On Defect of Plane Curves

Ian Biringer
Oberlin College

Andrew T. Barker
Wheaton College

Advisor: Juha Pohjanpelto
Oregon State University

August 15, 2003

Abstract

The combination $2St + J^+$ of Arnold invariants [3], called the defect, vanishes on treelike curves and in general indicates how much a curve differs from being treelike. In this paper we attempt to isolate the attributes of a curve that determine its defect, define operations that preserve defect, and show that defect can be regarded as an invariant of spherical curves and Gauss diagrams. Finally, a classification of the path components in the space of spherical immersions is presented.

1 Background

By a plane curve we mean a generic immersion of the circle into $\mathbb{R}^2$. A map $S^1 \to \mathbb{R}^2$ is an immersion if it is everywhere smooth, in the sense that the Jacobian never vanishes. An immersion $S^1 \to \mathbb{R}^2$ is generic if the only self-crossings are transversal double points, meaning that points of self-tangency and triple intersection are not allowed.

1.1 Index and the Whitney formula

The number ofcounterclockwise rotations made by the tangent vector when travelling along a curve K is called the Whitney index of K, and is denoted by $\text{index}(K)$. In [9], Whitney shows that

Theorem 1 *Two plane curves are homotopic in the space of immersions if and only if they have the same Whitney index.*

Whitney also proves that given any generic point $x \in K$, the index satisfies the equation

$$\sum_{\text{double points } d \in K} \varepsilon_d(x) = \text{index}(K) - 2\,\text{ind}_K(x). \qquad (1)$$

Given a double point $d \in K$, the sign $\varepsilon_d(x)$ of d with respect to x is +1 if the outgoing branches of K at d (in the order in which they are visited when travelling along the curve starting at x) orient the plane clockwise, and -1 otherwise. The index $\text{ind}_K(x)$ of a generic point $x \in K$ is the average

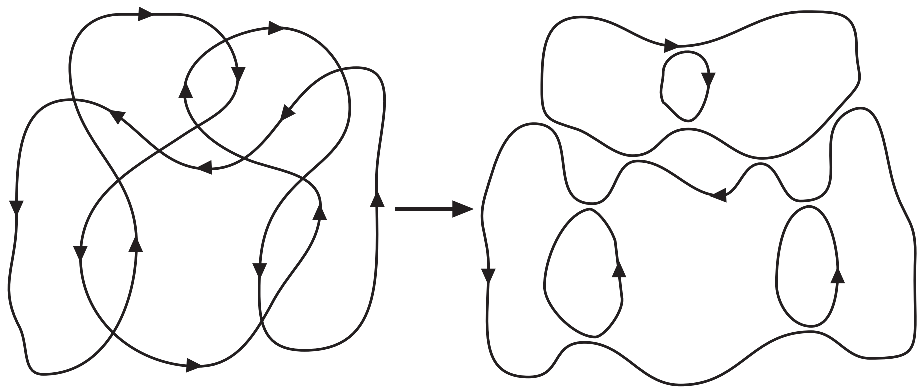

Figure 1: A curve and its Seifert decomposition.

of the region indices of the two regions of $\mathbb{R}^2 - K$ adjacent to x, where the region index of a region is simply the winding number of K about any point in that region. This definition can be extended to double points of K by averaging the indices of the four regions adjacent to the point. Note that the indices of points on K need not be integers.

1.2 The Seifert decomposition

A useful combinatorial representation of a plane curve is obtained by splitting the curve at each of its double points to produce a collection of circles, connected to each other tangentially at the locations of the double points of the original curve. See Figure 1. This representation of the curve is called its Seifert decomposition, and the circles are its Seifert cycles. If two Seifert cycles connect and one is contained inside the other, we say that they share an interior connection. If they connect and neither is contained inside the other, we say that they share an exterior connection.

The Seifert cycles of a curve have a natural orientation, induced by the orientations of the arcs of the curve. We define the index of a Seifert cycle c, denoted $\text{index}(c)$, to be +1 if the cycle is oriented counterclockwise and -1 otherwise. Note that if two cycles share an exterior connection then their indices must be opposite, while if they share an interior connection their indices must be the same. The Seifert cycles of a curve K provide a simple formula for the Whitney index:

$$\text{index}(K) = \sum_{\text{Seifert cycles } c \in K} \text{index}(c). \tag{2}$$

1.3 The Arnold invariants

Definitions of the Arnold invariants St, J^+ and J^- are given in [2] and [3]. In [8], Shumakovich shows that for any generic point $x \in K$

$$St = \sum_{d \in K} \varepsilon_d(x)\text{ind}_K(d) + \text{ind}_K(x)^2 - \frac{1}{4}, \tag{3}$$

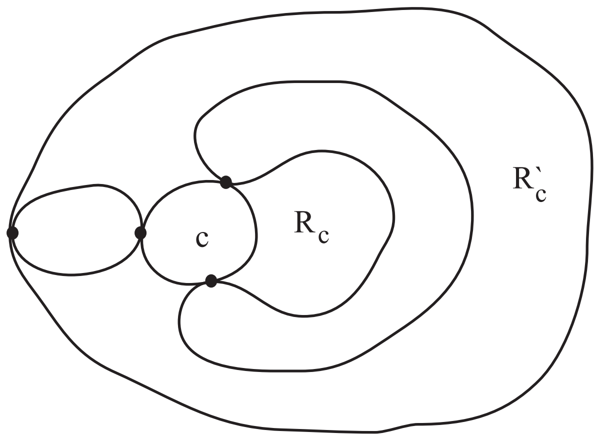

Figure 2: A Seifert cycle c and two regions R'_c and R_c which border an arc of c from the outside. The regions are different, but since their region indices are the same $\mathrm{ind}(R_c)$ is well defined for any Seifert cycle c.

where d ranges over the double points of K.

In [5], Luo derives explicit formulas for J^+ and J^- which involve the decomposition of a curve into Seifert cycles. The formula for J^+ uses a function $t(c)$ defined on the Seifert cycles of a curve K by

$$t(c) = \mathrm{index}(c)\mathrm{ind}_K(R_c), \tag{4}$$

where $\mathrm{ind}_K(R_c)$ is the region index of some (any) region which borders some arc of c from the outside. See Figure 2. We call this function the t-function of K. The following lemma shows that the value of $\mathrm{ind}_K(R_c)$ does not depend on the choice of the region R_c.

Lemma 1 (Luo [5]) *The index of a region is equal to the sum of the indices of Seifert cycles which contain the region.*

Luo proves that the t-function satisfies the recursive relation

$$t(c) = \begin{cases} 0 & \text{if } c \text{ is a cycle with exterior boundary,} \\ t(s)+1 & \text{if } s \text{ and } c \text{ share an interior connection and } c \text{ is inside } s, \\ -t(s) & \text{if } c \text{ and } s \text{ share an exterior connection.} \end{cases} \tag{5}$$

Finally, using the t-function and relation (5), he shows that the invariants J^+ and J^- are given by

$$J^+ = 1+n-s-2\sum_c t(c), \qquad J^- = 1-s-2\sum_c t(c) \tag{6}$$

where n is the number of double points and s is the number of Seifert cycle in the decomposition.

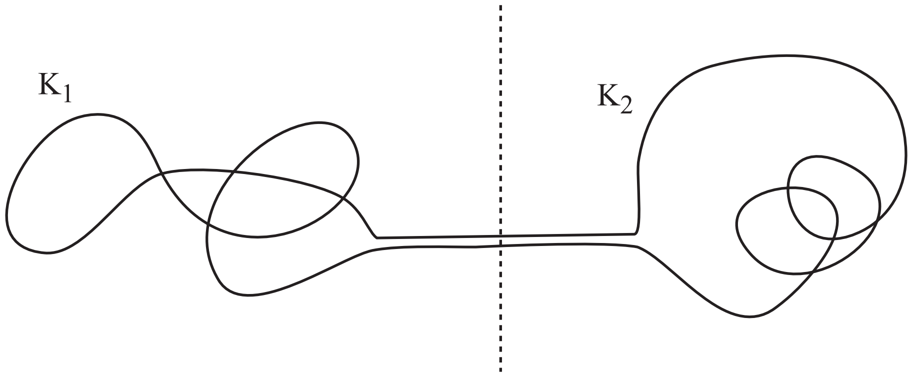

Figure 3: A connected sum of curves K_1 and K_2.

1.4 Connected sum

Arnold proves in [2] that all three invariants are additive under the connected sum; in fact, this additivity motivates much of the way in which he normalizes the invariants and coorients the discriminant hypersurface consisting of immersions with self-tangency and triple-point singularities. To form the connected sum of two curves K_1 and K_2, place one in the left half-plane and the other in the right half-plane and join them with a connecting bridge as defined in [3]. See Figure 3. Unfortunately, the connected sum is not a well-defined operation on plane curves, because for some pairs of curves it is impossible to construct the connected sum—consider two circles oriented in opposite directions—while for others there are multiple (inequivalent) ways to form the connecting bridge. But if a connected sum exists, it is additive on the invariants J^+, J^-, and St, and thus additive on defect.

1.5 Treelike curves and defect

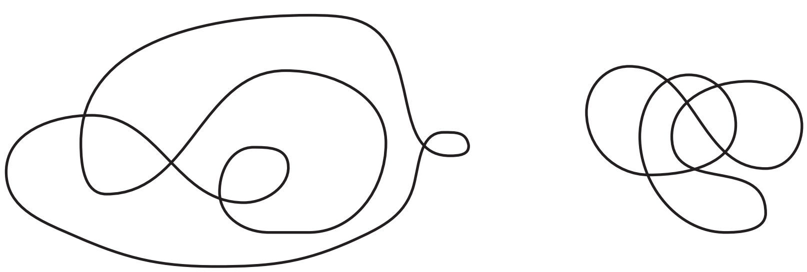

Figure 4: A treelike curve and a nontreelike curve.

A curve is treelike if each of its double points divides it into two pieces (see Figure 4). In [1], Aicardi shows that all treelike curves have planar Gauss diagrams and defect 0. She provides

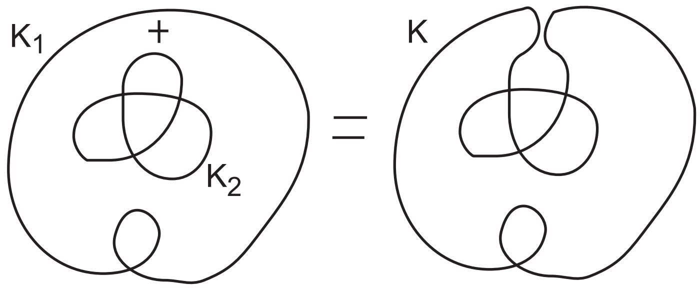

Figure 5: An interjected sum.

simple formulas for the Arnold invariants of treelike curves, one of which is $J^+ = -2St$, but since here our concern is primarily with non-treelike curves her formulas are of limited interest to us.

2 Interjected sum

The usual connected sum of plane curves K_1 and K_2 from topology or differential geometry is formed by embedding K_2 in a component of $\mathbb{R}^2 - K_1$ and connecting K_1 to K_2 with a bridge as defined earlier. Arnold's connected sum is a special case of this, requiring that K_2 be embedded into the unbounded component of $\mathbb{R}^2 - K_1$. He imposes this extra restriction because his invariants are not additive under the usual connected sum. We will show, however, that defect is additive under this sum, which from now on we will refer to as the *interjected sum*, to distinguish it from Arnold's connected sum.

