gamma_v) = 1$.*

Proof. Let us apply Lemma 2.7 to get a direct split decomposition $\Omega = \{\gamma_i\}_{i=1}^k \cup \{C\}$ of γ. Then $2 = whit(\gamma) = \sum_{i=1}^k whit(\gamma_i) \pm 1$ by Lemma 4.7, depending on the orientation of C. Either way, $\sum_{i=1}^k whit(\gamma_i) \ge 1$ and hence we have some γ_j with $whit(\gamma_j) \ge 1$.

We now form a non-extendable sequence $\{\gamma_i\}_{i=1}^k$ of direct splits on γ, so that γ_i is a direct split on γ_j for $i > j$ and $|\gamma_i| < |\gamma_j|$ and $whit(\gamma_i) \ge 1$ for each direct split. By the previous argument, we know such a sequence exists. Since γ is a finite curve, this sequence must eventually terminate. Suppose γ_k is the final curve, so that there is no direct split on γ_{k+1} such that we can extend our sequence. Then we cannot have $whit(\gamma_k) \ge 2$. Thus, $whit(\gamma_k) = 1$. Since direct splits of direct splits are direct splits on the original curve, γ_k is a direct split of γ. ☐

6.3 Classification of Elementary SO Curves

We are now ready to prove our main correspondence – elementary SO curves correspond with irreducible trees.

Theorem 6.11 (Classification of Elementary SO Curves)**.** *Let $\gamma \in \mathscr{S}$. Then γ is SO iff it is irreducible.*

Proof. We first prove necessity. This follows easily by contrapositive. In the proof of Lemma 6.7, we showed that a SO curve cannot have an empty positive SO subcurve. Thus, if an elementary curve is not irreducible, it is not SO, for all subcurves are empty.

Sufficiency is trickier to see. Let γ be a balanced elementary curve. We now proceed by induction on the size $2k$ of the curve. The base case $k = 0$ is trivial, for γ is a simple positive Jordan curve in this case.

For the general inductive case, suppose we have a curve $\gamma \in \mathscr{S}$ with $|\gamma| = 2k$ and assume that and any smaller balanced elementary curve is SO iff it is irreducible. We first note that the tree $\gamma(T) = \phi^{-1}(\gamma)$ is irreducible as well, by inductive hypothesis.

We claim that γ is actually top-heavy. Suppose otherwise. Then we have a direct split γ_v with $whit(\gamma_v) \ge 1$. Either $whit(\gamma_v) = +1$ exactly or $whit(\gamma_v) \ge 2$, in which case we can apply Corollary 6.10 to find a direct split γ_{v_u} on γ_v with $whit(\gamma_u) = 1$. A direct split of a direct split is a direct split of the original curve, so either way we can obtain γ_u as a proper direct split of γ with $whit(\gamma_u) = 1$. Let us then continue this process to form a maximal sequence $\{\gamma_i\}_{i=1}^k$ so that $whit(\gamma_i) = 1$, γ_i is a direct split on γ and $|\gamma_{i+1}| < |\gamma_i|$. Then γ_k must be top-heavy, or else we could extend our sequence. But then γ_k is irreducible. Since any direct split γ_v has v as its basepoint, on an elementary curve, we automatically have v has an outer basepoint on γ_v and v will be a positive/negative basepoint depending on $\mathrm{sgn}(v)$. We claim that the direct split γ_k has a positive outer basepoint. Suppose otherwise. Then when we take a direct split decomposition $\Omega = \{C_i\}_{i=1}^l \cup \{C\}$ of γ_k we will have $whit(C) = \pm 1$, depending on the sign of the basepoint. In the case that $\mathrm{sgn}(v) = -1$, then, we see that $\sum_{i=1}^l whit(C_i) \geq whit(\gamma_k) + 1 = 2$, so that $whit(C_j) \geq 1$ for some j. This is a contradiction, because either $whit(C_j) = +1$, or we can apply Lemma 6.10 again to get a direct split C' of C_j, with $whit(C') = 1$. Since C' which is a direct split of γ_k and $|C'| < |\gamma_k|$, this contradicts the maximality of our sequence. So, indeed, γ_k is a top-heavy, and hence irreducible curve, with $whit(\gamma_k) = +1$, with positive outer basepoint. By inductive hypothesis, it is SO. This contradicts the irreducibility of γ, so we conclude it must be top-heavy.

Let us now examine $\gamma(T)$. In particular, we consider the set S of maximally connected green nodes, containing the root r of $\gamma(T)$. We call S the *skeleton* of $\gamma(T)$. Obviously, S is non-empty. By Lemma 6.5, we know that T is indeed a tree. Thus, we have a unique path on the tree P_v for any node $v \in T$ from v to the root u. Let us do a post-order traversal on the tree and set v to be the last red node touched by this traversal. It follows that once we leave v we will touch at least one green node w_1. In fact, we must touch at least two green nodes, including possibly r itself, on the path P_v. This is true since v cannot be directly connected to the root by definition of G-R trees. Thus, we know P_v has a substring $vw_1w_2\cdots u$ where $w_1 \neq w_2$ are both green nodes. Since each of these nodes corresponds to a simple subcurve, it follows that we saw $j\cdots i\cdots ij$ as a substring in the intersection sequence, where p_j is the intersection point associated to w_1 and p_i is the intersection point associated to v. Since our curve is elementary, we may find a simple path from the negative loop γ_- corresponding to v to the positive subcurve γ_+ corresponding to w_2. Geometrically, it is clear that this cut is a Blank Cut. Indeed, we can trace out its path to see that it bounds a simple positive cycle. Suppose we perform this cut. Then we will delete the negative vertex p_i and the positive vertex p_j and leave all other vertices alone. By the form of the intersection sequence, for any vertex u, we either have $p_i, p_j \subset u$ or $p_i, p_j \; \mathrm{S} \; u$. Applying the Titus' formula from [21],

$$whit(\gamma_{p_k}) = \sum_{\sigma \subset p_k} \mathrm{sgn}(\sigma) + 2 \sum_{\sigma L_l p_k} \mathrm{sgn}(\sigma)$$

we can see that $whit(\gamma'_{p_k}) \leq whit(\gamma_{p_k}) \leq 0$ where $\gamma \to \gamma'$ by the simple cut. Here, we overload p_k to indicate the original vertex on γ and the corresponding vertex on γ' after the cut. Hence, top-heavyiness is preserved. Now, all Blank cuts can be performed by a regular left sense-preserving homotopy. Write $\gamma \xrightarrow{H} \gamma'$ where H performs the Blank cut. Since $|\gamma'| < |\gamma|$, it is SO, by inductive hypothesis, since it is also elementary. We can then reverse the homotopy to get a regular right sense-preserving homotopy $\overline{H}$ to see that as $\gamma' \xrightarrow{\overline{H}} \gamma$ and hence γ is SO by Lemma 3.7. $\square$

In the process of this proof, we actually showed the following:

Corollary 6.12. *Let γ be a balanced elementary curve with $whit(\gamma) = +1$ and positive outer basepoint. Then the following are equivalent:*

1. *γ is SO.*
2. *γ is irreducible.*
3. *$T(\gamma)$ is irreducible.*

4. γ is top-heavy.

We conclude this section with a nice alternate characterization of top-heavy curves, which applies generally.

Corollary 6.13. *Let $\gamma \in \mathscr{C}$ have $whit(\gamma) = +1$. Then the following are equivalent.*

1. γ is top-heavy.

2. For any SSD Ω of γ, γ_Ω is top-heavy.

3. For any SSD Ω of γ, γ_Ω is SO.

Proof. By Theorem 6.11, we already have (2) $\Leftrightarrow$ (3). Now, clearly (1) implies (2). We can see that (2) implies (1) by contrapositive: if γ were not top-heavy, then it would have a SSD Ω so that γ_Ω were not top-heavy, just by ensuring that Ω contains a SSD of a direct split γ_i on γ with $whit(\gamma_i) = 1$. □

7 Direct Splits of SO Curves

We are now prepared to show a few surprising and unintuitive results on the direct splits of positive SO curves, by applying our theory of elementary SO curves.

7.1 Arbitrarily Large Interior Boundaries on SO Curves

To begin, consider the following naive claim:

Claim 7.1. *Let γ be positive SO. Then any direct split $\tilde{\gamma}$ on γ with $whit(\tilde{\gamma}) < 0$ is a negative loop.*

In fact, this claim is quite wrong, and as evident from the bijection between elementary SO curves and irreducible G-R trees. In fact, this claim fails in two ways: we can have both $-k$-boundaries for arbitrary large k and extremely non-simple negative SO curves appear as direct splits. We now explicitly show the former.

Proposition 7.2 (Negative Interior Boundaries on Positive SO Curves)**.** *Take $k \in \mathbb{Z}_+$. Then there is a SO curve γ with a direct split that is a $-k$ boundary.*

Proof. We tackle the problem all at once by defining a family of SO curves $\{\gamma_i\}_{i=1}^k$ so that γ_i has a direct split that is a $-i$-boundary. We accomplish this task swiftly with our bijection between elementary SO curves and irreducible G-R trees. Let T_k be the G-R tree with a chain of $k+1$ green nodes connected vertically, each with one parent and one child, except for the root. Then, connect the deepest green node horizontally to a red node u, which in turn is the top of a vertical chain of k red nodes, connected in the same fashion. Then we can easily see that such a tree is irreducible and hence its associated curve $\gamma_k = \phi(T_k)$ is top-heavy and thus SO by Theorem 6.11. Now, the direct split γ_{k_u} on γ_k is a $-k$-boundary, as seen by the fact that the subtree $S(v_k)$ of the corresponding node in T_k is a chain of k red nodes. □

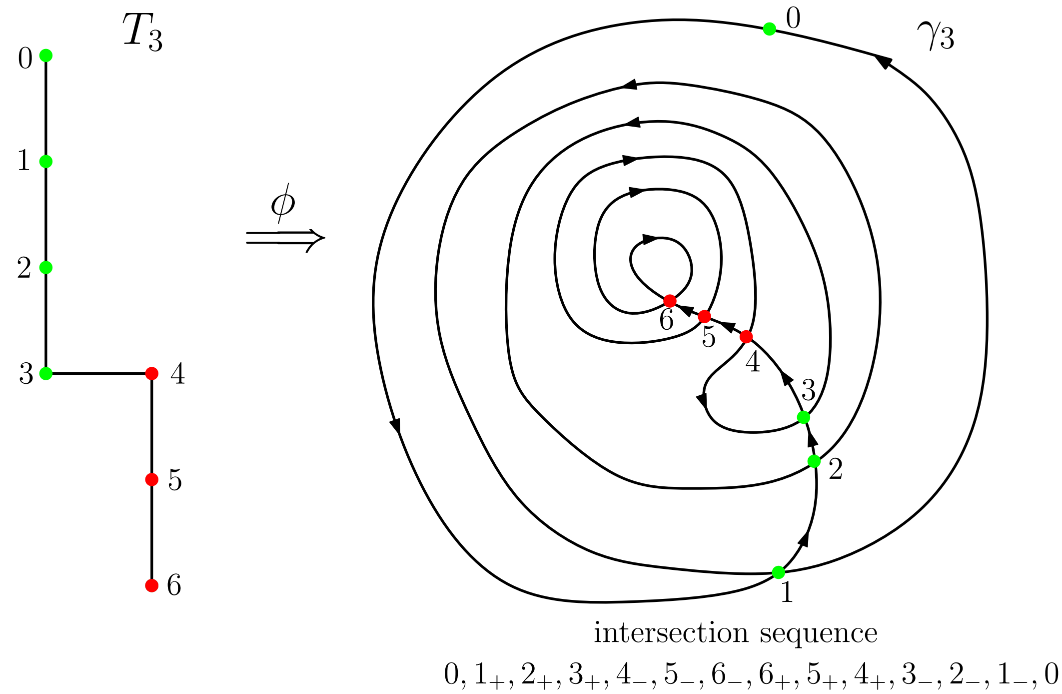


Figure 18: The tree T_3 and its corresponding elementary SO curve γ_3. The direct split of the 4^{th} vertex is an elementary -3 boundary.

