PLANE AND SPHERICAL CURVES: AN INVESTIGATION OF THEIR INVARIANTS

MICHAEL CHMUTOV, THOMAS HULSE, ANDREW LUM, AND PETER ROWELL

ADVISOR: JUHA POHJANPELTO
OREGON STATE UNIVERSITY

ABSTRACT. In plane and spherical curve geometry, the study of invariants is of particular interest as they help to identify relationships between curves and motivate a practical system of classification. We begin this work by dealing primarily with signed Gauss diagrams as an invariant of spherical curves, identifying original classes of them and how they are of particular use to us, and discussing an effective means for their construction and enumeration. From here we elaborate on a new structure for Gauss diagrams that form a complete invariant for plane curves, how it naturally arises, and how it can be used to find other invariants with the aid of Seifert decompositions and functions defined by Polyak. We then proceed to study symmetries of plane curves, what effects they have on their corresponding spherical curves and structured Gauss diagrams, and how they can be identified through matrix representations of their Gauss words. To conclude, we analyze one of Arnold's generally overlooked methods for plane curve decomposition and show that this gives rise to a ribbon graph representation and thus a polynomial due to Bollobás and Riordan; we study this polynomial and its specializations as an invariant of plane and spherical curves.

Date: August, 2006.
This work was done during the Summer 2006 REU program in Mathematics at Oregon State University.

CONTENTS

Part 1. **Introduction**

A *generic spherical curve* is defined to be a smooth immersion of the circle into the surface of the sphere where the finite number of self-intersections are transverse double points, so that self-tangency and triple point singularities are not allowed. A *generic plane curve* is defined similarly, but with an immersion into the plane. In this paper, we use the terms *spherical curve* and *plane curve* to denote the generic variety. A *region* of a curve is a maximal path-connected subset of the complement and an *edge* is an uncrossed strand between double points. Every plane curve can be uniquely lifted to a spherical curve through stereographic projection, and every spherical curve can be projected into the plane via the assignment of the north pole to some region, which we call *unbounded*. We consider plane or spherical curves to be equivalent if they are diffeomorphic; different projections of any given spherical curve typically yield inequivalent plane curves.

A *chord diagram* is a containing circle with a set of chords with distinct endpoints. Equivalence of chord diagrams is defined up to homeomorphism of the circle such that the order of the ends of chords about the circumference is unchanged. A *Gauss diagram* is a chord diagram associated with the preimage of a spherical or plane curve where each chord uniquely corresponds to a double point. Most chord diagrams are not Gauss diagrams; a chord diagram is called *planar* if it is also a Gauss diagram, and the problem of determining which chord diagrams are Gauss diagrams is commonly referred to as the *planarity problem* of chord diagrams [Ca]. An orientation and a base point on a curve and its corresponding Gauss diagram gives rise to an ordering of the directions of the outgoing arrows $(1, 2)$ at each singularity and so we can sign the double point and its corresponding chord positively (resp. negatively) if $(2, 1)$ orients the plane positively (resp. negatively), which defines a *signed* Gauss diagram.

If we label the endpoints of every chord with a letter on an n-chord diagram, pick a base point and orientation on the containing circle and then traverse it, we define the *chord word* to be the set of $2n$ letters ordered by how we encounter the endpoints on the containing circle. Every Gauss diagram gives rise to a *Gauss word*, and as its construction is reversible the Gauss diagram and the Gauss word of a curve are equivalent invariants, so we can refer to letters and chords interchangeably. Given a signed Gauss diagram, we can traverse the circle and label positively (resp. negatively) the first point encountered for a positively (resp. negatively) signed chord and the subsequent point negatively (resp. positively) to construct a *signed Gauss word* (Fig. 1). A signed Gauss diagram and a signed Gauss word are also equivalent invariants. Cycles of the Gauss word and permutations of the letters are equivalent as they correspond to rotations of the Gauss diagram and relabellings of the chords, respectively. As we frequently suppress orientation when considering the concept of equivalent curves, we allow the reversing of order or sign for all of the letters to yield an equivalent Gauss word.

In knot theory the concept of prime and composite knots has been thoroughly investigated and is employed as an intuitive means of construction and classification. However, this treatment has not, apparently, been extended to generic plane and spherical curves. Arnold briefly alluded to an immersed curve analogue in his seminal work on the subject, [Ar1], defining reducible and irreducible plane curves and later describing them as the "building blocks for all other immersions"[Ar3] though he left the subject largely unexplored.