Definition 1 *Let K_1, K_2 be plane curves. The* **interjected sum** *K of K_2 into K_1 is formed by embedding K_2 into a component of $\mathbb{R}^2 - K_1$, then connecting K_1 to K_2 with a bridge as defined in [3] . See Figure 5.*

Proposition 1 *Let K_1, K_2 be plane curves, and let K be an interjected sum of K_2 into a region R_1 of $\mathbb{R}^2 - K_1$. Assume that the bridge connects to an arc of K_2 bordering the (bounded) region R_2 of $\mathbb{R}^2 - K_2$. Then $J^+_K = J^+_{K_1} + J^+_{K_2} - 2[\mathrm{ind}_{K_1}(R_1)](\mathrm{index}(K_2) + \mathrm{ind}_{K_2}(R_2))$.*

Proof. Let γ_1, γ_2 be the Seifert cycles of K_1, K_2 containing the arcs on which the connecting bridge is formed. Considering γ_1 as the combined cycle formed in the interjected sum, we can identify the cycles of K with those of K_1 and K_2, other than γ_2 which we consider destroyed by the sum. Furthermore, if t_{K_1}, t_{K_2} and t_K denote the t-functions of K_1, K_2 and K, we can calculate J^+_K from $J^+_{K_1}$ and $J^+_{K_2}$ by analyzing how the values of t_K at cycles of K differ from the values of the other t-functions at the corresponding cycles of K_1 or K_2.

First, if c is a cycle of K identified with a cycle of K_1, then $t_K(c) = t_{K_1}(c)$. Since the cycles of K_2 are interjected into a single region and thus cannot contain any cycle of K, the cycles containing c in K correspond exactly to those containing c in K_1. So applying Lemma 1 we see that $\mathrm{ind}_K(R^K_c) = \mathrm{ind}_{K_1}(R_c)$, and the equality follows.

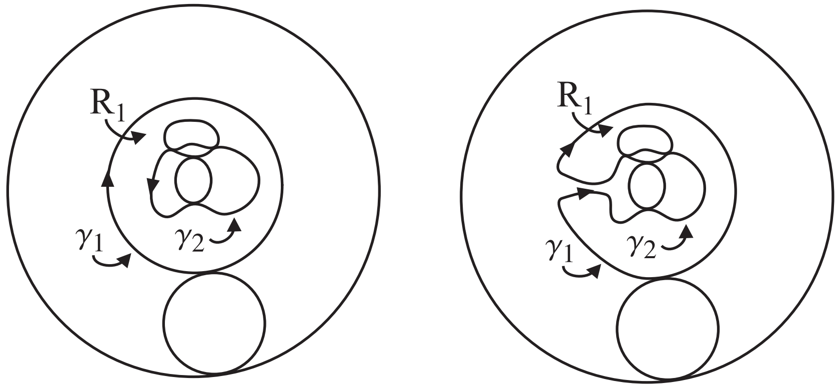

Figure 6: The case where γ_1 contains R_1.

Lemma 2 *For a cycle $c \in K_2$, $c \neq \gamma_2$, let $\Delta t(c) = t_K(c) - t_{K_2}(c)$. Assume that* $\text{index}(\gamma_2) = -1$. *Then*

$$\Delta t(c) = \begin{cases} -\Delta t(s) & \text{if } c \text{ and } s \neq \gamma_2 \text{ share an exterior connection in } K_2, \\ \Delta t(s) & \text{if } c \text{ and } s \neq \gamma_2 \text{ share an interior connection in } K_2, \\ -\text{ind}_{K_1}(R_1) & \text{if } c \text{ is contained in and connects to } \gamma_2 \text{ in } K_2, \\ \text{ind}_{K_1}(R_1) & \text{if } c \text{ and } \gamma_2 \text{ share an exterior connection in } K_2. \end{cases} \tag{7}$$

Proof. First, if c and $s \neq \gamma_2$ share an exterior connection in K_2, then they also do in K. So, (5) implies that $\Delta t(c) = t_K(c) - t_{K_1}(c) = -t_K(s) + t_{K_1}(s) = -\Delta t(s)$. Also, if c and $s \neq \gamma_2$ share an interior connection in K_2, then $\Delta t(c) = t_K(c) - t_{K_1}(c) = (t_K(s) \pm 1) - (t_{K_1}(s) \pm 1) = t_K(s) - t_{K_1}(s) = \Delta t(s)$.

Now, assume that the region R_1 is contained in the cycle γ_1. In this case, any cycle of K_2 which was contained in and connected to γ_2 now shares an exterior connection with γ_1, while any cycle previously sharing an exterior connection with γ_2 now shares an interior connection with γ_1. Moreover, since we have assumed that $\text{index}(\gamma_2) = -1$, in this case $\text{index}(\gamma_1) = +1$. See Figure 6.

If c is contained in and connects to γ_2 in K_2, then since γ_2 is an exterior cycle of K_2, $t_{K_2}(c) = 1$ by (5). But now we must have

$$\begin{aligned} t_K(c) &= -t_K(\gamma_1) \\ &= -\text{index}(\gamma_1)\text{ind}_K(R_{\gamma_1}) \\ &= -\text{ind}_K(R_{\gamma_1}) \\ &= -(\text{ind}_{K_1}(R_1) - 1) \end{aligned}$$

using the notation of equation (4). But then $\Delta t(c) = (-\text{ind}_{K_1}(R_1) + 1) - 1 = -\text{ind}_{K_1}(R_1)$.

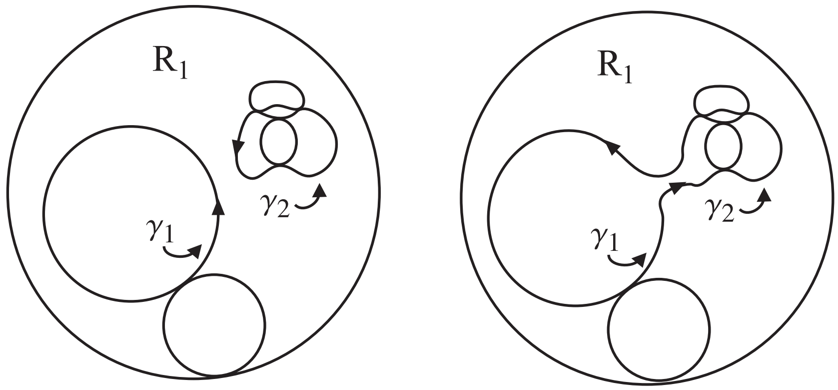

Figure 7: The case where γ_1 does not contains R_1.

On the other hand, if c shares an exterior connection with γ_2 in K_2, then since γ_2 is an exterior cycle of K_2, $t_{K_1}(c) = 0$ by (5). Then since c shares an interior connection with γ_1 in K,

$$\begin{aligned} t_K(c) &= t_K(\gamma_1) + 1 \\ &= \text{index}(\gamma_1)\text{ind}_K(R_{\gamma_1}) + 1 \\ &= \text{ind}_K(R_{\gamma_1}) + 1 \\ &= \text{ind}_{K_1}(R_1) \end{aligned}$$

but in this case $\Delta t(c) = \text{ind}_{K_1}(R_1)$.

If the region R is not contained in the cycle γ_1, then the way any cycle was connected to γ_2 in K_2 remains the same when the cycle is connected to γ_1 in K, and $\text{index}(\gamma_1) = -1$. See Figure 7. In this new case, a cycle c sharing an interior connection with γ_2 in K_2 must satisfy

$$\begin{aligned} t_K(c) &= t_K(\gamma_1) + 1 \\ &= \text{index}(\gamma_1)\text{ind}_K(R_{\gamma_1}) + 1 \\ &= -\text{ind}_K(R_{\gamma_1}) + 1 \\ &= -\text{ind}_{K_1}(R_1) + 1, \end{aligned}$$

implying that $\Delta t(c) = -\text{ind}_{K_1}(R_1)$. Similarly, a cycle c sharing an exterior connection with γ_2 in K_2 must satisfy

$$\begin{aligned} t_K(c) &= -t_K(\gamma_1) \\ &= -\text{index}(\gamma_1)\text{ind}_K(R_{\gamma_1}) \\ &= \text{ind}_K(R_{\gamma_1}) \\ &= \text{ind}_{K_1}(R_1), \end{aligned}$$

and it follows that $\Delta t(c) = \text{ind}_{K_1}(R_1)$. ■

Lemma 3 *Let $c \neq \gamma_2$ be a Seifert cycle of K_2. Then* $\Delta t(c) = \text{index}(c)\text{ind}_{K_1}(R_1)$.

Proof. Without loss of generality, assume that $\text{index}(\gamma_2) = -1$. If it is not, then reverse the orientations of K, K_1 and K_2. Proving the lemma for the new curves will prove it for the old curves, since $\Delta t(c)$ is unchanged by the reversal, while the new values of $\text{index}(c)$ and $\text{ind}_{K_1}(R_1)$ are opposite the old values. If c shares an exterior connection with γ_2 then its index must be +1, while if c is contained in and connects to γ_2 in K_2 then its index must be -1. The previous lemma then establishes the result for these cases. Inductively, assume that $\Delta t(c) = \text{index}(c)\text{ind}_{K_1}(R_1)$ holds for some cycle $c \in K_2$. Then it must also hold for any cycle $s \in K_2$ which shares a connection with c, for if the connection is interior then $\Delta t(c) = \Delta t(s)$ and $\text{index}(s) = \text{index}(c)$, and if it is exterior then $\Delta t(c) = -\Delta t(s)$ and $\text{index}(s) = -\text{index}(c)$. Since there is a path of Seifert cycles, each connected to the next, to any given cycle of K_2 from a cycle of the first two cases, this is sufficient to prove the lemma. ■

Now, Lemma 3 and equation (2) imply that

$$\sum_{c \in K_2, c \neq \gamma_2} \Delta t(c) = \text{ind}_{K_1}(R_1)(\text{index}(K_2) - \text{index}(\gamma_2)). \tag{8}$$

Finally, assume that n, n_1, n_2, s, s_1, s_2 are the numbers of double points and numbers of Seifert cycles of K, K_1 and K_2, respectively. Then Luo's formula (6) gives

$$\begin{aligned}
J_K^+ &= 1 + n - s - 2\sum_{c \in K} t_K(c) \\
&= 1 + (n_1 + n_2) - (s_1 + s_2 - 1) - 2\Big(\sum_{c \in K_1} t_K(c)\Big) - 2\Big(\sum_{c \in K_2, c \neq \gamma_2} t_K(c)\Big) \\
&= 1 + (n_1 + n_2) - (s_1 + s_2 - 1) - 2\Big(\sum_{c \in K_1} t_{K_1}(c)\Big) - 2\Big(\sum_{c \in K_2, c \neq \gamma_2} (t_{K_2}(c) + \Delta t(c))\Big) \\
&= J_{K_1}^+ + 1 + n_2 - s_2 - 2\Big(\sum_{c \in K_2, c \neq \gamma_2} t_{K_2}(c)\Big) - 2\,\text{ind}_{K_1}(R_1)(\text{index}(K_2) - \text{index}(\gamma_2)) \\
&= J_{K_1}^+ + 1 + n_2 - s_2 - 2\Big(\sum_{c \in K_2} t_{K_2}(c)\Big) - 2\,\text{ind}_{K_1}(R_1)(\text{index}(K_2) - \text{index}(\gamma_2)) \\
&= J_{K_1}^+ + J_{K_2}^+ - 2\,\text{ind}_{K_1}(R_1)(\text{index}(K_2) - \text{index}(\gamma_2)).
\end{aligned}$$

Noting that $\text{index}(\gamma_2) = \text{ind}_{K_2}(R_2)$ completes the proof. ■

Proposition 2 *Let K_1, K_2 be plane curves, and let K be an interjected sum of K_2 into a region R_1 of $\mathbb{R}^2 - K_1$. Assume that the bridge connects to an arc of K_2 bordering the (bounded) region R_2 of $\mathbb{R}^2 - K_2$. Then $St_K^+ = St_{K_1}^+ + St_{K_2}^+ + \text{ind}_{K_1}(R_1)(\text{index}(K_2) + \text{ind}_{K_2}(R_2))$.*