The curves $\{\gamma_k\}_{k=1}^{\infty}$ from the previous proof are quite beautiful, with the first k intersection points all positive, the last k negative. The curve γ_3 and its associated G-R tree T_3 are shown in Figure 18.

Just as $-k$-boundaries with k arbitrarily large can appear as a direct split on positive SO curves, so too can $+k$-boundaries.

Proposition 7.3. *Choose $k \in \mathbb{Z}_+$. Then there is a SO curve γ with a direct split that is a k-boundary.*

Proof. To prove this, we will first define a curve γ_k associated to an elementary G-R tree T_k. Let T_k be the G-R tree, which has a chain of $k+1$ green nodes so that every node, except the root, is connected to exactly 1 red node. Moreover, form the order preference of each green node, other than the root, with the red node first and the green child second. Now, we homotope γ_k to a curve γ_k', which does have a direct split that is a k-boundary. By the construction of the map $\phi : \mathscr{T} \to \mathscr{S}$, we have an edge e from the deepest red node u to its green neighbor v. During our traversal of the tree, to form the signed intersection sequence of γ_3, after we visit u twice, we will visit v, then the green parent w of v, by our order preference. Thus, we have a substring $uuvw$ in our intersection sequence. Since our vertices are in bijection with simple subcurves, we know that after the edge e we lie on an edge of some positive simple subcurve γ^w with basepoint w. Now, if we do a II_b on e to force e to intersect the edge (v, w), we will create a positive loop. By the homogeneous structure of the tree T_k, we can repeat this action. In our traversal of the tree T_k to for the signed intersection sequence, from the vertex u onwards, we will create a substring $uuw_1w_2 \cdots w_{k+1}$ where $w_1 = v, w_2 = w$, and the root $r = w_{k+1}$. Let P be the path from u to v which we just performed a II_b on. Then we can repeatedly perform II_b moves on P to force it to intersect each of the edges

(w_i, w_{i+1}), up to, and including the final edge (w_k, w_{k+1}). Each II_b move will create a new positive and negative vertex. Let v be the final positive vertex created and suppose that $\gamma_k \overset{H}{\to} \gamma_k'$ where H performs the k II_b moves. Then by our construction, the direct split $\tilde{\gamma}_+ = \gamma'_{k_v}$ is an elementary k-boundary. γ_k' is SO by Lemma 3.7, so we are done. $\square$

The curve γ_3' from Proposition 7.3 is shown in Figure 19.

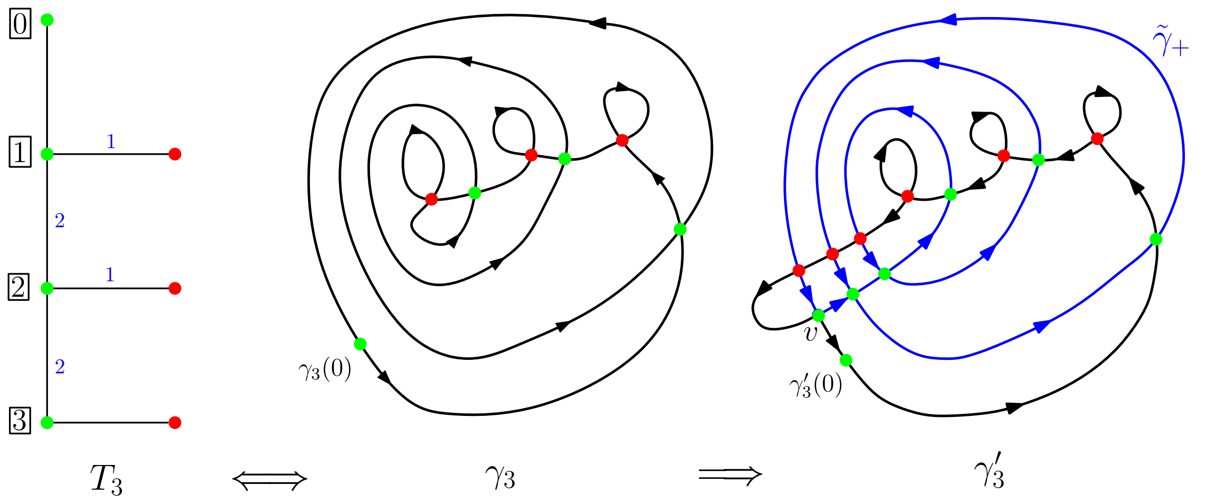


Figure 19: The tree T_3 from Proposition 7.3 and its associated elementary SO curve γ_3. After application of three II_b moves to the right, the direct split $\tilde{\gamma}_+$, seen to be a 3-boundary, is created.

7.2 Arbitrary Negative and Positive SO Direct Splits on SO Curves

We now show the other half of our prior claim – non-simple negative SO subcurves can appear as direct splits on SO curves. In fact, much more is true.

Proposition 7.4 (Negative SO Subcurves on Positive SO Curves)**.** *Let γ be a positive SO curve. Then there is a positive SO curve C so that $-\gamma$ is a direct split.*

Proof. Let us first take $-\gamma$ and a left ϵ neighbor of it, traversed in the opposite direction, $\gamma' = L_\epsilon(-\gamma)$. Choose an outer point p on $-\gamma$ then take its ϵ copy p'. Since these points are outer to both $-\gamma$ and γ', respectively, we can construct simple paths P_1, P_2 from p' to p and p to p', respectively, so that P_1, P_2 intersect only at their common endpoints. We now glue $-\gamma$ and γ' together along the paths P_1, P_2. Consider $C = P_1 * (-\gamma) * P_2 * \gamma'$, after smoothing at p and p', so C is normal and regular. We claim C is is our desired curve. First of all, the intersection point p has its direct split as $-\gamma$, by construction.

We now show that C is SO. First, note that the bigon with intersection points p, p' is empty, so we have an available II_a move. In fact, this move can be applied to the left. We claim that we actually have a sequence of II_a moves to the left to simplify C entirely. If we can construct such a regular homotopy, then C is SO by Lemma 3.4. Since we begin with a II_a move available, let us assume that after applying a II_a move the curve is non-simple, or else we are done. Then, look at the edge e where the last II_a move was applied, trace the path from the intersections of the bigon previously simplified, further inwards on the curve. But upon doing this we note that beyond the end of our bigon, the edges of the bigon extend to edges e_1, e_2 that are ϵ neighbors. It follows that once we apply our II_a move we will create another bigon. As these edges continue in the same direction as their predecessors, we must be able to again collapse this bigon with a II_a move to the left. Performing all such II_a moves gives us a regular left sense-preserving homotopy that simplifies C. We can then easily contract the simple positive Jordan curve remaining to obtain a regular left sense-preserving nullhomotopy of C. □

Given a positive SO curve γ, we will call the procedure in the above proof, the *doubling* of γ. See Figure 20 for an example of the doubling process. Iterating this doubling process, we can create some very strange SO curves.

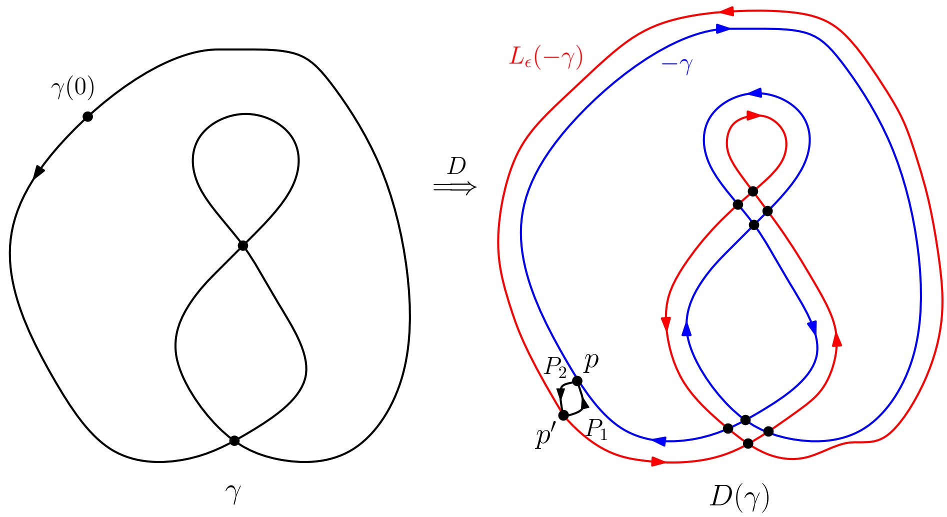


Figure 20: A small self-overlapping curve γ and its doubling $D(\gamma)$.

Corollary 7.5. *Given any positive integer k, we have a positive SO curve γ with a sequence of SO direct splits $\{\gamma_1\}_{j=1}^{2k}$ so that γ_{2j+1} is positive SO, γ_{2j} is negative SO, and γ_i is a direct split on γ_j for $i > j$.*

Proof. Let $D(C)$ denote the doubling of a positive SO curve C. Then we simply take $\gamma = D^k(C)$, for it has at least $2k$ free copies of non-simple SO subcurves, half of which are negative, half of which are positive. The direct split property is easily seen. □

We can actually play similar games with positive SO direct splits on positive SO curves.

Proposition 7.6 (Positive SO Direct Splits on Positive SO Curves)**.** *Let γ be a positive SO curve. Then there is a positive SO curve C such that γ is a direct split on C.*

Proof. Recall that with any positive SO curve, we can form a regular left sense-preserving homotopy H so that $\gamma \overset{H}{\to} \gamma'$ where γ' is a simple positive Jordan curve. This can be done by performing a sequence Blank cuts, each of which can be executed through a regular left sense-preserving homotopy. Thus, looking at $\overline{H}$, we know that for any SO curve ζ and any positive Jordan curve ζ' we have a regular right sense-preserving homotopy H_ζ which homotopes ζ' to ζ. Thus, we need only find a SO curve with a positive loop, and we can force our positive loop to become any SO curve ζ we like, while preserving SO-ness with H_ζ. Let us take our curve γ_1' from Proposition 7.3. Then, we can take the regular right sense-preserving homotopy H_γ with respect to our desired curve γ and homotope the positive loop with H_γ until it has the same signed intersection sequence as γ. Since this homotopy operates on the whole curve γ_1', let us write $\gamma_1' \xrightarrow{H_\gamma} C$. Then C is SO by Lemma 3.7 and has γ appear as a direct split. $\square$

8 Top Heavy Curves

In this section, we explore a few natural generalizations of elementary SO curves of Section 6, with the goal of seeing how to generalize the "Classification of Elementary SO Curves".

8.1 Top-Heaviness: Local vs. Global

We first recall the Corollary 6.12, which told us that for an elementary curve γ with positive outer basepoint and $whit(\gamma) = +1$, the following are equivalent:

(i) γ is SO.

(ii) γ is top-heavy.

(iii) γ is irreducible.

(iv) $T \iff \gamma$ is irreducible, where T is the elementary G-R tree associated to γ.

We saw in Section 7 by Proposition 7.6 that not every SO curve γ is top-heavy. Thus, the best we can hope is that the sufficiency of top-heaviness for SO-ness will generalize, assuming $whit = +1$ and positive outer basepoint. We proceed towards proving that this is indeed the case. It turns out that to find nice generalizations of elementary SO curves, we need to work with equivalence classes of plane curves with the same image and orientation. We call the set

$$E[\gamma] = \{C \in \mathscr{C} : [C] = [\gamma], orientation(C) = orientation(\gamma), C(0) \ outer\}$$

a **planar image**.

Given a planar image $E[\gamma]$, any two representatives $\gamma, \gamma' \in E[\gamma]$ will have the same free subcurves, despite the fact that the direct splits and wrap may differ. See Figure 21 and note that the direct splits γ_1 and γ_1' differ, though $\gamma, \gamma' \in E[\gamma]$. In fact, $\gamma_1 \cong \gamma_{1^*}'$ and $\gamma_{1^*} \cong \gamma_1'$. We now generalize relations to the set of free subcurves of a planar image, allowing us to circumvent basepoints altogether. When the planar image $E[\gamma]$ under discussion is clear, we will notate $\mathcal{S}$ as the set of free subcurves of the planar image and $\mathcal{S}_i = \{\tilde{\gamma} \in \mathcal{S} \,|\, whit(\tilde{\gamma}) = i\}$ as the free subcurves of Whitney index i.

Definition 8.1 (Free Subcurve Relations)**.** *Let $E[\gamma]$ be a planar image. Let us take any two free subcurves $\gamma_1, \gamma_2 \in \mathcal{S}$. We define $\gamma_1 \subset \gamma_2$ exactly as before: $\gamma_1 \subset \gamma_2$ iff $[\gamma_1] \subseteq [\gamma_2]$. We then define $\gamma_1 \ S \ \gamma_2$ iff $[\gamma_1] \cap [\gamma_2] \subset V(\gamma)$ and $\gamma_1 \ L \ \gamma_2$ otherwise. Note that our old formulation of relations $u \ R \ v$ in Definition 2.1 is exactly the same as the relations between direct splits – $\gamma_u \ R \ \gamma_v$.*

Remark: It's not too hard to see that this definition is equivalent to the following: $\gamma_1 \subset \gamma_2$ iff the image of γ_1 is completely contained in the image of γ_2, γ_1 L γ_2 iff they share edges, but there is not complete containment in either direction, and γ_1 S γ_2 iff they share at most vertices (possibly none).

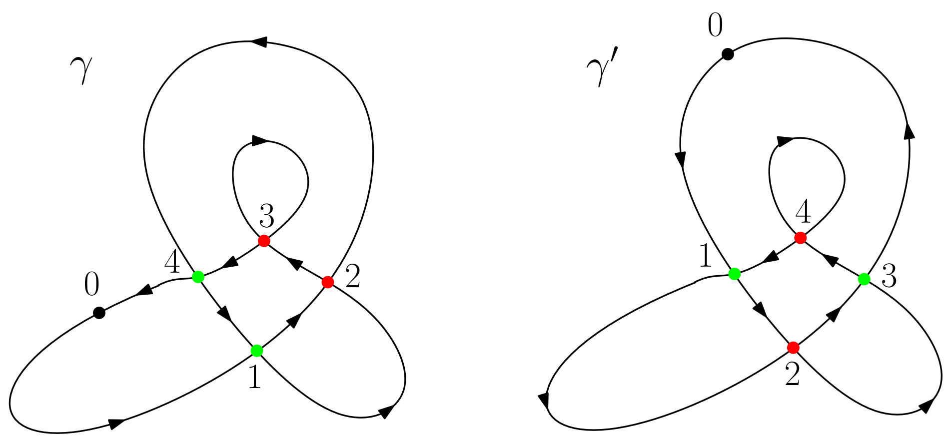


Figure 21: Two representatives γ, γ' of the same planar image so that γ is not top-heavy, but γ is.

Given an arbitrary planar image $E[\gamma]$, a representative $\gamma \in E[\gamma]$, and a vertex v on the planar image, we have no way of knowing a priori which of the two free subcurves with basepoint v, call them $\gamma_1, \gamma_2 \in \mathcal{S}$, will be the direct split of v and which will be the wrap split of v on γ. Indeed, one can think of choosing a basepoint $r \in [\gamma]$ as a way of *specifying* which free subcurves are direct splits and which are wrap splits. Thus, it is not too surprising that our old notion of top-heaviness, Definition 6.9, for a fixed plane curve, does not extend to a well-defined notion of top-heaviness for planar images. For any pair of representatives $\gamma, \gamma' \in E[\gamma]$, we would need γ to be top-heavy iff γ' is top-heavy. Unfortunately, this is not the case. See the example in Figure 21. The problem here is that top-heaviness depends on which free subcurves are direct splits, which is dependent on basepoint. In fact, top-heaviness is extremely dependent on basepoint. With this fact in mind, we now offer a new, more careful definition of top-heaviness for planar images, yielding two new classes of curves.

Definition 8.2 (Top-Heavy Planar Images)**.** *We call an equivalence class $E[\gamma]$* ***locally top-heavy*** *iff there exists a representation $C \in E[\gamma]$ such that C is top-heavy. In the case that each representative $C \in E[\gamma]$ is top-heavy, we call the planar image $E[\gamma]$* ***globally top-heavy****.*

Now, we precluded inner basepoints in our definition of planar images for good reason. Consider the example curve γ in Figure 22. Since γ is inconsistent, we can immediately see that it is not SO. Yet, γ is top-heavy. With respect to top-heaviness, inner basepoints can lead to illegitimate, or uninteresting curves, like γ, in the sense that they have no hope of being SO.

8.2 Properties of Globally Top-Heavy Planar Images

We now proceed to show the incredibly special nature of globally top-heavy planar images. These curves are actually generalizations of elementary SO curves. Since any elementary SO curve is

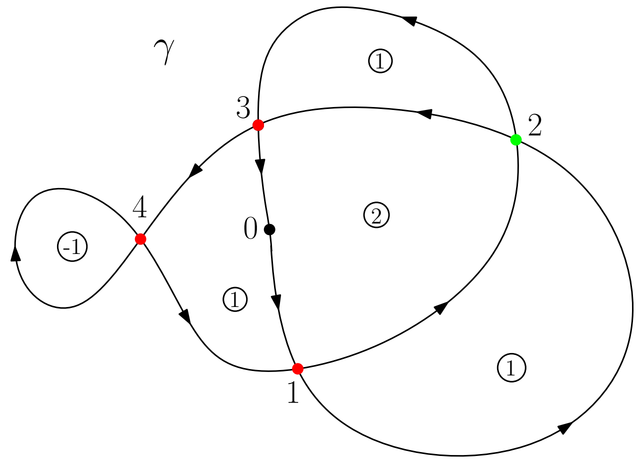


Figure 22: A plane curve γ that is top-heavy, but whose equivalence class $E[\gamma]$ is neither locally nor globally top-heavy.

wrapped, it has a consistent outer edge, $\gamma \cong \gamma'$ for any two representatives $\gamma, \gamma' E[\gamma]$, where γ is an elementary SO curve. We know γ is top-heavy by Theorem 6.11, therefore $E[\gamma]$ is globally top-heavy.

It turns out that unlike the general case, with globally top-heavy planar images, we know exactly which free subcurves are direct splits and which are wrap splits, independent of outer basepoint.

Proposition 8.3 (Properties of Globally T-H Planar Images)**.** *Let $E[\gamma]$ be a globally top-heavy planar image with $whit(\gamma) = 1$.*

(i) We have a well-defined direct split map $D : V(\gamma) \to \mathcal{S}$ and wrap split map $Wr : V(\gamma) \to \mathcal{S}$ categorizing the free subcurves of each vertex. That is, for any $C \in E(\gamma)$, we have the direct split $C_v = D(v)$ and the wrap split $C_{v^*} = Wr(v)$.

(ii) The boundary of the exterior face on $G(\gamma)$ is a CCW cycle $\mathscr{B}$.

(iii) $\bigcap_{v \in V}[Wr(v)] = \mathscr{B}$.

Proof. (i) Suppose that we had $C, C' \in E[\gamma]$ so that $C_v \neq C'_v$ for some vertex v. Since we have exactly two options for the direct and wrap splits of a curve with respect to a vertex, we must have $C_{v^*} = C'_v$ and $C_v = C'_{v^*}$. We now observe that $\Gamma = \{C_v, C'_v\} = \{C_v, C_{v^*}\}$ is a free subcurve decomposition of γ. Thus,

$$whit(\gamma) = whit(C_v) + whit(C'_v) \leq 0$$

by Lemma 4.7. Since this contradicts $whit(\gamma) = +1$, we conclude that $C_v = C_{v^*}$.The wrap split map is easily defined once the direct split map is, as the two are complementary.

(ii) Let $v \in V(\gamma)$ be exterior. Then v is adjacent to two exterior edges, e_1, e_2, which lie on the boundary of the exterior face. If both of these edges were directed towards v, then by choosing representatives $C_1, C_2 \in E[\gamma]$ with $C_1(0)$ on e_1 and $C_2(0)$ on e_2, we would have $[C_{1_v}] \neq [C_{2_v}]$, a contradiction by (i). The same argument holds if both exterior edges were outgoing from v; we would get different direct splits of v on C_1, C_2, when chosen in the same fashion. Thus, v must have exactly one incoming exterior edge and one outgoing in $G(\gamma)$. It follows that the boundary is a cycle in one direction, either clockwise or counterclockwise. Now, let us take any representative $C \in E[\gamma]$. We now form a direct split decomposition $\{C_i\}_{i=1}^k \cup \{\tilde{C}\}$. Surprisingly, it is actually easiest to form $\tilde{C}$ first. Let $\tilde{C}$ be the simple subcurve around the outer boundary. Consider $\{w_i\}_{i=1}^k$, the exterior vertices in the order they appear on C. Then we note that the only pieces of the curve remaining after removing $\tilde{C}$ are the direct splits $\{C_{w_i}\}_{i=1}^k$ of each exterior vertex. Hence, $\Omega = \{C_{w_i}\}_{i=1}^k \cup \{\tilde{C}\}$ is a direct split decomposition. By Lemma 4.7, we have $whit(\gamma) = \sum_{i=1}^m whit(C_i) + whit(C)$. Now, if $whit(C) = -1$, we would have $\sum_{i=1}^m whit(C_i) = 2$, which immediately implies $whit(C_i) \geq 1$ for some i, a contradiction. Thus, the boundary cycle must be positive.

(iii) This equality is easily seen by double inclusion. Take an outer basepoint $C(0)$ from a top-heavy representative $C \in E[\gamma]$. We now make a key observation: if $\gamma \in \mathscr{C}$ and v is a vertex, then the direct split γ_v will be the free subcurve $\tilde{\gamma}$ with basepoint v so that $\gamma(0) \notin [\tilde{\gamma}]$. Now, as $whit(Wr(v)) \geq 1$ for all wrap splits, since C is top-heavy, we must have $C(0) \in \bigcap_{v \in V}[Wr(v)]$. Conversely, take any interior point $w \in [\gamma]$. Then take u as the first exterior vertex before w on γ. Then we have $w \in [\gamma_u]$ and hence $w \notin [\gamma_{u^*}]$. So, $w \notin \bigcap_{v \in V}[\gamma_{v^*}]$. This means $\bigcap_{v \in V}[\gamma_{v^*}] \subset \mathscr{B}$. ☐

From this proposition, we can already begin to see the resemblance between globally top-heavy planar images and elementary SO curves. We encourage the reader to note that (i),(ii), and (iii) are all easily seen facts for elementary SO curves. We have one more nice result here – not only is any representation C with $C(0)$ outer top-heavy, but in fact, the converse is also true.

Corollary 8.4. *Let C have $whit(\gamma) = +1$ and $[C] = [\gamma]$, where $E[\gamma]$ is globally top-heavy. Then if C is top-heavy, $C(0)$ is an outer basepoint.*

Proof. This follows immediately from part (iii) of Proposition 8.3 – any wrap split will become a direct split if we do not choose our basepoint to lie on $\mathscr{B}$. ☐

Thus, if we take any representative C so that $[C] = [\gamma]$, with the same orientation as γ, where $E[\gamma]$ globally top-heavy and $whit(\gamma) = +1$, then C is top-heavy iff $C(0)$ is a positive outer basepoint.