Proof. Consider the Shumakovich formula (3) for the strangeness of K and pick a base point x on the connecting bridge. The signs $\varepsilon_d(x)$ of double points $d \in K$ coincide with the signs $\varepsilon_{d_1}(x_1), \varepsilon_{d_2}(x_2)$ of the corresponding double points $d_1 \in K_1, d_2 \in K_2$, where x_1, x_2 are the points on K_1, K_2 where the connecting bridge is attached. Also,

$$\text{ind}_K(x) = \text{ind}_{K_1}(x_1), \tag{9}$$

and since x_2 lies on the exterior contour of K_2,

$$\mathrm{ind}_{K_2}(x_2) = \pm\frac{1}{2}. \tag{10}$$

There exists a natural surjection $R \to \widehat{R}$ from the regions of $\mathbb{R}^2 - K_1$ and $\mathbb{R}^2 - K_2$ to those of $\mathbb{R}^2 - K$. The surjection is nearly one-to-one, except that the unbounded component of $\mathbb{R}^2 - K_2$ is sent to the same region of $\mathbb{R}^2 - K$ as is the region R_1 of $\mathbb{R}^2 - K_1$, and the region R_2 of K_2 is sent to the same region of $\mathbb{R}^2 - K$ as is the region of $\mathbb{R}^2 - K_1$ into which the connecting bridge opens. It is clear that if R is a region of $\mathbb{R}^2 - K_1$, then $\mathrm{ind}_{K_1}(R) = \mathrm{ind}_K(\widehat{R})$. Also, if R is a region of $\mathbb{R}^2 - K_2$, then $\mathrm{ind}_K(\widehat{R}) = \mathrm{ind}_{K_2}(R) + \mathrm{ind}_{K_1}(R_1)$. For, if we fix a point $p \in \widehat{R}$, it should be intuitively obvious that the winding number of K about p is the sum of the winding numbers of K_1 and K_2 about corresponding points in R_1 and R_2. A rigorous proof of this can be given using either the integral formula for winding number or by relating the Seifert cycles of K containing p with those of K_1 and K_2 containing the corresponding points. Finally, given a double point d of K_1 or K_2, let the corresponding double point of K be $\widehat{d}$. Recalling that the index of a double point is the average of the indices of the regions it borders, we have that

1. if $d \in K_1$, $\mathrm{ind}_K(\widehat{d}) = \mathrm{ind}_{K_1}(d)$, and
2. if $d \in K_2$, $\mathrm{ind}_K(\widehat{d}) = \mathrm{ind}_{K_2}(d) + \mathrm{ind}_{K_1}(R)$.

Now by the Shumakovich formula (3),

$$\begin{aligned}
St_K &= \sum_{\widehat{d}\in K} \varepsilon_{\widehat{d}}(x)\mathrm{ind}_K(\widehat{d}) + \mathrm{ind}_K(x)^2 - \frac{1}{4} \\
&= \sum_{d\in K_1} \varepsilon_d(x_1)\mathrm{ind}_K(d) + \mathrm{ind}_K(x)^2 - \frac{1}{4} + \sum_{d\in K_2} \varepsilon_d(x_2)(\mathrm{index}_{K_2}(d) + \mathrm{index}_{K_1}(r)) \\
&= St_{K_1} + \sum_{d\in K_2} \varepsilon_d(x_2)\mathrm{ind}_K(d) + \mathrm{ind}_{K_1}(r)\sum_{d\in K_2} \varepsilon_d(x_2) \\
&= St_{K_1} + St_{K_2} + \mathrm{ind}_{K_1}(r)(\mathrm{index}(K_2) - 2\,\mathrm{ind}_{K_2}(x_2)).
\end{aligned}$$

Noting that $2\,\mathrm{ind}_{K_2}(x_2) = \mathrm{ind}_{K_2}(R_2)$ establishes the desired equation. ■

Theorem 2 *Defect is additive under the interjected sum.*

Proof. The theorem follows from Propositions 1 and 2. ■

3 Flip

Given two generic points p, q on a curve that separate the curve into two disjoint pieces, we can flip one of the pieces over to produce a new curve. See Figure 8. Using the interjected sum, we can show that this operation preserves defect. Rigorously, we define the flip as follows:

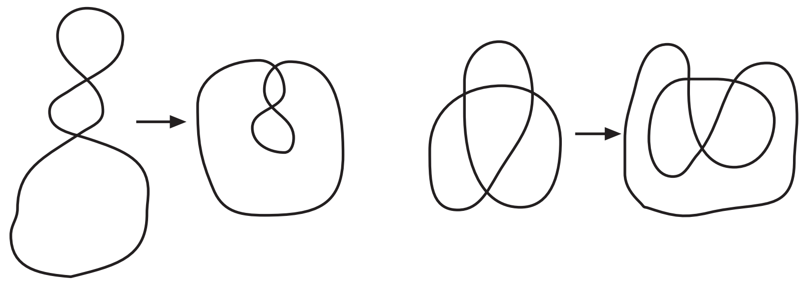

Figure 8: Two examples of flips.

Definition 2 *Let K be a plane curve. Choose two generic points p, q (called flip points) that separate the curve into two disjoint pieces, R and S. Assume that p and q are not contained in a region bounded by R. Since R does not intersect S, this implies that no point of S lies in a region of the plane bounded by R. So, we can find a diffeomorphism of the curve that sends R into the open unit disk, the flip points to (-1,0) and (1,0) and S outside the unit disc. Smooth the curve around the flip points so that the tangent vectors lie on the line pq. A* **flip** *of K is then a reflection of R about the line pq. See Figure 9.*

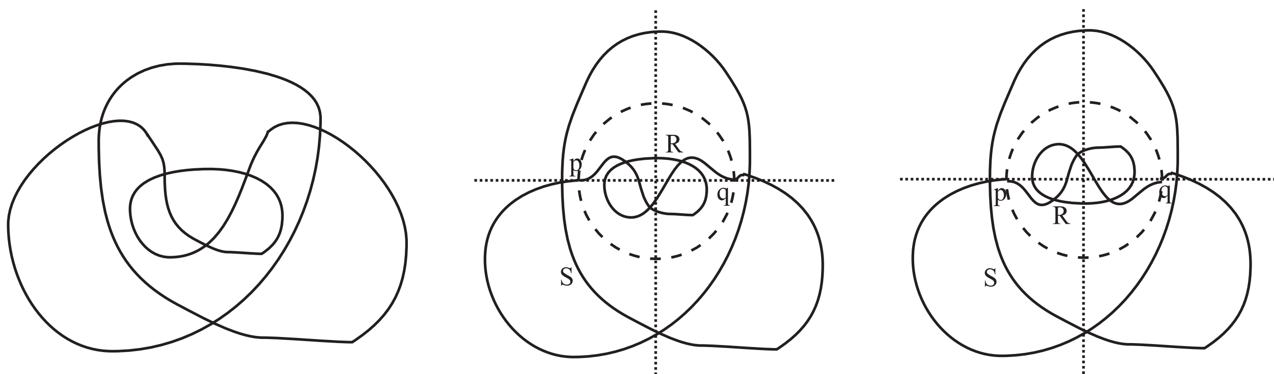

Figure 9: A curve before the diffeomorphism, after the diffeomorphism and after the flip.

Note that the conditions on this flip require that both flip points are on the same Seifert cycle of K. For if they are not, then neither can separate its Seifert cycle into two components and so there must exist a path along the curve from one side of p to its other side that does not pass through q. Also, note that the flip does not change the order in which double points are visited, so it does not affect the Gauss diagram of the curve.

Theorem 3 *The flip preserves defect.*

Proof. Let K be a plane curve and let K' be a flip of K about the flip points p, q. Cut K at p and q, and identify the ends of the segments S and R to form two new plane curves K_S and K_R. Then K

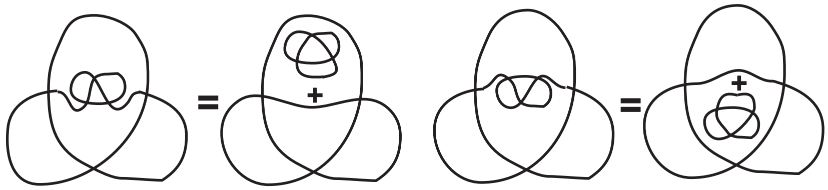

Figure 10: Two interjected sums forming two curves with the same defect, which differ by a flip.

and K' are both interjected sums of K_R into K_S. Since defect is additive under the interjected sum, K and K' must have the same defect. See Figure 10. ■

4 Gauss diagram and defect

In this section we prove that all curves corresponding to the same Gauss diagram have the same defect. In most cases, a curve K can be transformed into any other curve K' with the same Gauss diagram by a sequence of flips, a defect preserving operation. In some complicated cases this is not possible, but we show that it is always possible to transform K into a curve with precisely the same Arnold invariants as K' by a sequence of defect preserving moves.

To do so, we introduce a structure called the B-graph, which together with the Gauss diagram completely determines the Arnold invariants of a curve (Lemma 4). We then characterize which B-graphs correspond to real curves and define a flip operation on the B-graph which corresponds to a flip on its curve. Finally, we show that the B-graph of a curve K can be transformed through a sequence of flips into that of any other curve K' with the same Gauss diagram, showing that K and K' have the same defect.

Definition 3 *An* **arc** *of a Gauss diagram is an arc of the boundary circle between two adjacent endpoints of chords.*

Definition 4 *Let G be a Gauss diagram. Assign an orientation to the boundary circle of G. An* **S-cycle** *of G is a cyclicly ordered subset* $\{a_i, \ldots, a_k\}$ *of the chords of G such that there exists a labeling on the endpoints of* a_i *as* α_i, β_i *so that* $\overrightarrow{\alpha_i\beta_{i+1}}$ *is an arc of G that follows the orientation of the boundary circle.*

Intuitively, an S-cycle is an alternating sequence of chords and arcs on the Gauss diagram where the arcs all follow a given orientation of the boundary circle. See Figure 11. For any Gauss diagram, a given arc connects the chords of exactly one S-cycle. For once the chords a_i, a_{i+1} and the arc $\overrightarrow{a_i a_{i+1}}$ between them are given, the rest of the cycle is completely determined. Each chord, on the other hand, belongs to two different S-cycles.

A chord in the Gauss diagram of a curve represents a double point. In the rest of this section, when considering Gauss diagrams which correspond to real curves we will regard the elements of an S-cycle as double points, rather than chords.

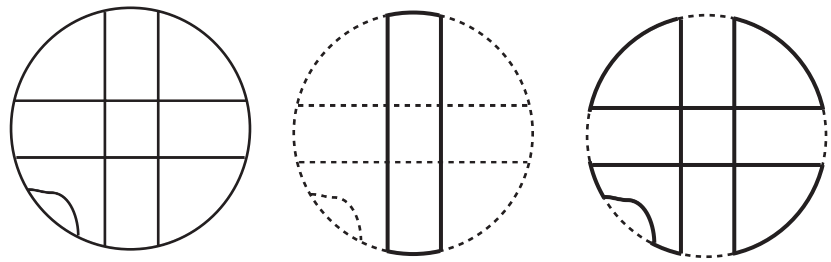

Figure 11: A Gauss diagram alone and then with two different S-cycles highlighted. This Gauss diagram has two S-cycles not highlighted.

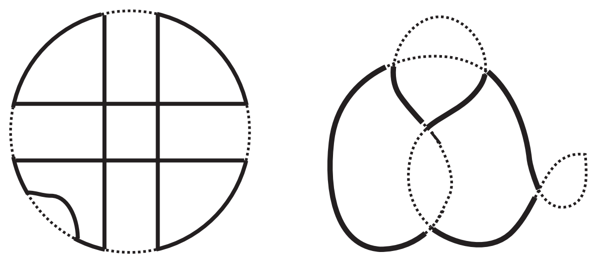

Figure 12: An S-cycle and the arcs of its corresponding Seifert cycle.

Theorem 4 *Let K be a plane curve and G be the Gauss diagram corresponding to it. Then there exists a natural bijection between the Seifert cycles of K and the S-cycles of G such that the cyclicly ordered subset of double points bordering a Seifert cycle is exactly the corresponding S-cycle.*

Proof. The arcs on such an alternating path through G are precisely the arcs of K which lie on some Seifert cycle: the pairing of the four arcs adjacent to a given double point induced by the S-cycles of G is exactly that given by the Seifert splitting algorithm. See Figure 12. ■

From now on, when considering Gauss diagrams of curves we will use the terms S-cycle and Seifert cycle interchangeably, referring to s as an S-cycle when considering it as a set of double points and as a Seifert cycle otherwise. It is worth noting that the above construction of S-cycles gives us a way to calculate the number s of Seifert cycles in a plane curve using only the Gauss diagram. Using this, we can use the formula in Luo [5] to calculate the Euler characteristic $\chi = s+1-n$ of the Seifert surface and alternating knot corresponding to a plane curve directly from its Gauss diagram.