We now continue to show similarities between globally top-heavy planar images and elementary SO curves. We will now begin to exploit the poset $P(\gamma)$, on the vertex set $V(\gamma)\backslash\{p_0\}$, with partial order $\subset$. See Section 2.2.

Proposition 8.5 (Properties of Globally T-H Planar Images (Part II))**.** *Let $E[\gamma]$ be a globally top-heavy planar image with $whit(\gamma) = +1$. Let $O \subset V(\gamma)$ be the subset of outer vertices. Then*

(i) For $u \neq v \in O$, we have $u\ S\ v$. (Outer vertices are pairwise separate.)

(ii) Each $u \in O$ is maximal in $P(\gamma)$. (Outer vertices are maximal.)

(iii) Each $u \in V\backslash O$, u is not maximal. (Interior vertices are not maximal.)

(iv) Each $v \in O$ is positive, independent of outer basepoint, and $whit(D(v)) = 0$, where $D : V(\gamma) \to \mathcal{S}$ is the direct split map.

Proof. We can see (i) topologically. Let us regard the boundary $\mathscr{B}$ of $G(\gamma)$ as a Jordan curve, ignoring that the curve is not smooth at the vertices in E. Then by Proposition 8.3, we know that at our first and second visits to each exterior vertex v, we go from outside to inside $\mathscr{B}$ to inside $\mathscr{B}$, then inside to outside, respectively. Let $u \neq v \in O$. We claim that u S v. Let us consider the curve

$C \in E[\gamma]$ with basepoint $C(0)$ on the unique outer edge incoming to u. Let $\{e_1, e_2\}$ be the pair of edges incident to u so that e_1 is incoming to u and outer, and e_2 is outgoing from u and inner. Both of these edges are uniquely defined by (ii) of Proposition 8.3. We proceed by contradiction, in two cases.

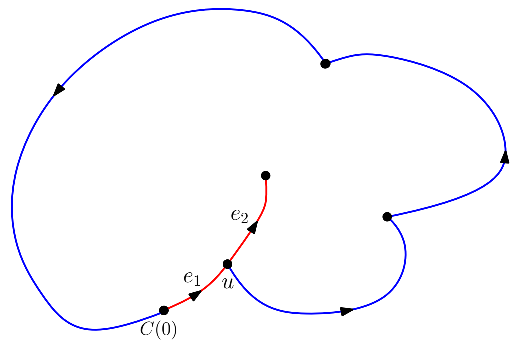


Figure 23: Curve C chosen in the proof of Proposition 8.5 part (i). The boundary $\mathscr{B}$ is shown in blue.

Case 1: Suppose that $u \mathrm{~L~} v$. Then on C, u links v on the left, i.e., the intersection sequence has a subsequence $uvuv$. Now, suppose that no other exterior vertex $w \neq u, v \in O$ lay on the direct split C_u. Then we must directly return from v to u on C_u along the outer boundary $\mathscr{B}$. Additionally, these edges are traversed consecutively by C due to regularity. As a direct split, C_u contains the edge e_2, and we saw that e_1 lies on C_u as well since we return to u from v along the boundary $\mathscr{B}$. But then since the direct split C_u, with basepoint u, contains the pair e_1, e_2, it follows that C_u is regular at u since C is. We now have a contradiction. As mentioned in the preliminaries, direct splits of curves in $\mathscr{C}$ are necessarily irregular at their basepoint.

Otherwise, there is a vertex $w \neq u, v \in O$ which lies on C_u. In this case, all three vertices in $\{u, v, w\}$ are pairwise linked. Thus, the intersection sequence of C has a subsequence $uvwuvw$ with u, v, w exterior. We claim this is impossible. If we consider the curve C' with outer basepoint $C'(0)$ on the outer edge incoming to w, then $[C'_u] \neq [C_u]$, a contradiction to (i) of Proposition 8.3. We conclude that no two exterior vertices may be linked.

Case 2: Suppose $u \subset v$. Again, if v is the last exterior vertex on C_u, we will have both $\{e_1, e_2\}$ lie on C_u, since we must travel along the boundary $\mathscr{B}$ from v to u on C_u. We already argued in Case 1 that this is impossible. Thus, we must have a different exterior vertex $w \neq u, v$ as the last exterior vertex on C_u. Now, we cannot have $w \subset u$, or else, we again have $\{e_1, e_2\}$ on C_u. As w lies on the direct split C_u, we know w is not separate to u. Hence, we must have $w \mathrm{~L~} u$. We now have our final contradiction by Case 1.

(ii) We now address maximality. Let us look at the posets $(D(\gamma), \subset)$ and $P(\gamma)$, where $D(\gamma) = \mathrm{im}(D)$ with D the direct split map. We note that for any curve $\gamma \in E[\gamma]$, the poset of direct splits $(D(\gamma), \subset)$ is isomorphic to the poset $P(\gamma)$, with the obvious order isomorphism between the two being the map which sends each direct split to its basepoint. Since our planar image $E[\gamma]$ is globally top-heavy, every poset $(D(C), \subset)$ is isomorphic to any other $(D(C'), \subset)$ where $C, C' \in E[\gamma]$, by (i) of Proposition 8.3. It follows that $P(C) \cong P(C')$ as well for any two representative curves $C, C' \in E[\gamma]$. Take $v \in E$ and select a curve $C \in E[\gamma]$ so that $C(0)$ lies on the unique outer edge incoming to v. Then since any inner vertex $p \in V(\gamma)$ comes after v on C it follows that v is maximal in the poset $(A \cup \{v\}, \subset)$ where A is the set of inner vertices. By (i), we are done, then, for each element $v \in O$ is maximal in $(O, \subset)$ as the poset is trivial (every pair of elements is not comparable).

(iii) Let choose such any initial representative $\gamma \in E[\gamma]$ and let v be an inner vertex. Choose $t_v \in (0,1)$ so that $\gamma(t_v) = v$, since v is inner. Then, since $\gamma(0)$ is outer, we can set $t_0 = \sup\{t \in [0, t_v] \;:\; \gamma(t) \text{ is outer}\}$ and then $\gamma(t_0) = w$ will be an outer vertex by construction. As v lies on the direct split $D(w)$, we either have $v \text{ L } w$ or $v \subset w$. Suppose the former occurred, or else v is automatically not maximal. Then we must have the direct split $D(v)$ eventually touch the boundary $\mathscr{B}$ after $D(v)$ reaches w, since we currently lie inside the outer boundary and our curve is normal and regular. It follows that there is another exterior vertex $w' \neq w \in [D(v)]$. Since the outer vertices are maximal, we must have $v \text{ L } w'$ as well. But then if we take $C \in E[\gamma]$ so that $C(0)$ lies on the unique outer edge incomding to w', we have $[C_v] \neq [\gamma_v]$.This is a contradiction to (i) of Proposition 8.3. So, we must have $v \subset w$ on $P(\gamma)$. Since all the posets $P(\gamma)$ for any representation $\gamma \in E[\gamma]$ are equivalent, we are done.

(iv) Here, we employ a direct split decomposition. In particular, we look at the decomposition $\Omega = \{\gamma_i\}_{i=1}^k \cup \{C\}$, where γ_i is the direct split of the i^{th} boundary vertex and C is the boundary $\mathscr{B}$ regarded as a simple subcurve. Since the exterior vertices are the unique maximal vertices in $P(\gamma)$ by (ii) and (iii) and the exterior vertices are pairwise separate by (i), we know Ω is a bonafide direct split decomposition. Now, $whit(C) = +1$, by (ii) of Proposition 8.3. Applying Lemma 4.7, we then have $\sum_{i=1}^k whit(\gamma_i) = 0$. If we had $whit(\gamma_j) \leq -1$ for any $j \in [k]$, then we must have some other γ_l with $whit(\gamma_l) \geq 1$, an impossibility by the global top-heaviness of $E[\gamma]$. Thus, $whit(\gamma_i) = 0$ exactly for each of these direct splits.

Choose any exterior vertex v. The positivity of v is seen through the positive boundary cycle. Let e_1 be the edge after the (unique) exterior edge incoming to v. Note that such an edge naturally has a pair by the normality of the curve γ. Then take e_2 to be the unique exterior edge outgoing from v. We can use these edges to approximate the tangent vectors $\gamma'(t_v)$ and $\gamma'(t_{v^*})$ where $t_v < t_{v^*} \in [0,1]$ are the times so that $\gamma(t_v) = \gamma(t_{v^*}) = v$. From (ii) of Proposition 8.3 , we know the point $\gamma(t_v + \epsilon)$ is interior and hence must lie on e_1. Similarly, $\gamma(t_{v^*} + \epsilon)$ must lie on e_2 for ϵ small. As e_1 turns clockwise to e_2, v is a positive vertex. The signs of all vertices are independent of outer basepoint by (i) of Proposition 8.3. □

Corollary 8.6. *Take γ with $whit(\gamma) = +1$ and $E[\gamma]$ globally top-heavy. Suppose $\gamma(0)$ is outer so that γ is top-heavy. Then if $w_1 \to \cdots \to w_k \to w_1$ is the boundary cycle on $\mathscr{B}$, then $\{w_i\}_{i=1}^k$ are the k unique maximal elements in $P(\gamma)$. Moreover, we have*

$$\gamma = Wr_+(D(v_i)) \# \cdots \# Wr_+(D(v_k))$$

where # is the connected sum operator.

Proof. All of the key ingredients for this proof lie in the proofs of Propositions 8.3 and 8.5. Let us first consider the decomposition Ω of γ by $\Omega = \{D(w_i)\}_{i=1}^k \cup \{\alpha\}$ where α is the outer boundary $\mathscr{B}$ regarded as a closed subcurve of γ. Then by the proof of (iii) from Proposition 8.5, we know that each inner vertex v has $v \subset w_i$ for some unique w_i. It follows that for $i \neq j$, we have $D(w_i) \cap D(w_j) = \varnothing$, otherwise we would have some inner vertex u with $u \text{ L } w_i$ and $u \text{ L } w_j$, which we argued cannot happen while proving (iii) in Proposition 8.5. Hence, we can take an ϵ-neighborhood around the outer boundary of each $D(w_i)$, which contains no extra intersection points of γ besides the ones on $D(w_i)$. Then, since w_i is an outer point on γ, we can find a simple path in this ϵ-neighborhood from some point $p_i = \gamma(t_i - \varepsilon)$ to another point $q_i = \gamma(t_i^* + \varepsilon)$ where again $0 < t_i < t_i^* < 1$ are the pre-images of w_i. By construction, the right cut of each of the paths P_i from p_i to q_i will be combinatorially equivalent to the curve $Wr_+(D(w_i)$. See Definition 5.1 and Figure 24. □