Definition 5 *Let G be a Gauss diagram, and let $s_1,\ldots,s_n$ denote its S-cycles. Then the* **B-graph** *of G is the graph $B_G = (V,E)$, where $V = \{s_1,\ldots,s_n\}$, and $E = \{(s_i,s_j) : s_i \cap s_j \neq \emptyset\}$.*

By the definition above, two S-cycles are connected by an edge if they contain the same chord on the Gauss diagram; that is, each Seifert cycle becomes a vertex of the B-graph and any two Seifert cycles that are connected by one or more double points are connected by an edge on the B-graph. The B-graph is in some sense an artificial structure—it is simply a convenient representation

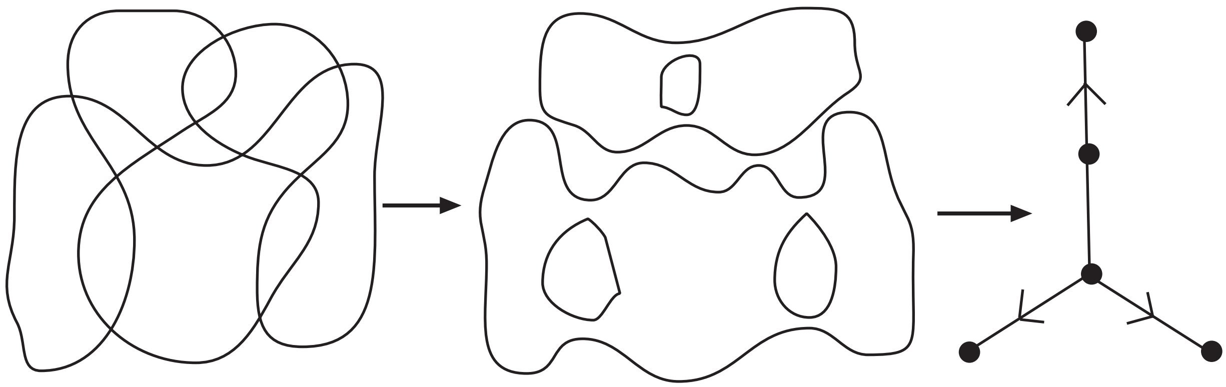

Figure 13: A curve, its decomposition into Seifert cycles, and the associated directed B-graph.

of the Seifert cycle decomposition that discards the information which is not important in this section. In fact, thinking about the following results in terms of Seifert cycles will be very helpful. However, since the B-graph also provides a more convenient language in which to phrase some of the our results, we feel that its use is merited.

So for the Seifert cycle decomposition of a curve K with Gauss diagram G, the B-graph B_G indicates the Seifert cycles of K and which cycles share a connection. The S-cycles themselves show the number of double points connecting two cycles and the order in which all the connections to a given Seifert cycle occur. This information is the same for all curves with the same Gauss diagram.

To more precisely determine a curve with a given Gauss diagram G, we must include information describing the ways in which two cycles connect. We can express this by adding directions to the edges of a B-graph so that

$$E = \{ \ (s_i, s_j) : \text{either } s_i, s_j \text{ share an exterior connection} \\ \text{or } s_i \text{ contains } s_j \text{ and } s_i, s_j \text{ share an interior connection } \}.$$

By this definition, an edge containing vertices s_i and s_j is directed towards s_j if s_j is inside s_i and undirected if neither contains the other. The result is a directed B-graph; see Figure 13.

Much of the following section is spent developing conditions that allow us to “flip” the directions of the edges of a B-graph in such a way that the corresponding operation on its curve is a flip as described in the last section. Flips will be made on sets of edges adjacent to a given vertex.

Lemma 4 *Let K, K' be plane curves with the same Gauss diagram and directed B-graph. Then the Arnold invariants of K and K' are identical.*

Proof. We show that all three invariants can be calculated using only the Gauss diagram and the directed B-graph.

For J^+ and J^-, we use the Luo formulas (6). The B-graph includes all information necessary to calculate the t-function; it follows that two curves with identical B-graphs have the same values for J^+ and J^-.

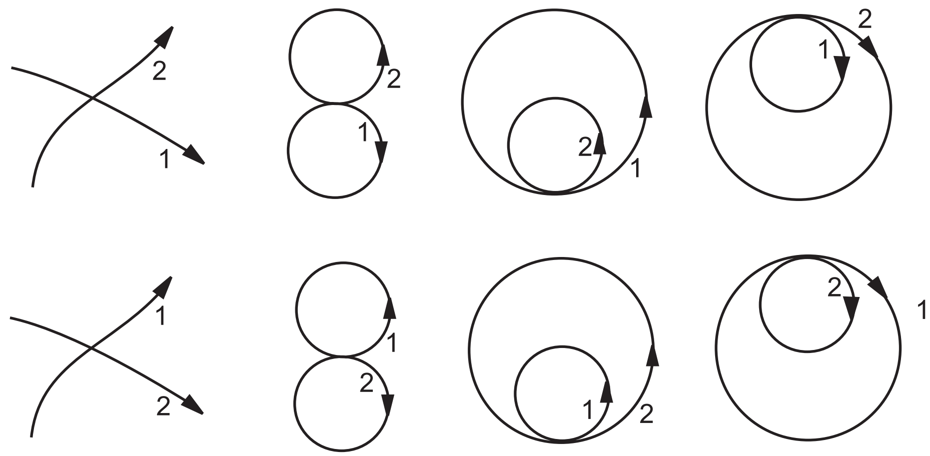

Figure 14: Double points with positive and negative sign, and the possible connections and orientations of the Seifert cycles each connects.

For strangeness, we use the Shumakovich formula (3). Pick a base point x on some Seifert cycle of K, and a point $x' \in K'$ that lies on the same arc of the corresponding cycle. Since the cycles of K connect in the same way as those of K', we can assign the same set of orientations to the cycles of K as to those of K'. Now, the sign of a double point is completely determined by the orientations of the two Seifert cycles it connects and the branch of the curve at the double point that a path from the base point takes first; see Figure 14. However, given a double point $d \in K$ and its corresponding point $d' \in K'$, the paths from x to d and from x' to d' are provided by the Gauss diagram, so the branches at d and d' taken first lie on corresponding Seifert cycles and the signs of d and d' with respect to x must coincide. The region indices of each Seifert cycle, and therefore of the whole curve, can be calculated from the B-graph using Lemma 1, so it follows that two curves with the same Gauss diagram and directed B-graph have the same strangeness. ■

Not all directed B-graphs correspond to actual curves. In particular, not all assignments of directions to the edges containing a vertex s will result in a B-graph of a curve, no matter what the other directions in the graph may be. The goal of much of what follows is to formalize constraints on these directions that determine which B-graphs correspond to real curves.

Definition 6 *Let s be a vertex of B_G. A* **primitive link** *of B_G at s is a set L of edges containing s such that*

1. *For every two edges $e = (s, s_1)$ and $f = (s, s_2)$ in L there exists a path from s_1 to s_2 which does not contain s and*
2. *L is maximal in the sense that there is no proper superset of L which satisfies the first condition.*

A primitive link can contain one edge or more, and the primitive links partition the edges containing s.

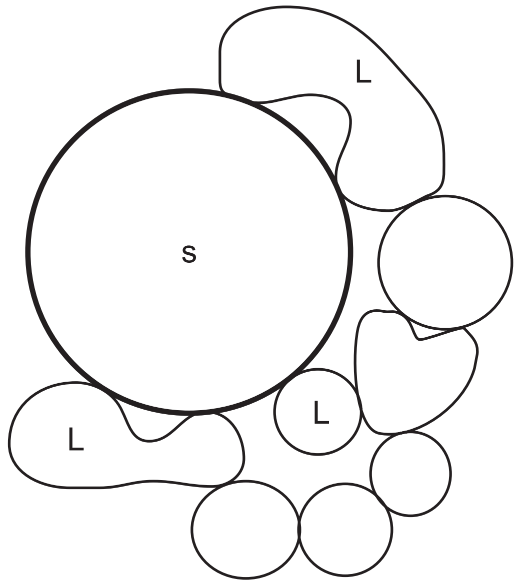

Figure 15: A vertex s and the vertices connected to s by edges of a primitive link L.

From now on, we will call an edge e containing s *inside* with regard to s if it is directed away from s and *outside* with regard to s if it is undirected or directed towards s. This tag is designed to reflect the connections of Seifert cycles represented by a directed B-graph.

If K is a curve and B_G is its directed B-graph, then given a primitive link at a vertex s, either all edges of the primitive link are inside with regard to s or they are all outside. In terms of Seifert cycles, this means that either all cycles connected to s by edges of the primitive link are contained in s, or they all are connected outside of s. This is evident if you consider s_1, s_2 connected to s from inside and outside, respectively, with edges that belong to a primitive link; for then there must exist a path of Seifert cycles starting inside s and ending outside without passing through s, a clear contradiction.

We can impose further restrictions on the directions of a B-graph by combining primitive links in the following process.

Definition 7 *If X is a set of edges connected to a vertex s in a B-graph, we denote by $D(X)$ the set of all double points represented by the edges of X, and similarly for a collection, $D(C) = \bigcup_{L \in C} D(L)$. Given a set $Y \subset s$ of double points with s an S-cycle, we denote the set of all edges e containing s such that $D(e) \cap Y \neq \emptyset$ by the symbol $E(Y)$.*

Note that D and E are almost inverses, in the sense that $E(D(X)) = X$. Also, if $Y \subset s$ is a set of double points which is the union of all the double points of certain edges, then $D(E(Y)) = Y$.

Begin with $C_1 = \{A_1\}$ where A_1 is an arbitrary primitive link at the vertex s. Repeatedly add primitive links to C_1 where at each step a primitive link B is added to C_1 if there exist $a_1, a_2 \in D(C_1), b_1, b_2 \in B$ such that in the cyclic order of s,

$$a_1 < b_1 < a_2 < b_2. \tag{11}$$

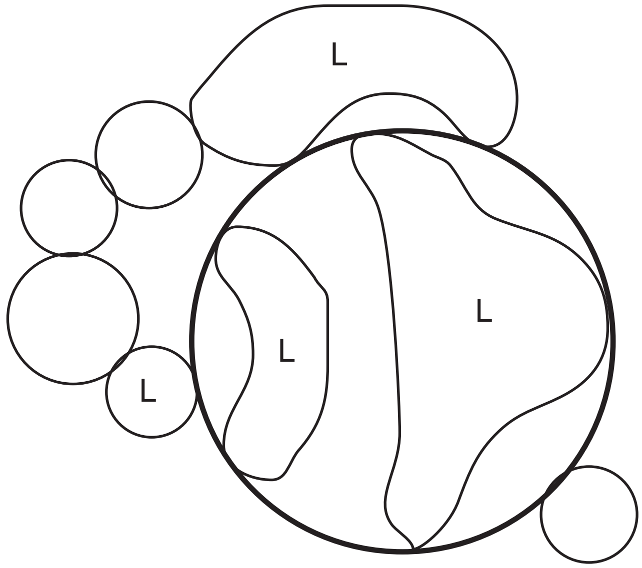

Figure 16: A vertex s and the vertices connected to s by edges of a link L.

When no more primitive links can be added to C_1, pick a new primitive link A_2 not in C_1 and begin the process again, with $C_2 = \{A_2\}$ considering now only the primitive links not in C_1.

Definition 8 *A* **link** *of B_G at s is a set $\bigcup_{L\in C_i} L$, where the collection C_i is constructed in the discussion above.*

This link construction process will result in a partition $C_1, \cdots, C_n$ of the primitive links, and since the primitive links partition the edges containing s, the set of links $\{\bigcup_{L\in C_i} L : i = 1, \cdots, n\}$ is a partition of the edges containing s.

The directed B-graph of a curve must in some sense respect its links. This is formalized in the following lemma.

Lemma 5 *Let L be a link of B_G at s. Then there is a labeling or tag of "inside" and "outside" to the edges of L such that the directions given to these edges by any plane curve with Gauss diagram G either agree with the tags of every edge or disagree with the tags of every edge.*

Proof. Assume $L = \bigcup_{i=1}^{n} A_i$ where each A_i is a primitive link of B_0. We proceed by induction on k.

If $k = 1$, then L is a primitive link, and as discussed earlier a plane curve gives either all outside tags or all inside tags to elements of L.

Assume that $\bigcup_{i=1}^{n-1} A_i$ satisfies the tag property. Then let $L = \bigcup_{i=1}^{n} A_i$. By construction there exist $a, a' \in D(\bigcup_{i=1}^{n-1} A_i), b, b' \in A_n$ such that in the cyclic order of s, $a < b < a' < b'$. Then b and b' define a natural partition of $D(\bigcup_{i=1}^{n-1} A_i) = \{a_1, \ldots, a_m\}$ induced by the sequence

$$a_1, \ldots, a_i < b_1 < a_{i+1}, \ldots, a_m < b_2.$$

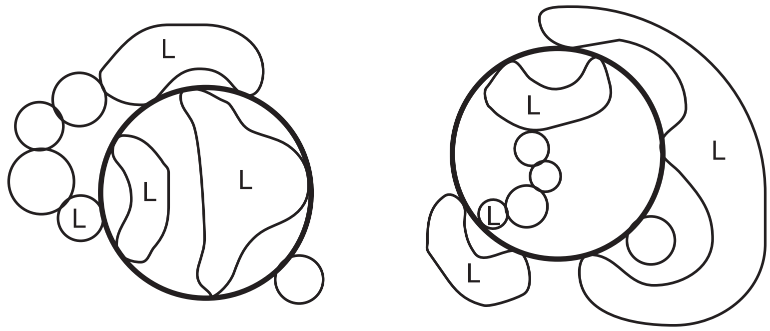

Figure 17: Two curves with the same Gauss diagram but different B-graphs. Notice that the tag on one curve is precisely the opposite of the other's tag with regard to the link L.

Pick a primitive link A_k containing elements of both $\{a_1,\ldots,a_i\}$ and $\{a_{i+1},\ldots,a_m\}$. There must be such an A_k, for if not the primitive links $A_1,\ldots,A_{n-1}$ can be partitioned into $\mathcal{A}_1,\mathcal{A}_2$ according to whether their double points lie in $\{a_1,\ldots,a_i\}$ or $\{a_{i+1},\ldots,a_m\}$. Since there is no way to satisfy inequality (11) using a subset of $\{a_1,\ldots,a_i\}$ and a subset of $\{a_{i+1},\ldots,a_m\}$, the link construction algorithm could then only choose primitive links from the set $\mathcal{A}_j$ containing A_1.

So, we have two primitive links A_k,A_n with $a,a'\in A_k$ and $b,b'\in A_n$ such that $a<b<a'<b'$. The edges of $A_1,\ldots,A_{n-1}$ have an assignment of inside and outside that satisfies the requirements of the lemma. Note that since A_k is a primitive link, all its edges have the same tag. If the edges of A_k and A_n are given the same tags, then the curve cannot possibly be embedded in the plane—somewhere it must "overlap" itself; see Figure 18. So if we give the edges of A_n the label opposite to those of A_k, this set of tags will also meet the requirements of the lemma. ■

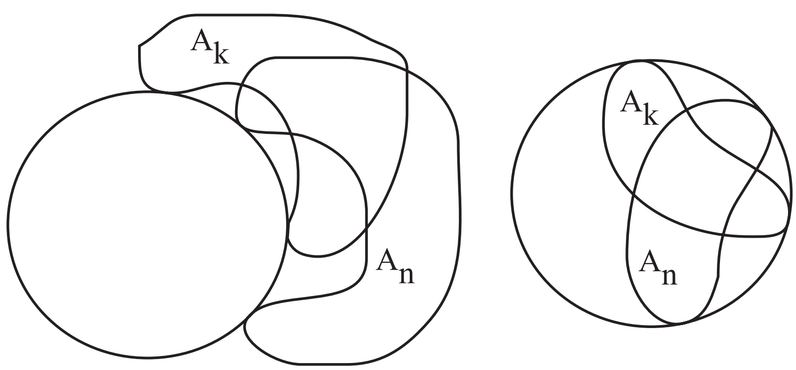

Figure 18: If the same tags are given to A_k and A_n, then two parts of the curve must overlap.

The following is a more technical lemma which says that for any two links, all the double points of one must lie between two adjacent double points of the other. See Figure 19.

Lemma 6 *Let A,B be links of B_G at s, and let $D(A)=\{a_1,\ldots,a_k\},D(B)=\{b_1,\ldots,b_\ell\}$. Then there* *cannot be $a_1, a_2 \in A, b_1, b_2 \in B$ such that $a_1 < b_1 < a_2 < b_2$ in the cyclic order of $D(A) \cup D(B)$ induced by s.*

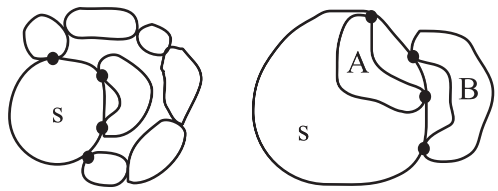

Figure 19: Two links at a vertex s, and vertex s with two sets of edges A and B whose double points overlap. In the second case, Lemma 6 implies that the edges connecting these two cycles to s must belong to the same link.

Proof. Suppose not. Then there exist $i, j \in \{1, \ldots, k\}$ such that

$$a_i < b_1, \ldots, b_k < a_j < b_{k+1}, \ldots, b_\ell.$$

Now there must exist a primitive link containing some element of $\{b_1, \ldots, b_k\}$ and some element of $\{b_{k+1}, \ldots, b_\ell\}$, because otherwise these two sets would be separate links. Call two members of this primitive link b_c and b_m, with $b_1 \leq b_c \leq b_k$ and $b_{k+1} \leq b_m \leq b_\ell$. Then $a_i < b_c < a_j < b_m$ and so by the construction defined above, A must contain the primitive link that includes b_c and b_m, and we have reached a contradiction. ■

Definition 9 *A vertex $s_1 \in B_G$* **contains** *a vertex $s_2 \in B_G$ if there is a path from s_1 to s_2 that begins with a directed edge and follows the direction of each edge it travels through.*

Again, the motivation for the above definition is to establish a correspondence between a curve and its directed B-graph. Vertices on B-graphs contain other vertices in precisely the same way that the Seifert cycles they correspond to contain each other.

Definition 10 *Let K be a curve, E be a set of edges containing a vertex $s = \{d_1, ..., d_n\}$, and let $C = \{x \in V(B_G)$: there exists a simple path from s to x starting with an edge of $E\}$. Then*

$$A(E) = \{\overrightarrow{d_i d_j} : \text{the part of } K \text{ corresponding to } \overrightarrow{d_j d_i} \cup [\bigcup_{C \in C} C] \text{ separates } \overrightarrow{d_i d_j} \text{ from the unbounded component of } \mathbb{R}^2 - K.\}.$$

Intuitively, $A(E)$ is the set of arcs of E which are completely surrounded by cycles of C. See Figure 20. Note that if E is a set of edges such that $D(E) = d_{\alpha(1)}, \ldots, d_{\alpha(k)}$, then $A(E)$ contains an arc covered by $\overrightarrow{d_{\alpha(i)}, d_{\alpha(i+1)}}$ if and only if it contains $\overrightarrow{d_{\alpha(i)}, d_{\alpha(i+1)}}$ itself. In other words, the arcs covered by E can be written as a union of arcs of the form $\overrightarrow{d_{\alpha(i)}, d_{\alpha(i+1)}}$.

Definition 11 *Let $s \in V(B_G)$ be the S-cycle $\{d_1, \ldots, d_n\}$. A set E of edges containing s is called* **complete** *if*

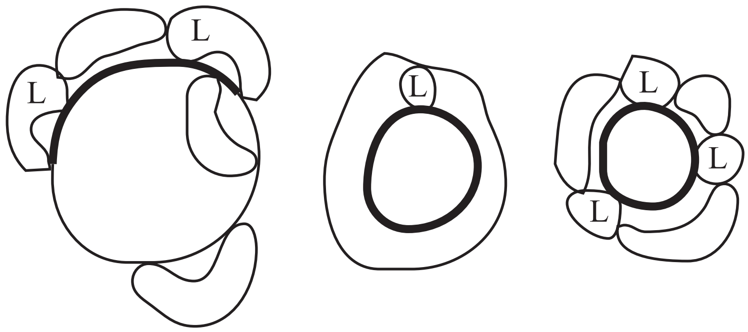

Figure 20: Examples of Definition 10. In each case, the arcs that make up $A(L)$ are highlighted.

1. *for every link L, $E \cap L \neq \emptyset$ implies that $L \subset E$,*
2. *there exist $a, b \in \{1, \ldots, n\}$ such that $D(E) = \{d_a, \cdots, d_b\}$, and*
3. *for each arc $\overrightarrow{d_i d_j} \in A(E), d_i < d_j$ in the order of $d_a < \cdots < d_b$.*

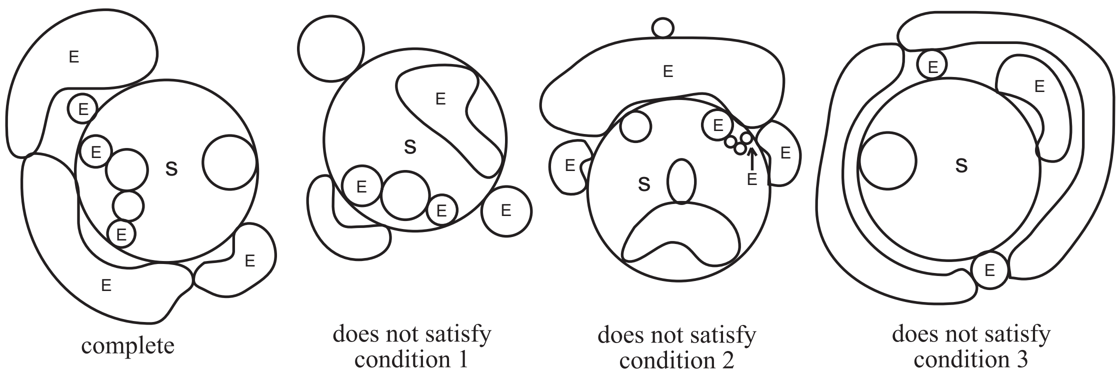

Figure 21: Examples of Definition 11; labeled cycles are connected to s by edges of E.

Definition 11 completely characterizes the conditions of a flip on a curve in terms of its directed B-graph; that is, a complete set of edges represents a portion of a curve that can be flipped according to Definition 2. We will prove this in Lemma 10, but first we need to establish some additional results.

Definition 12 *Let L be a link, and assume $D(L) = \{d_{\alpha(1)}, \ldots, d_{\alpha(k)}\}$. Choose i from $\{1, \ldots, k\}$ such that $E(\{d_{\alpha(i)}, d_{\alpha(i)+1}, d_{\alpha(i)+2}, \ldots, d_{\alpha(i-1)}\})$ is a complete set of edges containing s. Then we say that $E(\{d_{\alpha(i)}, d_{\alpha(i)+1}, \ldots, d_{\alpha(i-1)}\})$ is a* **completion** *of L, and we denote it $C(L)$.*

Definition 13 *A set E of edges containing a vertex s is called* **bad** *if $A(E)$ contains every arc between double points of s. See Figure 22.*

The motivation for this definition is that a link which is bad cannot be flipped. This is formalized in Lemma 8. In some cases we will be able to get around this by showing that we do not want to flip the link, but to deal with other cases we need to perform an operation that makes the link flippable. This operation is described in Lemma 9.

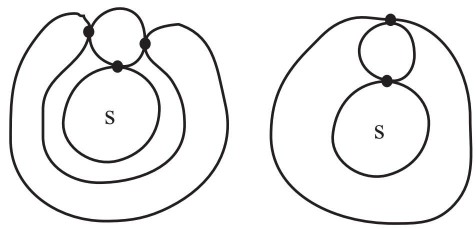

Figure 22: Two examples of a vertex s and a vertex connected to s by a bad edge. The last two parts of Figure 20 also show bad sets of edges.