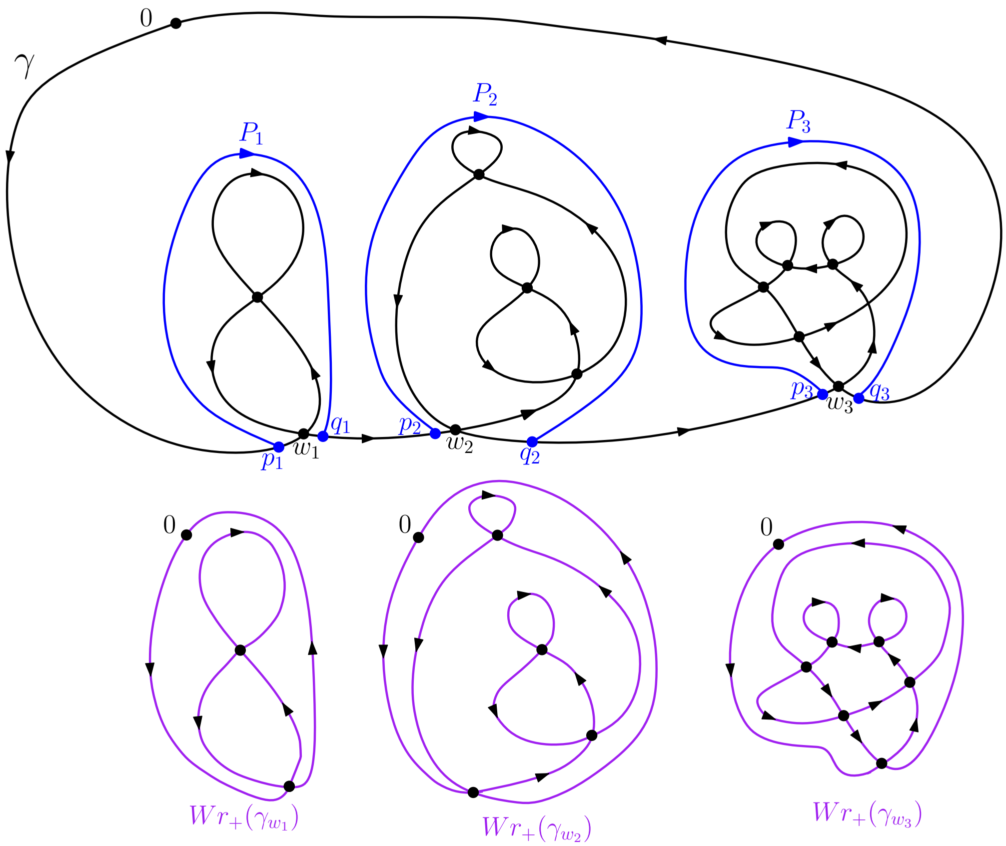


Figure 24: A globally top-heavy planar image along with the cuts P_1, P_2, P_3 from the proof of Corollary 8.6. It is geometrically clear here that $\gamma \cong Wr_+(\gamma_{w_1})\ \#Wr_+(\gamma_{w_2})\#Wr_+(\gamma_{w_3})$.

This corollary shows us the striking similarities between globally top heavy planar images $E[\gamma]$ and our elementary SO curves, for which this statement is also true. With this in mind, one can see that globally top-heavy planar images are a natural generalization of elementary SO curves.

We can now generalize irreducibility to planar images in the same fashion that we did with top-heaviness.

Definition 8.7. *We call a planar image $E[\gamma]$* ***locally irreducible*** *iff there exists a representative curve $C \in E[\gamma]$ that is irreducible. We then call a planar image $E[\gamma]$* ***globally irreducible*** *iff any representative $C \in E[\gamma]$ is irreducible.*

8.3 The Main Theorems

We proceed towards proving locally irreducible planar images $E[\gamma]$ with $whit(\gamma) = +1$ are SO. It turns out that attacking this result directly is the wrong way to go about it. Instead, as we will see, the result, Theorem 8.9, follows quite nicely from Theorem 8.8, a result that is seemingly about minimum homotopy area.

We now introduce a new combinatorial decomposition of a curve called a *simple path decomposition*. Let $\gamma \in \mathscr{C}$. Suppose we trace out the path of $[\gamma]$ from time $t = 0$ until γ first self-intersects. It's not hard to see that the point we touch will be the basepoint of a simple loop. More precisely, suppose that $t_1 < t_1^* \in [0,1]$ are the smallest values of t so that $\gamma_1 = \gamma_{[t_1, t_1^*]}$ is a loop. By Lemma 4.2, we know such a loop is outwards. Now, any path along the orientation of the curve, between two vertices, can be represented by subsection of the intersection sequence. Moreover, a vertex v_i labeled i in the intersection sequence will be a vertex of the path iff i appears in the (sub)intersection sequence exactly twice. It follows that the sub-intersection sequence from $r = \gamma(0)$ until $p_1 = \gamma(t_1) = \gamma(t_1^*)$ is of the following form: $r \cdots p_1 \cdots p_1$, where no intersection point occurs twice before the second occurrence of p_1. Geometrically, this means the first portion of γ, from r to p_1, is necessarily simple. Let us call P_1 the path from r to p_1. Then $\gamma_{[t_1, t_1^*]} = P_1 * \gamma_1$. We can naturally continue this decomposition. If we continue tracing the next piece of the curve, looking at $\gamma_{[t_1^*, t]}$ then either we find $\gamma_{[t_1^*, t]}$ is non-injective for some $t < 1$ or the whole path $\gamma_{[t_1^*, 1]}$ is simple and we return to r. In the former case, let us continue by setting $t_2^* = \sup\{t \in [0,1] \mid t > t_1^*, \gamma_{[t_1^*, t]} \text{ is injective }\}$. By the same logic, the point $\gamma(t_2^*) = p_2$ must be the basepoint of an outwards loop, for we have some t_2 with $t_1^* < t_2 < t_2^* < 1$ so that $\gamma(t_2) = \gamma(t_2^*)$, by construction of t_2^* and hence $\gamma_{[t_2, t_2^*]}$ is also a loop. Continuing in this fashion, we will achieve a decomposition of γ into alternating simple paths and loops. Let us write $\gamma = P_1 * \gamma_1 * \cdots * P_k * \gamma_k * P_{k+1}$ where $*$ is meant to indicate path concatenation and each P_i is a simple path, each γ_i is an outwards loop. We call such a decomposition a *simple path decomposition* of the curve γ. Following the notation from earlier, we would have $P_i = \gamma_{[t_{i-1}^*, t_i]}$ and $\gamma_i = \gamma_{[t_i, t_i^*]}$ with $t_0^* = 0$. See Figure 25 for an example.

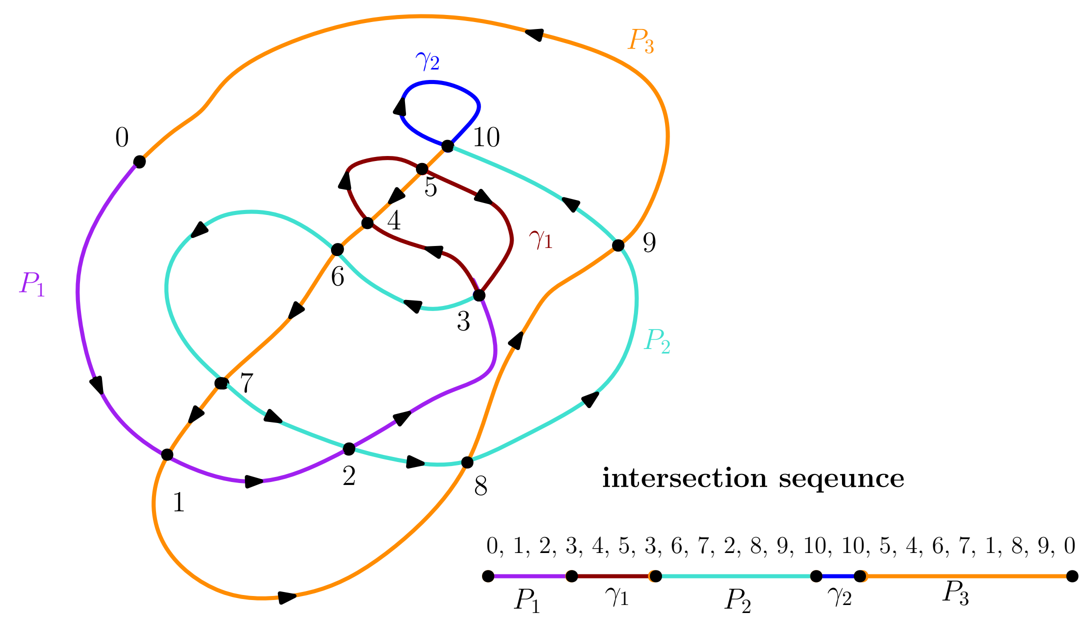


Figure 25: A top heavy plane curve with its simple path decomposition shown. The intersection sequence is also shown to the side, with the intersection sequence broken into the the simple path decomposition as well.

We now employ our simple path decompositions to prove our first main theorem.

Theorem 8.8. *Let $\gamma \in \mathscr{C}$ have positive outer basepoint.Then there is a positive integer k so that $Wr_+^k(\gamma)$ is a positive interior boundary.*

Proof. Then let l be the number of negative vertices in $V(\gamma)$. Set $k = l+1$. We claim that $Wr_+^k(\gamma)$ is an interior boundary. We will show this by iteratively building a left sense-preserving nullhomotopy H for γ. It then follows that $\sigma(Wr_+^k(\gamma)) = W(Wr_+^k(\gamma)$ by Lemma 3.6. Then the reversal $\overline{H}$ of this homotopy is right sense-preserving. Such a homotopy is winding number increasing. That is, if $wn_i(x)$ is the winding number of $x \in \mathbb{R}^2$ on $\overline{H}(i,\cdot)$, then $wn_i(x)$ is monotonically increasing. It follows that γ, the target curve of $\overline{H}$ is positive consistent. Thus, $Wr_+^k(\gamma)$ is a positive interior boundary by part (3) of Theorem 3.5.

We now present a trick called *balanced loop deletion*, which will be instrumental in our construction of the left sense-preserving nullhomotopy of $Wr_+^k(\gamma)$. Suppose that C is a curve that is positively wrapped, $C = Wr_+(C')$ for some curve C' and also that the first loop γ_- in the simple path decomposition of C is negative. Let b be the basepoint of γ_-, the first loop in the simple path decomposition of C. The act of balanced loop deletion will be performed by a left sense-preserving homotopy H so that $C \overset{H}{\to} C\backslash\{\alpha,\gamma_-\}$ where α is the positive outer wrap on C. See Figure 26, where all of the following objects are shown, in the context of balanced loop deletion. Let P be the simple path along the orientation of C from the positive outer point $C(0)$ to b. We are interested in the subpath P' from a to b, where a is the unique outer intersection point on $[C]$, the basepoint of the wrap α. Let us take a right ϵ copy P'' of P'. Then, for ϵ sufficiently small, we can make an arbitrarily small deformation to P'' to transform it into a path P'_ϵ between points a' and b' which are on C and are in a right ϵ neighborhood of a and b, respectively. Let us take $t_a < t_a^*$ and $t_b < t_b^*$ so that $C(t_a) = C(t_a^*) = a$ and $C(t_b) = C(t_b^*) = b$. Then, since the direct split of a is a negative loop and the wrap split of b is a positive loop, we will have $a' = C(t_a^* + \varepsilon_1)$ and $b' = C(t_b^* - \varepsilon_2)$ for some $\varepsilon_1, \varepsilon_2$ small. We now consider the path $\widetilde{P}$ between a' and b' along the orientation of C. This path consists of the concatenation of the following paths: a simple path from a' around the wrap back to a, the path P' from a to b, and the simple path along the direct split of b from b to b'. These paths are shown in purple, blue, and red, respectively in Figure 26. Now, not only are each of these subpaths simple, but none of them intersect each other since b is the first self-intersection point of the curve. Thus, $\widetilde{P}$ is a simple path. We now make a crucial observation: $\widetilde{P} * P'_\epsilon$ is a simple, positive, closed curve. It follows that we can perform a Blank cut along C, which will replace the path $\widetilde{P}$ on C with the path P'_ϵ. The effect of this cut on curve is that both the outer wrap α and the negative loop γ_- are deleted, and the path P' is replaced by its right ϵ copy P'_ϵ. For ϵ sufficiently small, the trade of P' for P'_ϵ will not affect the portion of the signed intersection sequence along which P' corresponds. Thus, considering only the intersection sequence, the cut will delete both the direct split $C_b = \gamma_-$ and the wrap $C_{a^*} = \alpha$, and leave everything else unaffected. This Blank cut can be performed by a left sense-preserving homotopy, so we have established the existence of left sense-preserving balanced loop deletion.