The following lemma is instrumental in proving Lemma 11. Given that complete sets can be flipped, this will allow us to flip a link's completion, then flip back all the edges between the double points of the link so that only the edges of the link itself are changed.

Lemma 7 *Let L be a link, and let $C(L)$ be a completion of L. Assume that $s = \{d_1, \ldots, d_n\}$ and $D(L) = \{d_{\alpha(1)}, \ldots, d_{\alpha(k)}\}$ is in the order induced by that given to D(C(L)) by condition 2 of the definition of completeness. Then it follows that for every $i \in \{1, \ldots, k-1\}$, $D_{\alpha(i)} := E(\{d_{\alpha(i)+1}, \ldots, d_{\alpha(i+1)-1}\})$ is a complete set of edges.*

Proof. Let $C = E(D_{\alpha(i)})$. Lemma 6 implies that no edge can contain some double points in $D_{\alpha(i)}$ and some not in $D_{\alpha(i)}$, so $D(C) = D_{\alpha(i)}$. Also, if A is a link such that $A \cap C \neq \emptyset$, Lemma 6 also implies that $A \subset C$. Now condition 3 of the completeness of $C(L)$ implies that no set of Seifert cycles contained in the edges of C can enclose the arc of $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}}$. But the arc $\overrightarrow{d_{\alpha(i+1)-1}d_{\alpha(i)+1}}$ contains the arc $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}}$ which means that the arc $\overrightarrow{d_{\alpha(i+1)-1}d_{\alpha(i)+1}}$ cannot be in $A(C)$, which in turn implies that C satisfies condition 3 of Definition 11. ■

Lemma 8 *If L has no completion, then L is bad.*

Proof. We will prove the contrapositive. Assume that L is not bad. Then without loss of generality, assume that $A(L)$ does not contain the arc $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}}$ where $D(L) = \{d_{\alpha(1)}, \ldots, d_{\alpha(k)}\}$. Furthermore, assume that either (1) no link $Y \neq L$ encloses any of the arcs of $\overrightarrow{d_{\alpha(i)}d_{\alpha(j)}}$, or (2) $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}}$ contains the double points of such a link. In case (2), notice that the link in question must enclose the entire arc $\overrightarrow{d_{\alpha(1)}d_{\alpha(k)}}$.

Let $C = E(\{d_{\alpha(i)}, d_{\alpha(i)+1}), \ldots, d_{\alpha(i-1)}\})$. We claim that C is complete, with $\{d_{\alpha(i)}, d_{\alpha(i)+1}, \ldots, d_{\alpha(i-1)}\}$ in the order of condition 2. Lemma 6 implies that any link which has double points on the arc $\overrightarrow{d_{\alpha(k)+1}d_{\alpha(1)-1}}$ has only such double points, and is therefore contained in C, so C satisifies condition 1 of Definition 11; condition 2 follows from the construction.

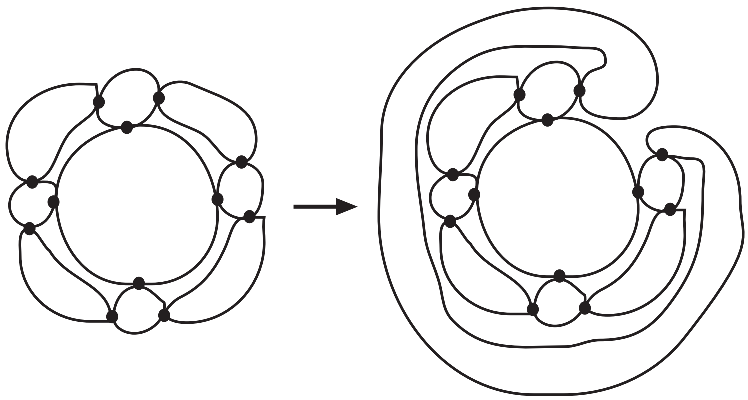

Figure 23: An example of modifying a link so it no longer completely surrounds a Seifert cycle.

Assume condition 3 is not satisfied. Then there must be a link $X \subset C$ such that $A(X)$ contains an arc covered by $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}}$. Since $D(X) \subset \{d_{\alpha(1)}, d_{\alpha(1)+1}, \ldots, d_{\alpha(i-1)}\}$, this implies that $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}} \in A(L)$. If (1) holds, then this is a contradiction.

Otherwise, assume that the double points of X lie on the arc $\overrightarrow{d_{\alpha(i)}d_{\alpha(i+1)}}$. (2) tells us that there exists a link Y whose double points lie on the arc $\overrightarrow{d_{\alpha(k)}d_{\alpha(1)}}$ which encloses some arc between two double points of L. As mentioned before, since the double points of Y lie on $\overrightarrow{d_{\alpha(k)+1}d_{\alpha(1)-1}}$, this is equivalent to saying that Y encloses the entire arc $\overrightarrow{d_{\alpha(1)}d_{\alpha(k)}}$, which implies that it encloses $\overrightarrow{d_{\alpha(i)}d_{\alpha(i+1)}}$.

We now have two links, each enclosing the arcs on which the other's double points lie. This is clearly impossible (the curve would overlap in the same way as in part 1 of Figure 19), so C must satisfy condition 3 of the completeness definition. ■

Definition 14 *A set E of edges containing a vertex s of a directed B-graph B_G* **surrounds** *s if it is bad with regard to s and there exists no path from s to a vertex containing s which begins with an edge of E.*

Note that such a set of edges can surround s only if s is not contained by any other vertex. For any path from s to a vertex containing it must begin with an edge of E because E is bad. A vertex that is surrounded corresponds to a Seifert cycle at which we cannot flip; to overcome this difficulty, we prove the following lemma, which essentially modifies the curve without changing the B-graph. See Figure 23. Lemma 4 implies that such a modification preserves the Arnold invariants, including, of course, the defect.

Lemma 9 *Let s be a Seifert cycle of a curve K, and assume there is a link L of edges containing s which surrounds s. Then it is possible to modify K so that L no longer surrounds s without altering the directed B-graph.*

Proof. The idea is as follows. Choose a sequence $R_0, c_0, \ldots, R_n, c_n$ of Seifert cycles c_i followed by components R_i of $\mathbb{R}^2 - K$ such that

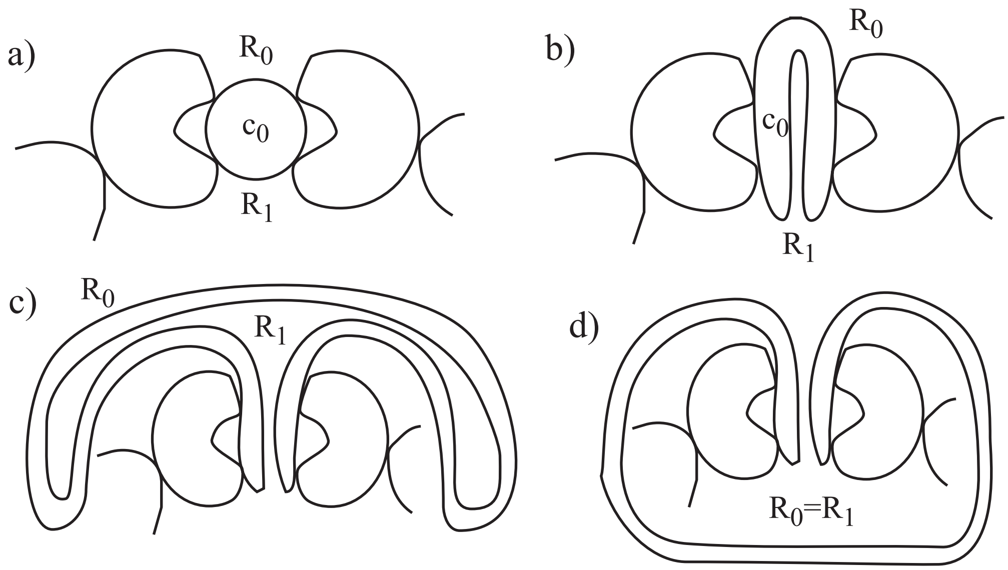

Figure 24: Merging two regions in a way that does not change the Arnold invariants.

1. c_n is s and R_0 is the unbounded component of $\mathbb{R}^2 - K$,
2. $R_i \neq R_j$ and $c_i \neq c_j$ for $i \neq j$,
3. R_i borders c_j iff $j = i$ or $j = i-1$,
4. no R_i is contained in a Seifert cycle, and no c_i is contained in some other Seifert cycle.

Such a sequence certainly exists. Simply choose a point on s, draw a ray to the unbounded component which does not intersect s more than once and does not go through any double points of K. Then read off each c_i and R_i that the ray intersects, possibly reducing afterwards to satisfy conditions 2 and 3.

To modify the link, we begin with c_0. We push the arcs of c_0 bordering R_0 and R_1 into R_0, and continue pushing c_0 into R_0 until c_0 assumes the shape shown in part c) of Figure 24. Then we reflect the upper handle of c_0 downwards around the rest of the curve to produce part d) of Figure 24. This last step merges the regions R_0 and R_1 without affecting any of the connections of the Seifert cycles of K, and therefore without changing the directed B-graph of K.

To be more precise, the regions R_0, R_1 determine arcs α_0, α_1 of c_0. If $A_0 = \{$ arcs on c_0 which bound $R_0\}$ and $A_1 = \{$ arcs on c_0 which bound $R_1\}$, we cannot have arcs of A_0 interspersed with those of A_1, so α_0, α_1 will be arcs if defined as

$$\alpha_i = [\bigcup A_i] \cup \{ \text{ points } p \in c_0 : \text{any path along } c_0 \text{ starting at } p \text{ must intersect an arc of } A_i \text{ before one of } A_{i+1}\}$$

We claim that there is no path of Seifert cycles (not passing through s) starting with a double point in the interior of α_i and ending with a double point of s not contained in α_i. But this is fairly elementary—any such path would cut R_i in two.

So, the Seifert cycles reachable from α_0, α_1 are in a sense separate from the rest of the cycles of K, and we can drag and reflect α_0 and α_1 as pictured.

After applying the process to c_0, c_1 will border R_0, and the resulting sequence $R_0, c_1, R_2, c_2, \ldots, R_n, c_n$ will satisfy all the conditions enumerated above. Repeating this process we will end with R_0, c_n satisfying all the conditions, which implies that $c_n = s$ borders the unbounded component of $\mathbb{R}^2 - K$, which in turn implies that L no longer surrounds s. ■

Definition 15 *Let $s \in V(B_G)$ and let E be a complete set of edges at s. Then the* **flip** *of E at s is a redirection of the edges of E such that an edge directed away from s becomes undirected, and an undirected edge is directed away from s.*

Note that if any of the edges in E are directed toward s, a flip is impossible. The flip of a B-graph is defined to correspond to the flip of a curve, which preserves defect. The next lemma proves that correspondence, and then Lemma 11 shows that each link can be flipped individually to have the directions we require.

Lemma 10 *Assume that B_G is directed to correspond to a curve K. Let B'_G be obtained from B_G by a flip of a complete set E at a vertex s. Then there is a curve K' with directed B-graph B'_G, and K' can be obtained from K by a flip as described in Definition 2.*

Proof. Suppose all the conditions of the B-graph flip (Definition 15) are met. We show that the conditions of the curve flip (Definition 2) are met, and that this flip results in a curve whose B-graph is B'_G.

Suppose that $D(E) = \{d_a, \ldots, d_b\}$ is the order given by condition 2 of the completeness of E. Choose flip points p, q on the Seifert cycle s of K with $d_b < p < d_{b+1}, d_{a-1} < q < d_a$ which separate the double points of s into those in $D(E)$ and those not in $D(E)$. Condition 2 of the completeness of E guarantees that such a choice is possible. We claim that these flip points separate K into two disjoint pieces.