With balanced loop deletion in hand, we are ready to form a left sense-preserving nullhomotopy H of $Wr_+^k(\gamma)$. We can iteratively build H as the concatenation of many easily defined left sense-preserving subhomotopies, so $H = \sum_i H_i$. We proceed inductively as follows. Suppose $H_1, \ldots, H_{i-1}$ have been defined and γ_i is the current curve. Then if the first loop in the simple path decomposition of γ_i is positive, we let H_i be a left sense-preserving nullhomotopy of this loop. Otherwise, the loop

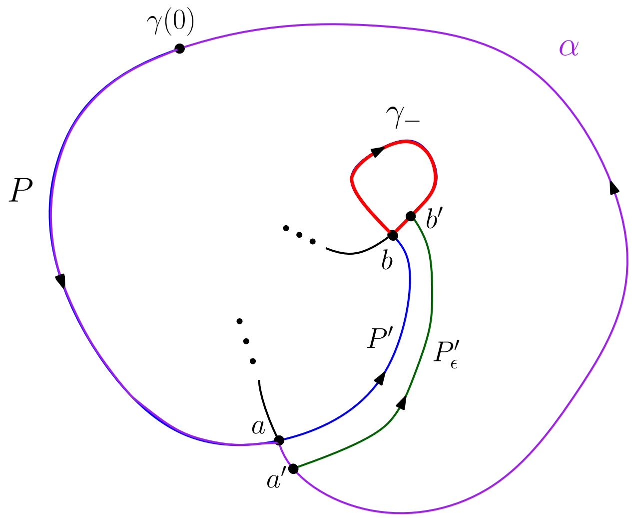


Figure 26: The combinatorial structure necessary to apply balanced loop deletion - a wrapped curve, with outer wrap α and a negative loop γ_- as first loop in the simple path decomposition.

is negative and we let H_i be the homotopy performing balanced loop deletion. We claim that we will always have a wrap available, so we can perform balanced loop deletion. We show this now. Since each of the homotopies H_j is equivalent to the action of deleting a free subcurve on γ_j for $j \in [i-1]$, the intersection points left over will not have their signs affected. Indeed, all the vertices created by wrapping will remain positive, and since our basepoint will still lie on a wrap, the signs of the vertices originally on γ will be preserved as well. Thus, we note that the number of negative vertices n_i on any curve γ_i is bounded: $n_i \leq l$. Also, if a vertex v is the basepoint of an outwards loop γ_v, then the sign $\operatorname{sgn}(v)$ is naturally given by the orientation of the loop: $\operatorname{sgn}(v) = +1$ iff γ_v is counterclockwise oriented and $\operatorname{sgn}(v) = -1$ iff γ_v is clockwise oriented. It follows that we can have at most l distinct integers s so that γ_s has its first loop as a negative loop, since the first loop is always outwards. We wrapped γ $l+1$ times, so we will always have a wrap available.

Now, since our algorithm will never get stuck, and $|\gamma_{i+1}| < |\gamma_i|$, we must eventually reach a point when the current curve γ_m has $|\gamma_m| = 0$. We now show that this final curve γ_m is a positive Jordan curve. First, let us note that the basepoint of the first loop on $\gamma_0 = Wr_+^k(\gamma)$ is a vertex from γ, unless γ is simple. In the case that γ is simple, it must be a positive Jordan curve since it has a positive outer basepoint. Thus, $Wr_+^j(\gamma)$ is a basic $(j+1)$-boundary for any $j \in \mathbb{Z}_+$. Thus, we need not consider this case, so we assume γ is non-simple. In the case that γ is non-simple, eventually the basepoint of the first loop of some γ_j will not come from γ. This holds as we have at least one extra vertex w on γ_0 created by wrapping, and $v \subset w$ for all vertices $v \in V(\gamma)$, so there is no way either possible step, balanced loop deletion or positive loop contraction, will delete w. Suppose that γ_{n_0} is the first curve in the sequence to have the basepoint of its first loop not be a vertex from γ. We now make an observation: each γ_j has its intersection sequence of the form $w_1 w_2 \cdots w_n I w_n \cdots w_2 w_1$ where w_i are the remaining basepoints of the wraps and I is the subsequence of the remaining portion of γ. The only way that the basepoint of the first loop of γ_{n_0} is not a vertex from γ is if no vertices from γ are vertices on γ_{n_0}. It follows that the only vertices remaining are the vertices of the wraps. Thus, our algorithm will employ only positive loop contraction $|\gamma_{n_0}|$ times and then our final curve γ_m will indeed be a positive Jordan curve. Thus, we can perform a left sense-preserving nullhomotopy H_m contracting it to its basepoint. It follows that $H = \sum_{i=1}^{m} H_i$ is our desired left sense-preserving nullhomotopy. □

Our main theorem now falls out quite nicely:

Theorem 8.9 (Locally Irreducible Planar Images are SO)**.** *Let γ have $whit(\gamma) = +1$ with positive outer basepoint. If γ is irreducible, then it is SO.*

Proof. Let us wrap around γ so that $Wr_+^k(\gamma)$ is a positive interior boundary. By Theorem 3.5 part (5), we have a SO decomposition Ω of $Wr_+^k(\gamma)$ into only positive SO subcurves. Additionally, since this is a free subcurve decomposition, by Corollary 2.6, we know Ω has a SO direct split of $Wr_+^k(\gamma)$. Let w_i be the vertex created by the i^{th} wrap. Thus, when traversing $Wr_+^k(\gamma)$, we will see $w_k w_{k-1} \cdots w_1$ as the first substring of the intersection sequence. Then the direct split of w_i on $Wr_+^k(\gamma)$ has $whit(Wr_+^k(\gamma)_{w_i}) = 1 + (i-1) = i$. This follows from Lemma 4.7, since we can construct a free subcurve decomposition of $whit(Wr_+^k(\gamma)_{w_i})$ into the direct split corresponding to γ, along with $(i-1)$ wraps. Thus, each $Wr_+^k(\gamma)_{w_i}$ is not a SO direct split, for $i \in \{2, 3, \dots, k\}$ since SO curves C have $whit(C) = +1$. Note that by our notation w_1 is the vertex corresponding to the original basepoint $\gamma(0)$. Now, no direct split on $Wr_+^k(\gamma)$ which was also a direct split on γ can be SO, since γ is irreducible. The only direct split we have not excluded is $Wr_+^k(\gamma)_{w_1}$. Consequently, it must be SO. As $Wr_+^k(\gamma)_{w_1} \cong \gamma$, by Lemma 3.1 we conclude that γ is SO. □

Corollary 8.10 (Locally Top-Heavy Planar Images are SO)**.** *Let γ have $whit(\gamma) = +1$ and positive outer basepoint. If γ is top-heavy, then it is SO.*

Proof. Any top-heavy curve is irreducible, since SO curves C have $whit(C) = +1$ and we have no direct splits γ_v with $whit(\gamma_v) = +1$. □

Corollary 8.11. *Let γ have $whit(\gamma) = +1$ and positive outer basepoint. Then if γ is not SO, it has a self-overlapping direct split γ_v.*

We originally conjectured that curves γ with positive outer basepoint and $whit(\gamma) = +1$ are top-heavy iff they are irreducible. It turns out this is not the case, as we now show. See Figure 27.

Let us now view Theorem 8.8 in full generality.

Theorem 8.12 (Wrapping Resolves Badness)**.** *Let γ be a plane curve with outer basepoint and set $n = |\gamma|$. Then there are constants $k_-, k_+ \leq n+1$ so that $Wr_-^{k_-}(\gamma)$ and $Wr_+^{k_+}(\gamma)$ are negative and positive interior boundaries, respectively.*

Proof. If $\gamma(0)$ is a positive basepoint, then k_+ exists directly by Theorem 8.8. To prove the existence of k_-, consider the intermediary curve γ' obtained from γ by performing a I_b move on the outer edge to create a negative loop that is entirely outer to the curve. Additionally, let us set the basepoint on γ' on the new negative loop γ_-. Then γ' has a negative outer basepoint. Here is the key observation: $Wr_+(\overline{\gamma'})$ has the same signed intersection sequence as $Wr_-(\gamma)$. Thus, by the existence of $\tilde{k}_+$ for $\overline{\gamma'}$, we have k_- for γ.

If $\gamma(0)$ is negative, then since k_- and k_+ exist for the reversal $\overline{\gamma}$, we are done, for $Wr_-^k(\gamma)$ has the same signed intersection sequence as $Wr_+^k(\overline{\gamma})$. □

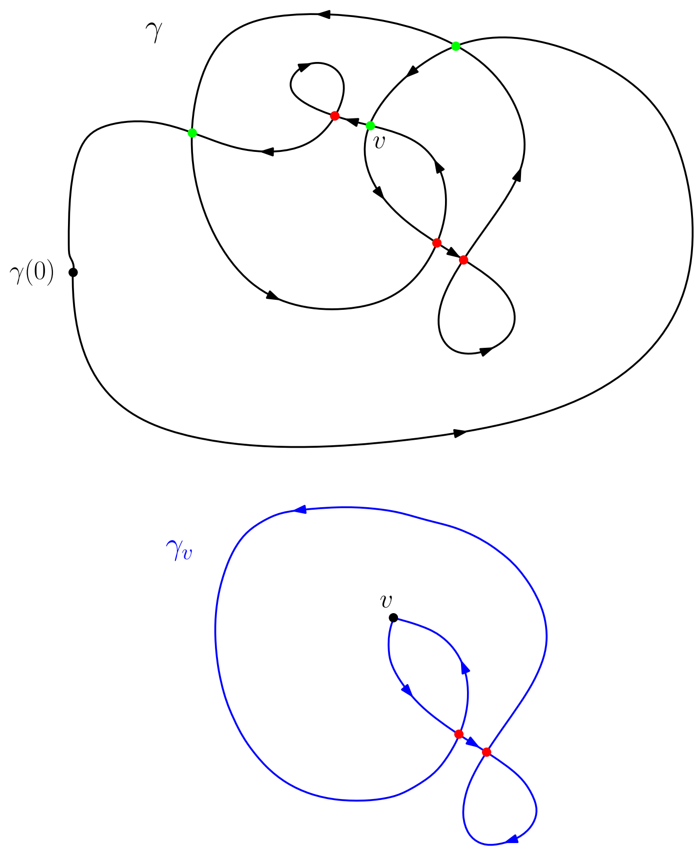


Figure 27: A locally irreducible curve γ with positive outer basepoint and whit$(\gamma) = +1$ that is not top-heavy. The perpetrator here is the direct split γ_v which has whit$(\gamma_v) = +1$ but is not SO. Note that γ_v is locally top-heavy, but that its basepoint is not outer.

Let us make a simple observation: once $Wr^k_\pm(\gamma)$ is an interior boundary, so too is $Wr^j_\pm(\gamma)$ for any integer $j \geq k$. This holds because we can simply add the extra wraps to our old SO decomposition, since they are positive/negative Jordan curves and hence are automatically positive/negative SO. Consequently, we can interpret Theorem 8.12 as saying interior boundaries are the equilibrium point for plane curves with respect to the action of wrapping. No matter where we begin, we will always eventually land and stick within the set of interior boundaries.

Consider the value $B(\gamma) = W(\gamma) - \sigma(\gamma) \geq 0$. We call $B(\gamma)$ the *badness* of γ. One can interpret $B(\gamma)$ as a measure of how pathological γ is with respect to minimum homotopy area. An alternate interpretation of Theorem 8.12, which inspired the name, is that wrapping very quickly removes all badness of a curve. Thus, we can think of wrapping as a powerful and efficient way to rectify a curve, with respect to minimum homotopy area.