Now, the flip points clearly separate the Seifert cycle s into two arcs, A_1 and A_2, and a path along the curve from one piece to the other would appear in the Seifert decomposition as a sequence of Seifert cycles, each connected to the next, which starts with a cycle s_1 connected to s by a double point in A_1 and ends with a cycle s_2 connected to s by a double point in A_2. Reducing if necessary by eliminating segments of the sequence which start and end on the same arc A_i, we obtain such a sequence of Seifert cycles which does not contain s. However, such a sequence implies that s_1 and s_2 lie in a primitive link, which is a contradiction by the completeness of E; see Figure 25. So, there can be no path along K from A_1 to A_2, implying the flip points do actually separate K.

Next, the two flip points cannot be contained in a region bounded by the part of the curve being flipped (that containing the double points of $D(E)$), for condition 3 of the completeness of E guarantees that all arcs $\overrightarrow{d_i d_j}$ of $A(E)$ satisfy $d_a < d_i < d_j < d_b$, and thus cannot contain the flip points, which satisfy $d_b < p, q < d_a$.

By Lemma 5, the links that make up E have a consistent tag of inside or outside on their edges that corresponds to the curve K and the directions of the B-graph B_G. Doing a curve flip at our two flip points results in a new curve, call it K'. The reflection inherent in the flip from K to K' reverses whether a Seifert cycle is inside or outside in precisely the same way that the B-graph flip reverses directed edges. Thus, B'_G corresponds to K'. ■

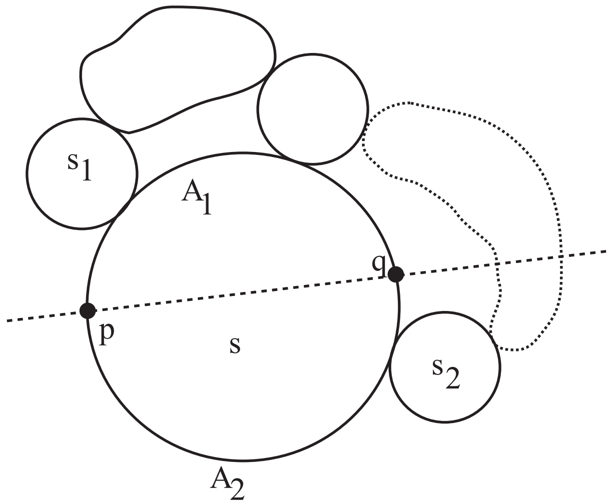

Figure 25: Flip points p and q separate the cycle s into pieces A_1 and A_2; it must be true that either p and q separate the entire curve into disjoint pieces or that there is a connection of Seifert cycles from s_1 to s_2, which means they are in a primitive link and thus p and q do not determine a complete set of edges.

Lemma 11 *Let $s \in V(B_G)$, L be a link at s, and $C(L)$ be a completion of L. Then there is a sequence of flips of B_G at s which*

1. *leaves edges not in L unchanged and*
2. *for edges in L, exchanges undirected edges with edges directed away from s and vice versa.*

Proof. Flip $C(L)$; all the edges in $C(L)$ including those in L are redirected according to Definition 15. Let $D(C(L)) = \{d_a, \dots, d_b\}$ in the order given in the definition of completeness, and let $D(L) = \{d_{\alpha(1)}, \dots, d_{\alpha(k)}\}$ be ordered in the order induced by $D(C(L))$. For any $i \neq k$, the set $D_{\alpha(i)}$ of edges determined by the double points between $d_{\alpha(i)}$ and $d_{\alpha(i+1)}$ is complete by Lemma 7, implying that we can flip $D_{\alpha(i)}$ so that its edges have their original directions. Flipping the edges of all such $D_{\alpha(i)}$ back to their original directions results in a B-graph which differs from the original only in that the directions of the edges of L have been modified as required in the statement of the lemma. ■

Lemma 12 *Given two directed B-graphs B_G and B'_G corresponding to curves K and K' with Gauss diagram G, there exists a sequence of flips on B_G which redirects the edges of B_G to match those of B'_G.*

Proof. Note that if B_G and B'_G correspond to the same Gauss diagram, they have the same graph structure and differ only by the directions on the edges. Consider the Seifert decomposition of K.

Select an exterior Seifert cycle c_0 of K. We proceed by changing the directions of the edges of B_G containing c_0 to match those of B'_G, then continuing along paths from c_0 changing the directions of the edges in a similar way as we visit each Seifert cycle of K.

This algorithm relies on the fact that the direction of an edge is determined by the inside/outside tags given to it by each of the two vertices it contains. Thus we can change the directions of the edges of B_G to match those of B'_G by at each vertex changing certain inside edges to outside and vice versa. Moreover, we show that if we want to change the inside/outside tag of an edge with respect to the current vertex, we can always do so with a flip at that vertex.

A problem could arise if we wish to switch the tag of an edge directed towards the vertex in consideration. In this case the link containing the edge is bad, and cannot be flipped. Note that in this case the edge is given an inside tag by the current vertex c_i and an outside tag by the other vertex c_j it contains. Certainly, either c_j contained c_i in the original graph B_G, or at some point during the algorithm we modified B_G so that c_j contains c_i. In both cases we have already visited c_j. For in the first case, considering the actual Seifert cycles corresponding to the graph B_G, we see that to reach c_i from c_0 we must pass through any cycle which contains c_i.

The fact that we have already visited the vertex c_j implies that the edge's inside/outside tag with respect to c_j matches that in B'_G. However, there is no direction that can be given to an edge such that the tag with respect to each of the vertices it contains are inside. So, in B'_G, the edge's tag with respect to c_i must be outside coinciding with the direction it currently has.

The initial case of our algorithm is simple. Let L_0 be a link of edges containing c_0. If it has a completion, we can toggle the tags of all edges in L_0 to match those of B'_G without changing the tags of any other edges in B_G. If L_0 has no completion, then since no vertex in B_G contains c_0 we can apply Lemma 9 and then perform the same process. Doing this with each link of edges containing c_0, we modify B_G so that all the directions of the edges containing c_0 match the corresponding directions in B'_G.

Inductively, assume that for each previously visited vertex, the inside/outside tags with respect to that vertex of all its edges match those of the corresponding edge in B'_G. Let L be a link of edges containing the current vertex c_i. If L has a completion, we can flip it (or not) to match the tags on B'_G. Assume L has no completion. Then Lemma 8 tells us that L is bad. If there exists a path from c_i to some vertex c_j containing c_i which starts at an edge $e \in L$, then the inductive hypothesis tells us that c_j contains c_i in B'_G, and since such a path connecting c_i to c_j also exists in B'_G, e is outside with respect to c_i in B'_G. And since to begin a simple path from c_i to c_j, e must be outside with respect to c_i in B_G, there is no need to flip L.

On the other hand, if there is no such path from c_i to c_j, then c_i is exterior and L must surround c_i in the sense of Definition 14. Then we can apply Lemma 9 so that L has a good completion and proceed as before. ■

Theorem 5 *All curves corresponding to a given Gauss diagram have the same defect.*

Proof. Let K and K' be curves with a Gauss diagram G and directed B-graphs B_G and B'_G. By Lemma 12 B_G can be transformed into B'_G by a sequence of flips. Applying the corresponding sequence of curve flips and Lemma 9 modifications to K, we produce a curve K_1 with the same defect as K and the directed B-graph of K'. Finally, by Lemma 4, since K_1 and K' have the same

Gauss diagram and the same directed B-graph, they have the same Arnold invariants, and therefore the same defect. ∎

5 Spherical curves

So far in this paper we have focused on plane curves, immersions of S^1 into $\mathbb{R}^2$. In terms of defect, though, it is also instructive to look at spherical curves, immersions of S^1 into S^2. A spherical curve may be represented by its stereographic projections into the plane; different planar projection of the same spherical curve K can be obtained by locating the north pole inside different regions of $S^1 - K$. However, the planar representatives of a spherical curve do not necessarily have the same invariants. In fact, the Whitney Theorem 1 does not hold for spherical immersions. In the following section, we characterize the path components in the space of spherical immersions and show that although St, J^+ and J^- are not spherical invariants, the defect is.

Definition 16 *The* **parity** *of a spherical curve is the parity of the indices of its planar representatives.*

Since all planar representatives of a spherical curve have the same Gauss diagram, the following theorem shows that parity is well defined.

Lemma 13 *The Whitney indices of plane curves corresponding to the same Gauss diagram have the same parity.*

Proof. From equation (2), the Whitney index of a plane curve is the sum of the indices of its Seifert cycles. That is, $\text{index}(C) = (s-k)(1) + (k)(-1) = s - 2k \equiv s \pmod 2$, where s is the number of Seifert cycles of C, k of which have negative indices. By Theorem 4, all curves with a given Gauss diagram have the same number of Seifert cycles, so the lemma follows. ∎

Lemma 14 *Parity is preserved by spherical homotopy.*

Proof. A spherical homotopy can be thought of as a sequence of spherical diffeomorphisms and planar homotopies. Any movement of the curve which does not pass through singularities such as self-tangencies and triple crossings can be simulated by a spherical diffeomorphism, while other parts of the homotopy can be isolated as $J^{\pm}$ and St moves on one of the curve's planar representatives. Arnold's three moves must preserve index, and from above spherical diffeomorphisms preserve index parity. ∎

Lemma 15 *Every spherical curve is homotopic to either the circle or the figure eight.*

Proof. Let K be a spherical curve. Choose a planar representative of K by locating the north pole inside a component of $S^1 - K$. This plane curve has a well-defined index and is homotopic to a base curve of the same index $(K_0, K_1, K_2, \ldots, K_n)$, by the Whitney theorem 1. Without loss of generality, we assume this index is non-negative.

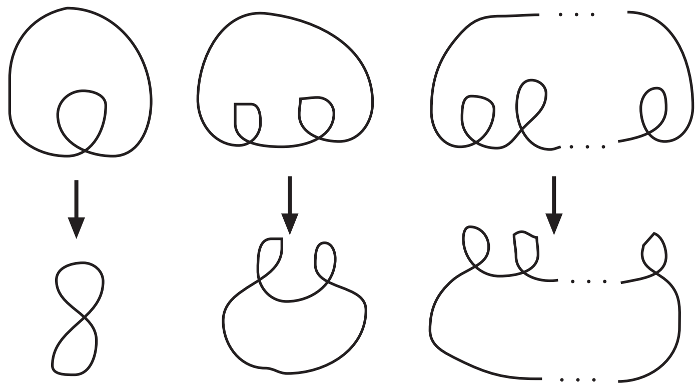

Figure 26: The plane curves $K_2, K_3, \ldots, K_n$ are equivalent on the sphere to plane curves with index $0, 1, \ldots (n-2)$.

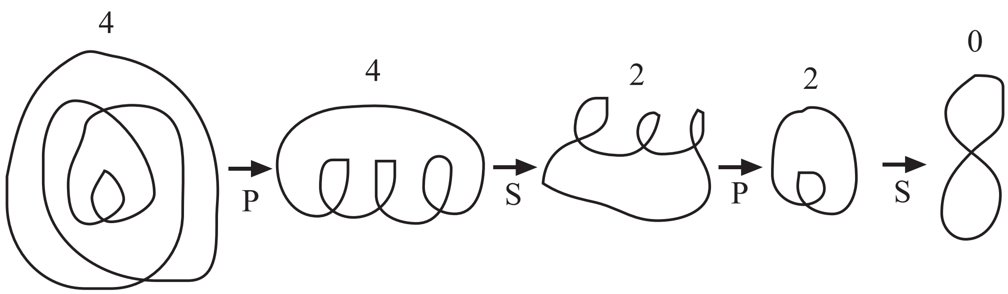

Figure 27: Homotopy of a spherical curve to the figure eight. The indexes of the planar representative shown are noted above each curve, spherical motions are represented with an S and planar homotopy is noted by P.

Now consider the base curve obtained by a plane homotopy of our representative. If the index is 0 or 1 then the theorem holds; otherwise, move the base curve along the sphere so that the north pole lies inside the large exterior loop. The new projection will have index 2 less than the previous projection, since the orientation of one loop was changed from positive to negative. See Figure 26.

Using a plane homotopy, deform this new curve into the base curve of the same index. Repeat the above process, at each step reducing the index of the planar representative by 2 and deforming the result into a base curve. The process will terminate with either a K_0 or a K_1. See Figure 27. ■

Since the circle and the figure eight have opposite parity, the preceding lemmas prove the following theorem.

Theorem 6 *Two spherical curves are homotopic in the space of immersions if and only if they have the same parity.*

Theorem 7 *Defect is an invariant of spherical curves.*

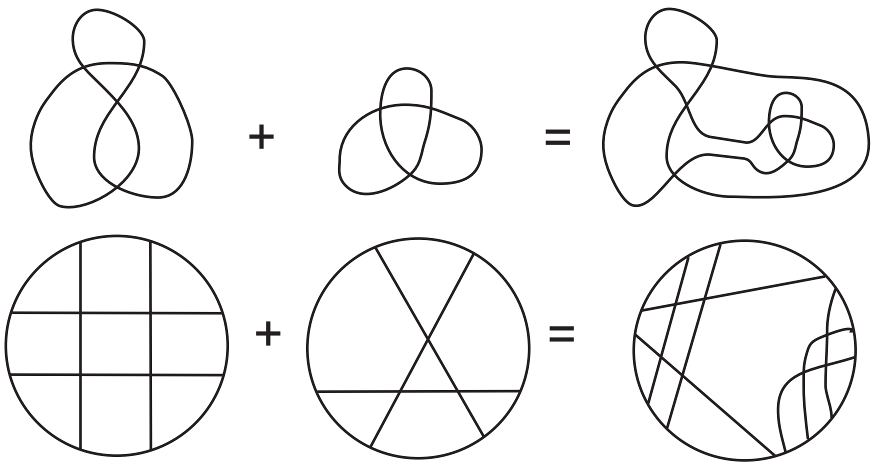

Figure 28: Gauss diagrams of curves and their interjected sum.

Proof. All planar representatives of a spherical curve have the same Gauss diagram, and all plane curves with a given Gauss diagram have the same defect. ■

6 Gauss diagrams and defect

Definition 17 *A Gauss diagram is* **bushy** *if the chords of the Gauss diagram partition into sets* $A_0, A_1, \ldots A_n$ *such that*

1. *no two chords in any* A_i *intersect,*
2. *given any* $i < j \in \{0, \ldots, n\}$*, there exists a subset* $B_i \subset A_i$ *such that every chord in* A_j *intersects exactly those chords of* A_i *which lie in* B_i*, and*
3. $|A_1| = |A_2| = \cdots |A_n| = 2.$

Note that A_0 may have any number of elements.

Conjecture 1 *A curve has defect 0 if and only if its Gauss diagram is bushy.*

Our conjecture seems plausible because bushiness is preserved by many defect-preserving operations. It is preserved by interjected sum—the Gauss diagram of an interjected sum is obtained by inserting the Gauss diagram of one curve into some arc of the Gauss diagram of the other. See Figure 28. More importantly, however, it is preserved by J^- moves and defect preserving combinations of J^+ and St, such as pulling certain sections of the curve through others. See Figure 29.

Proposition 3 *Bushiness is preserved by* J^- *perestroikas.*

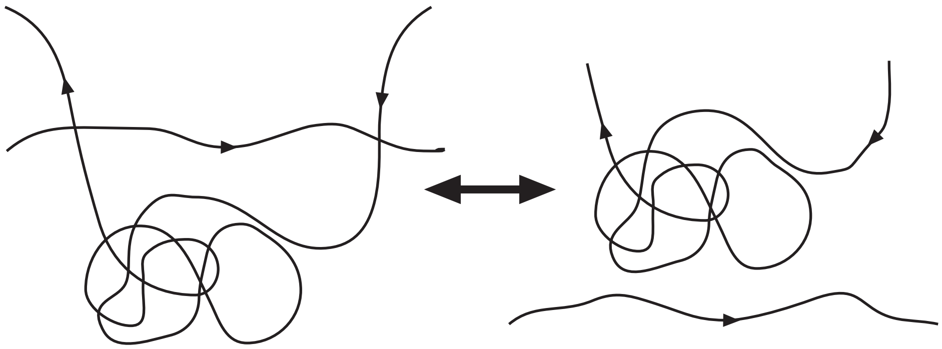

Figure 29: Pulling part of the curve through itself. The orientations must be as indicated.

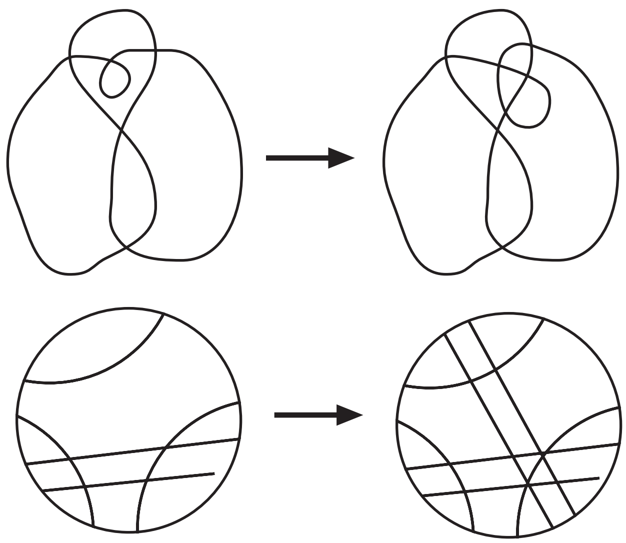

Figure 30: The effect of a J^- on a Gauss diagram.

Proof. Consider the effect of J^- perestroikas on the Gauss diagram. A J^- that reduces the number of double points simply deletes two adjacent parallel chords from the diagram. Clearly this will not affect its bushiness. A J^- that increases the number of double points adds two adjacent chords to the Gauss diagram. The chords are adjacent in the sense that there can be no endpoints of chords on the arcs of the diagram corresponding to the arcs of the curve created in the J^-. See Figure 30. Given a bushy partition $A_1, A_2, \ldots A_n$ of the diagram before the J^-, we create a new set A_{n+1} containing the new chords. For any $i \in \{0, \ldots, n\}$ the two new chords must intersect the same elements of A_i, so bushiness is preserved. ■

Many defect preserving operations on a curve can be regarded as a combination of a J^- and an interjected sum, both of which preserve bushiness. For example, pulling a section of the curve through an arc as in Figure 29 can be rephrased by decomposing the original curve with an interjected sum, executing a J^- perestroika and then interjecting again to create the final curve. See

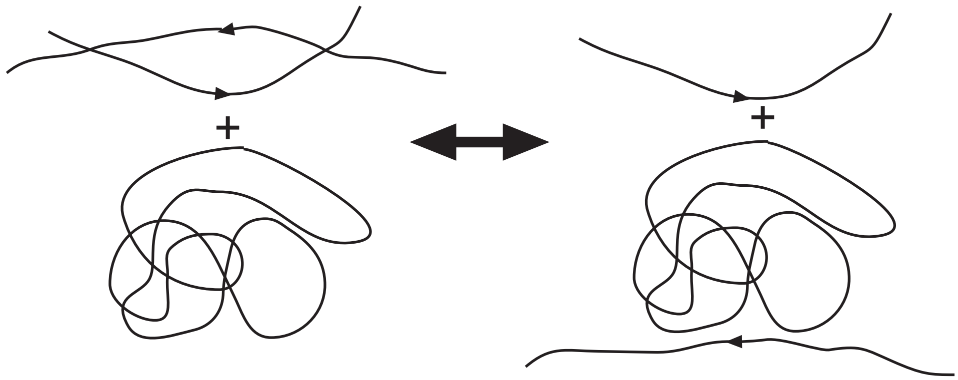

Figure 31: A pull-through expressed as a composition of J^- and interjected sum.

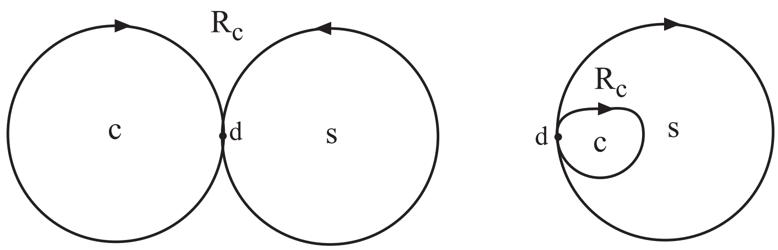

Figure 32: The index of a double point expressed as a single region index.

Figure 31. We believe that any defect 0 curve can be generated from treelike curves using only J^- operations and interjected sums, although we are not so bold as to state this in a formal conjecture.

7 Concluding remarks

Theorem 5 indicates that it might be possible to derive a formula for the defect of a curve which depends only on information given in the Gauss diagram. Upon examination, we see that much of the information used in the Luo and Shumakovitch formulas is the same. In particular, the t-function used in Luo makes use of the same region indices that are used in the summation in Shumakovich. To see this relationship, we consider the Shumakovich formula in terms of the Seifert decomposition of a curve. For a given double point, we see that (using the notation of equation 4) if d connects the Seifert cycles c and s,

$$\mathrm{ind}_K(d) = \begin{cases} \mathrm{ind}_K(R_c) & \text{if c is inside s,} \\ \mathrm{ind}_K(R_c) = \mathrm{ind}_K(R_c) & \text{if d is part of an exterior connection.} \end{cases} \tag{12}$$

For a Seifert cycle c, let E_c be the set of double points on c involved in exterior connections and let I_c be the set of double points on c connecting s to a cycle containing s. Then (12) implies that

$$St = \sum_{c \in K} \left[\text{ind}_K(R_c) \Big(\sum_{d \in E_c} \varepsilon_d(x) + 2 \sum_{d \in I_c} \varepsilon_d(x) \Big) \right] + \text{ind}_K(x)^2 - \frac{1}{4}. \tag{13}$$

Also, we can use the definition of the t-function to write Luo's formula as

$$J^+ = 1 + n - s - 2 \sum_{c \in K} \Big(\text{index}(c)\ \text{ind}_K(R_c) \Big). \tag{14}$$

Now, n and s in Luo's formula depend only on the Gauss diagram of the curve, and x in (13) can be chosen so that $\text{ind}_K(x)^2 - 1/4 = 0$. The hope is then that when we combine the summations in the two formulas, the information in each that cannot be calculated from the Gauss diagram will dissapear.

Initially, we had hoped to prove Theorem 5 by showing that two curves with the same Gauss diagram differ only by a sequence of flips. Although we were unable to prove it, we have not found a counterexample to this stronger result. The configuration of Seifert cycles in Lemma 9 is very complicated, and might not be possible in an actual curve. Additionally, instead of using Lemma 4 we might be able to flip the structure of a curve not represented by its B-graph.

References

[1] F. Aicardi. Tree-like curves. *Advances in Soviet Mathematics*, 21:1–31, 1994.

[2] V. I. Arnold. Plane curves, their invariants, perestroikas, and classifications. *Advances in Soviet Mathematics*, 21:33–39, 1994.

[3] V. I. Arnold. *Topological invariants of plane curves and caustics*, volume 5 of *University Lecture Series*. American Math Society, Providence, 1994.

[4] V. I. Arnold. Remarks on the enumeration of plane curves. In *Topology of real algebraic varieites and related topics*, volume 173 of *American Math Society Translation Series 2*, pages 17–32. American Math Society, Providence, 1996.

[5] Chenghui Luo. *Proof of Arnold's conjectures about plane curves*. PhD thesis, Brown University, 1997.

[6] Michael Polyak. New whitney-type formulas for plane curves. In *Advances in mathematical sciences*, volume 190 of *American Math Society Translation Series 2*, pages 103–111. American Math Society, Providence, 1999.

[7] Boris Shapiro. Tree-like curves and their number of inflection points. In *Advances in mathematical sciences*, volume 190 of *American Math Society Translation Series 2*, pages 113–129. American Math Society, Providence, 1999.

[8] A. N. Shumakovich. Explicit formulas for strangeness of plane curves. *St. Petersburg Mathematics Journal*, 7:445–472, 1996.

[9] H. Whitney. On regular closed curves in the plane. *Compositio Mathematica*, 4:276–284, 1937.