8.4 Balanced Loop Insertion

We now introduce an operation called *balanced loop insertion*, which is complementary to the balanced loop deletion applied in the proof of Theorem 8.8. We will show that any curve γ with positive outer basepoint and $whit(\gamma) = +1$ can be transformed into a SO curve, more specifically, a top-heavy SO curve, through a sequence of balanced loop insertions. This result is a nice parallel to Theorem 8.8.

Definition 8.13 (Balanced Loop Insertion)**.** *Let $\gamma \in \mathscr{C}$ have positive outer basepoint p_0. Then given any edge e from $G(\gamma)$, we define balanced loop insertion on γ with respect to e as follows. First, perform a I_b move to insert a negative loop, on the right side of the edge e, and smooth the resulting curve γ' until is normal and regular. Then, wrap around γ' to create $\gamma'' = Wr_+(\gamma')$. We call γ'' the curve obtained through balanced loop insertion on γ with respect to e. See Figure 28.*

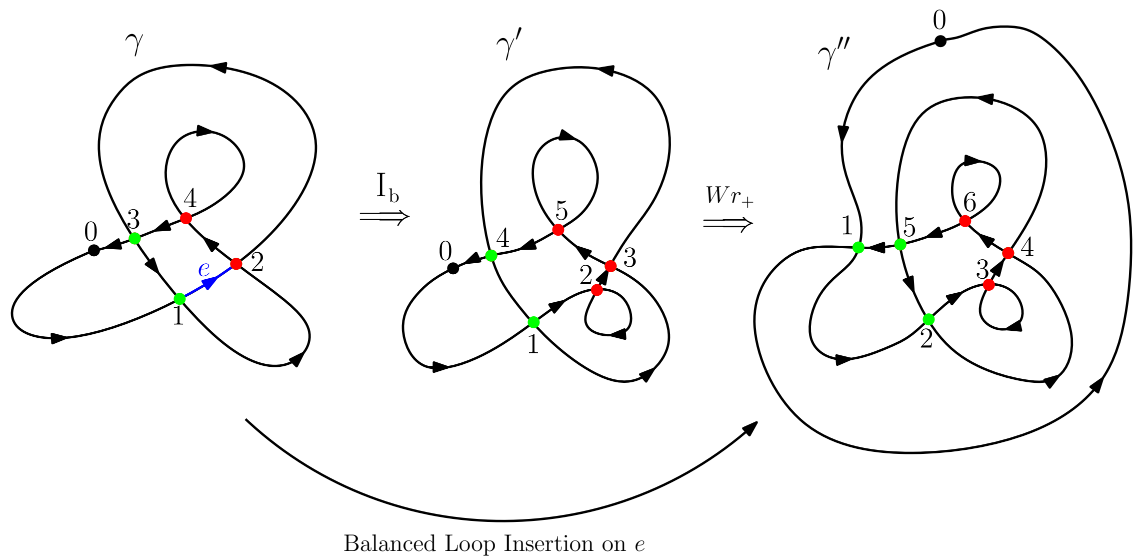


Figure 28: Balanced loop insertion on a SO curve γ with respect to an edge e. Note that the curve produced, γ'', is also SO.

We are particularly interested in applying balanced loop insertion to curves with $whit(\gamma) = +1$. In fact, we now define an iterative process of balanced loop insertion which will transform any such curve, with positive outer basepoint, into a top-heavy, SO curve. First, let us note an interesting property of balanced loop insertion.

Proposition 8.14. *Let γ have $whit(\gamma) = +1$ and positive outer basepoint. If γ is top-heavy, and γ' is obtained from γ by balanced loop insertion, then γ is top-heavy as well.*

Proof. Let $i : V(\gamma) \to V(\gamma')$ be the inclusion map, sending vertices on γ to the corresponding vertices on γ'. Then the only possible change to the direct split was that we added a negative loop, so

$$whit(\gamma'_{i(v)}) \leq whit(\gamma_v) \leq 0$$

The only other direct splits we need to check are those of the vertices u, v created by balanced loop insertion. We note immediately that $whit(\gamma'_u) = -1$ where u is the basepoint of the new negative loop. Meanwhile, if v is the vertex created by the wrap, then $whit(\gamma'_v) = whit(\gamma) - 1 = 0$. Hence, γ' is top-heavy. □

While top-heaviness is preserved by balanced loop insertion, and consequently, so too is SO-ness, if we have positive outer basepoint and $whit(C) = +1$, we need a stronger operation to transform a non-SO curve into a SO curve. We define this operation now.

Definition 8.15 (Global Balanced Loop Insertion)**.** *Given a curve $\gamma \in \mathscr{C}$, there are $2|\gamma| + 1$ edges on $G(\gamma)$, seen by the intersection sequence. Let us apply balanced loop insertion all at once $2|\gamma| + 1$ times on every edge of γ. Equivalently, we perform a I_b move to the right of every edge, adding a new negative loop, then wrap the curve $2|\gamma| + 1$ times. Suppose this yields a new curve γ'. We write M for the operator performing global balanced loop insertion, so that $M(\gamma) = \gamma'$.*

We now claim that this operation transforms any curve C with positive outer basepoint and $whit(C) = +1$ into a top-heavy curve and hence a SO curve by Corollary 8.10.

Theorem 8.16. *Let γ have positive outer basepoint and $whit(\gamma) = +1$. Then $M(\gamma)$ is top-heavy and SO.*

Proof. The only fact we will need here is a formula of Titus and Whitney [21, 23]. Let C be a curve with outer basepoint. Then

$$whit(C) = \sum_{v \in V(C) \setminus \{0\}} \operatorname{sgn}(v) + a,$$

where $a = +1$ if the basepoint is positive and $a = -1$ if the basepoint is negative. It follows immediately from this fact that for any curve C, we have $whit(C) \leq |C| + 1$. Thus, for any direct split γ_v on our original curve, we have $whit(\gamma_v) \leq |\gamma_v| + 1$. On the other hand, since every edge of γ_v received at least one negative loop, as edges of γ_v may be further subdivided on γ, we note that we inserted at least $2|\gamma_v| + 1$ negative loops on the direct split γ_v. Thus, if we write $i : V(\gamma) \to V(M(\gamma))$ for the natural inclusion map and set $\tilde{v} = i(v)$, then we see $whit(M(\gamma)_{\tilde{v}}) \leq -|\gamma|_v \leq 0$. Hence, all the vertices on $M(\gamma)$ that came from γ will not yield direct splits of Whitney index 1 or greater. Now, the only other vertices are the basepoints of the new negative loops and the basepoints of the wraps. Clearly, for any vertex u of the former kind, we have $whit(M(\gamma)_u) = -1$, so these are no problem. The basepoints of the wraps are no problem either. Set w_i for the i^{th} vertex on $M(\gamma)$, or, equivalently, the vertex created by wrap number $2|\gamma| + 1 - i$. Then we have $whit(M(\gamma)_{w_i^*}) = i$, for $i \in \{0, 1, \ldots, 2|\gamma| + 1\}$. Since direct splits and wrap splits are complementary, and $whit(C) = whit(C_v) + whit(C_{v^*})$ for any curve C and vertex $v \in V(C)$, we have

$$whit(M(\gamma)_{w_i}) = 1 - i \leq 0$$

for any $i \in \{0, \ldots, 2|\gamma| + 1\}$. By definition, $M(\gamma)$ is top-heavy, so we are done. $\square$

See Figure 29 for an example of global balanced loop insertion, which transforms a non-SO curve into a SO curve. In a sense, the curves $M(\gamma)$ produced by global balanced loop insertion are paragons of top-heavy curves. All of the negative subcurves lie at the 'back' or 'top' of the curve, stuffed inside the direct split corresponding to the original curve.

With global balanced loop insertion, we can prove our final surprising fact about SO curves.

Theorem 8.17 (Flexibility of Top-Heavy SO curves)**.** *Let C be a curve with $whit(C) = +1$ and positive outer basepoint. Then there is a top-heavy SO curve γ so that $G(C)$ is an induced subgraph of $G(\gamma)$.*

Proof. We simply set $\gamma = M(C)$. Then if $i : V(C) \hookrightarrow V(\gamma)$ is the natural inclusion map, we note that the induced subgraph H of $G(\gamma)$ with vertex set $\operatorname{im}(i)$ has H isomorphic to $G(C)$. $\square$

Hence, top-heavy SO curves are an extremely non-trivial class of SO curves. Loosely speaking, we can find any combinatorial structure embedded inside of some top-heavy SO curve.

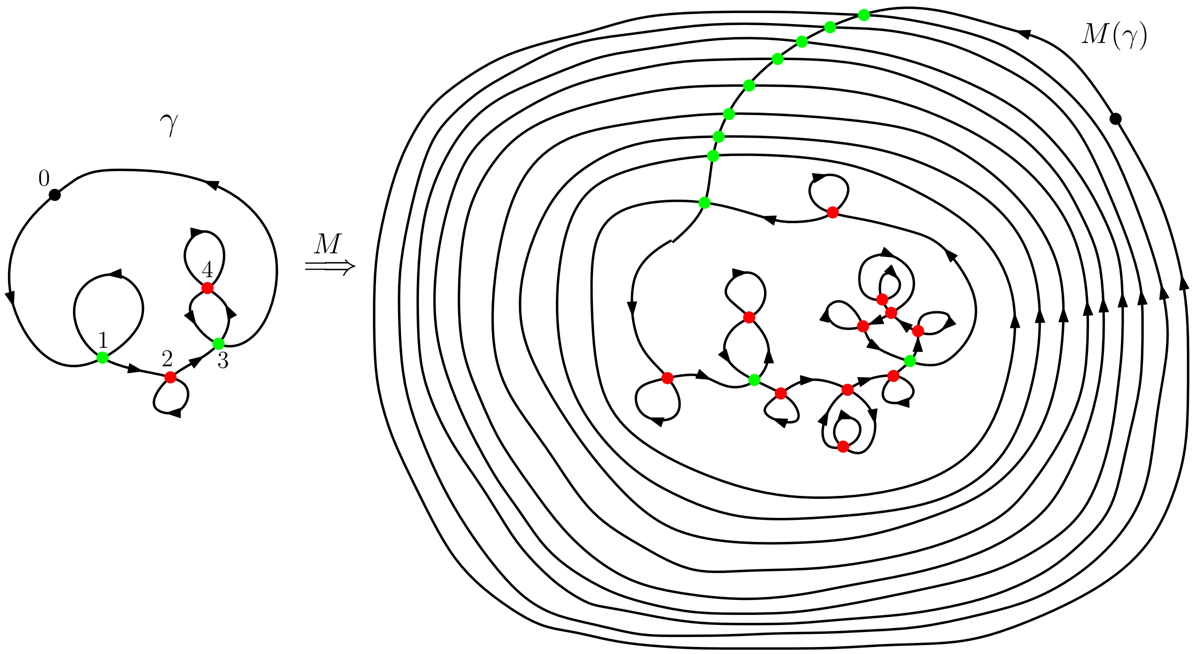


Figure 29: Global balanced loop insertion applied to a non-SO curve γ, seen by its empty positive loop.

9 Conclusion

In this thesis, we studied self-overlapping curves, employing interior boundaries and minimum area homotopies as tools for our investigation. The duality between self-overlapping curves and minimum area homotopies was exploited, to show that not only can self-overlapping curves inform us about minimum area homotopies, but so too can minimum area homotopies reveal to us surprising properties of self-overlapping curves. These connections are especially evident through Theorem 8.8, Wrapping Resolves Badness, which told us that after only $|\gamma| + 1$ wraps, any curve with outer basepoint will become an interior boundary. All of our main results on self-overlapping curves followed swiftly from this theorem.

We showed that an elementary curve γ with $whit(\gamma) = +1$ and positive outer basepoint is SO iff it is top-heavy in Theorem 6.11. Additionally, in Lemma 6.4 we formed a bijection between trees with green and red nodes and elementary SO curves. We then showed more generally with Corollary 8.10 that top-heaviness, plus unit Whitney index and positive outer basepoint are sufficient, in general, to be self-overlapping. In fact, this result was a corollary to Theorem 8.9, which said that a curve γ with $whit(\gamma) = +1$, positive outer basepoint and no SO direct splits is SO. The connections between SO curves, interior boundaries, and minimum area homotopies was particularly on display by how easily Theorem 8.9 followed from Theorem 8.8, two seemingly unconnected results.

Our two other main results were about transforming plane curves into interior boundaries and top-heavy SO curves. We showed that any curve C with outer basepoint can be transformed into a negative interior boundary and a positive interior boundary by less than $|C| + 1$ wraps. Similarly, we showed that with $2|C| + 1$ applications of balanced loop insertion, we can turn a curve with $whit(C) = +1$ and positive outer basepoint into a top-heavy SO curve.

In future work, we hope to extend the approach of studying direct splits to interior boundaries and additionally to further investigate the relationship between plane curves γ and their set of elementary forms $\{\gamma_\Omega\}_\Omega$.

10 References

[1] Helmut Alt. The computational geometry of comparing shapes. Lecture Notes in Computer Science 5760. Springer, 2009.

[2] Chanon Aphirukmatakun. An approach to polynomial curve comparison in geometric object database. *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, pages 386–392, 2007.

[3] Vladimir Arnold. Plane curves, their invariants, perestroikas and classifications. *Steklov Mathematical Institute, Moscow*, 1993.

[4] Samuel J. Blank. *Extending Immersions of the Circle*. PhD thesis, Brandeis University, 1967.

[5] Erin Chambers and David Letscher. On the height of a homotopy. In *21st Canadian Conference on Computational Geoemetry*, pages 103–106. 2009.

[6] Erin Chambers and Yusu Wang. Measuring similarity between curves on 2-manifolds via homotopy area. In *Proc. Proc. 29th Symposium on Computational Geometry*, pages 425–434, 2013.

[7] Hsien-Chih Chang and Jeff Erickson. Untangling planar curves. In *Proc. 32nd International Symposium on Computational Geometry*, 2016.

[8] David Eppstein and Elena Mumford. Self-overlapping curves revisited. *SODA '09 Proceedings of the twentieth annual ACM-SIAM Symposium on Discrete Algorithms*, 2009.

[9] Jeff Erickson. Generic and regular curves. University of Illinois at Urbana Champaign, CS 589 Course Notes, 2013.

[10] Brittany Terese Fasy, Selcuk Karakoc, and Carola Wenk. On minimum area homotopies of normal curves in the plane. aXiv:1707.02251, 2017.

[11] Jack Graver and Gerald Cargo. When does a curve bound a distorted disk? *SIAM Journal on Discrete Mathematics*, 25(1):280–305, 2011.

[12] Joel Hass and Peter Scott. Shortening curves on surfaces. *Topology*, 33(1):25–43, 1994.

[13] Selcuk Karakoc. *On Minimum Homotopy Areas*. PhD thesis, Tulane University, 2017.

[14] Morris Marx. Extending immersions of $\mathbb{S}^1$ to $\mathbb{R}^2$. *Transactions of the American Mathematical Society*, 187:309–326, 1974.

[15] Zipie Nie. On the minimum area of nullhomotopies of curves traced twice. arXiv:1412.0101v2, 2014.

[16] Tahl Nowik. Complexity of planar and spherical curves. *Duke Math. J.*, 148(1):107–118, 2009.

[17] Herbert Seifert. *Konstruktion dreidimensionaler geschlossener Raume*. PhD thesis, Saxon Academy of Sciences Leipzig, 1931.

[18] Peter W. Shor and Christhoper Van Wyk. Detecting and decomposing self-overlapping curves. *Computational Geometry: Theory and Applications*, 2(1):31–50, 1992.

[19] Ernst Steinitz. Polyeder und Raumeinteilungen. *Enzyklopädie der mathematischen Wis- senschaften mit Einschluss ihrer Anwendungen*, 1916.

[20] Ernst Steinitz and Hans Rademacher. *Vorlesungen über die Theorie der Polyeder: unter Einschluß der Elemente der Topologie*. Springer-Verlag, 1934.

[21] Charles Titus. A theory of normal curves and some applications. *Pacific J. Math*, 10(3):1083–1096, 1960.

[22] Charles Titus. The combinatorial topology of analytic functions of the boundary of a disk. *Acta Mathematica*, 106(1):45–64, 1961.

[23] Hassler Whitney. On regular closed curves in the plane. *Compositio Math. 4*, pages 276–284, 1937.

A Appendix

Let $\gamma \in \mathscr{C}$ be a curve with (unsigned) intersection sequence $1\,2\,3 \cdots n\,1\,2\,3 \cdots n$. Then we call γ a *flower*. The following argument is technical, but the lemma is needed to show that our G-R trees are trees.

Lemma A.1. *Let $\gamma \in \mathscr{C}$ be a flower. Then $|\gamma|$ is odd.*

Proof. Suppose we had an even flower γ with intersection points $\{p_i\}_{i=1}^n$. We write $t_i < t_i^*$ as the two values of $t \in \gamma^{-1}(p_i)$ and we notate the path $P_i^{(1)} = \gamma|_{[t_i, t_i^*]}$ and $P_i^{(2)} = \gamma|_{[t_i^*, t_i]}$ as the first and second direct paths, respectively, from P_i to P_{i+1}. Note that both these paths are simple. Since the path $P = \gamma_{p_1} = P_1^{(1)} * P_2^{(1)} * \cdots * P_{n-1}^{(1)} * P_n^{(1)}$, where $*$ is meant to indicate concatenation, is simple and closed, by the Jordan Curve Theorem, it has an interior $\mathrm{int}(P)$ and an exterior $\mathrm{ext}(P)$. Now, the key claim is that at t_i^*, we have $\gamma(t)$ switch from inside P to outside P. In other words, we claim there is some $\epsilon_0 > 0$ so that $\gamma(t_i^* - \epsilon) \in \mathrm{int}(P)$ and $\gamma(t_i^* + \epsilon) \in \mathrm{ext}(P)$ for all $\epsilon < \epsilon_0$. Now, since $P_i^{(2)}$ does not intersect P, except at p_i and p_{i+1}, we must have it either entirely contained in the interior or the exterior of P. γ is normal, so $\gamma'(t_i)$ and $\gamma(t_i^*)$ are linearly independent. It follows that the tangent line L of γ at t_i^* intersects both the interior and exterior of P, since the only line through p_i entirely contained in $\mathrm{ext}(P)$ or $\mathrm{int}(P)$ is the the tangent of γ at t_i. Parametrize L so that $L(0) = p_i$. Then $L(0)$ is the point at which parity switches on L, with respect to P. It then follows that for $\epsilon > 0$, we have $\gamma(t_i^* + \epsilon) \in \mathrm{int}(P)$ and $\gamma(t_i^* - \epsilon) \in \mathrm{ext}(P)$. Thus, we do switch parity with respect to P along γ at t_i^*.

We now make a simple observation – any intersection point p_i is adjacent to exactly four strands. We notate this set by $S(p_i) = \{P_i^{(1)}, P_i^{(2)}, P_j^{(1)}, P_j^{(2)}\}$, where $j = i - 1$ for $i > 1$ and $j = n$ for $i = 1$. Now, these strands have natural pairs since our curve is normal and regular. Write $S(p_i) = S_1(p_i) \cup S_2(p_i)$ where $S_1(P_1)$ as the two strands traversed when p_i is first crossed and $S_2(p_i)$ as the two strands traversed when p_i is reached the second time. Note that a subcurve $\gamma' \subset \gamma$ proceeds along γ at P_i iff the strands used are exactly S_1 or S_2. Lastly, we note two strands

$P_i^{(s)}$ and $P_{i+1}^{(t)}$ are of the form $S_k(p_{i+1})$ iff $P_i^{(s)}$ and $P_{i+1}^{(t)}$ are on opposite sides of P. Now, since we alternate sides along P by proceeding directly along γ, we note that the curve $\gamma' = P_1^{(2)} * \cdots * P_n^{(2)}$ alternates sides of P, switching at each p_i. Thus, we have $P_1^{(2)}$ and $P_n^{(2)}$ on opposite sides of P. But this is a contradiction to the normality of γ, since γ should then proceed along $P_1^{(2)}$ after $P_n^{(2)}$, instead of $P_1^{(1)}$. □

Proof of Lemma 6.5 (G-R Trees are Trees). *Let $T \in \mathscr{T}$. Then T is a tree.*

Proof. Suppose we had a cycle $\{v_i\}_{i=1}^{j}$. Then we first claim that we must have two unique vertices v_i, v_j which lie at the same depth. Suppose otherwise. Then since our graph T is simple, it has no self-edges or multi-edges, we must have $j \geq 3$. Then let us reorder the cycle if necessary so that the node with the minimal depth is v_1. Let this depth be d. Then we must have the depth of v_2 as $d+1$. Continuing in this fashion, we must then have v_j at depth $d+j > d+1$, but v_j is connected to v_1, which is a contradiction to property (ii).

Now, let us take two nodes v_i, v_j of the same depth from the cycle. Also, let $\gamma(T) = \phi(T)$ be the curve associated to T from the bijection ϕ, by Lemma 6.4. Let γ_i, γ_j be the unique subcurves corresponding to v_i, v_j in the unique SSD of $\gamma(T)$. The direct splits $\gamma(T)_i$ and $\gamma(T)_j$ are disjoint since v_i, v_j lie at the same depth of T. This fact is geometrically clear by the construction of the bijection ϕ. This means that we must have v_i and v_j adjacent, for there can be no path between the two which goes to a deeper level in the tree. By definition 6.1, v_i and v_j must differ in color and connected by a horizontal edge. Without loss of generality, let v_i be the green node and v_j the red node. Now, our cycle must go upwards because it cannot be entirely contained in one depth of the tree. Indeed, if this were the case, we would have a flower, which cannot occur, for any flower decomposes into just two simple curves, yet it must have at least three intersection points. This is a contradiction to the minimal structure of one intersection point per curve in the SSD. Thus, we must have some node w in the cycle, above v_i, v_j. Without loss of generality, suppose w is green, for the argument is symmetric if it is red. We cannot have two nodes at any level above v_i, v_j, for we would either have two nodes at the same whose subtrees are not disjoint, or three total nodes at different depths in our cycle. We conclude that that w is the only node at its level on our cycle. It then follows that w is connected to two green nodes u_1, u_2 in the cycle at the same level as v_i, v_j. We now know exactly what our purported cycle must look like, since it cannot go above w in depth, nor below v_i and v_j. Our cycle is of the following form: w on top, connected to a chain of alternating green and red nodes at the same depth as v_i, v_j, which eventually returns back to w. We claim we cannot have such a cycle. Let d be the depth of v_i, v_j. Let us delete the vertices from each direct splits of each node v_k at depth d in the cycle, until the direct split becomes simple. This is equivalent to deleting the subtrees of each of the nodes v_k at depth d in the cycle. Let us also delete all neighbors of w that are not on the cycle, excluding the parent of w. Now, deleting direct splits of a normal curve and then smoothing at the irregularities preserves the normality and regularity of the curve. In this case, though, after all such direct split deletion and smoothing, the direct split of w on the remaining curve is now a flower of even size. We now have our desired contradiction by Lemma A.1. □