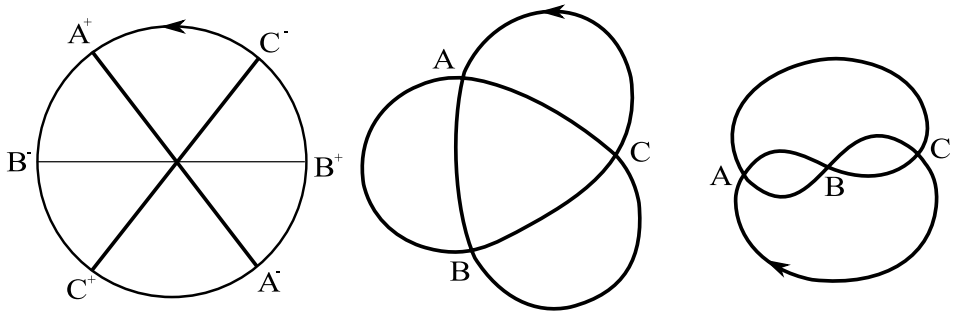


FIGURE 1. An example of a signed Gauss diagram (left) with Gauss word $(A^+,B^-,C^+,A^-,B^+,C^-)$ and two diffeomorphically inequivalent plane curves (middle-right) that correspond to the same spherical curve. The darker chords on the Gauss diagram indicate positive signing.

Carter [Ca] has shown that signed Gauss diagrams are a complete invariant for spherical curves and Polyak [Po] has demonstrated that methods for calculating invariants of plane curves could be devised from the signed Gauss diagram and the *Whitney index*. Inspired by this and knot theory, we define *prime* Gauss diagrams, in which the chords form a connected space, and show that such Gauss diagrams uniquely define a class of spherical curves, which we also call *prime*, that can be used to construct and classify all generic spherical curves.

Once this is established, we are faced with the difficult problem of constructing and classifying all such prime Gauss diagrams which is a simplification of the planarity problem for chord diagrams [Ca]. We give a systematic method for constructing all chord diagrams that could potentially be prime and planar by considering only chord diagrams in which the chords form a connected space and satisfy the even chord intersection criterion due to Gauss. The construction method ensures every such chord diagram is constructed and no equivalent chord diagrams are duplicated. By only considering less spurious cases, we are capable of running an algorithm due to Cairns and Elton [CE] that determines the planarity of chord diagrams in less time. Also importantly, the method enables one to catalog and classify all such chord diagrams in a specific and systematic manner; a method for the construction and classification of all prime Gauss diagrams is a special case of this. To this end, we introduce a complete invariant for chord diagrams termed the *primary fingerprint*, which will allow us to assign a complete numerical invariant to all prime spherical curves.

We will then focus a great deal of our attention to plane curves and a their classification. Arnold [Ar1] was aware that his invariants, along with the Whitney index, were insufficient for differentiating all inequivalent plane curves. Polyak [Po] tackled the problem of plane curve classification by demonstrating that n-th order invariants could be constructed by bracket functions on the signed Gauss diagram and the index of a plane curve, and successfully defined the Arnold invariants in terms of these functions. We review this function and these equations and explicitly construct a general n-th order invariant of our own, which we call the *n-th Zen Invariant*.

Polyak went so far as to claim that the signed Gauss diagram along with the index is a complete invariant for plane curves, though it is not difficult to construct any number of counterexamples to his assertion (Fig. 2). In this work, we construct and justify our own complete plane curve invariant. We define *plane structures* on signed Gauss diagrams corresponding to regions on spherical curves, and using the fact that a plane curve is defined by the choice of an unbounded region of a spherical curve we prove that these plane-structured, signed Gauss diagrams form a complete invariant. We also show that they can be naturally identified by the signed Gauss word, and that many other important invariants can be explicitly deduced from them.

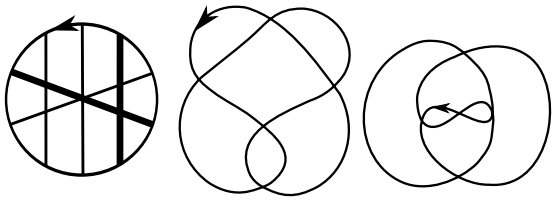


FIGURE 2. A based, oriented, and signed Gauss diagram (left) corresponding to two inequivalent plane curves with index zero (middle-right). Darked chords on the Gauss diagram are positively signed.

In the signing of composite Gauss diagrams, it is important to account for symmetries when determining whether the corresponding curves are inequivalent. Along similar lines, we say that different regions of spherical curves are *symmetrically equivalent* if they define equivalent plane curves when they are chosen to be unbounded. We will discuss how such symmetries can be easily identified from the plane structure, particularly for prime spherical curves, with implications for the enumeration of plane curves.

We then describe how the type of symmetry possessed by a plane curve has implications for the set of plane curves obtained from its spherical ancestor. Arnold [Ar1] defined four symmetry transformations and began classifying plane curves according to whether or not a curve was invariant under one, three, or four of these transformations. We argue that the range of symmetries exhibited by plane curves is not limited to the four investigated by Arnold; we examine symmetry classes generated by orientation transformations and develop methods for identifying whether a curve belongs to one of these classes.

Finally, we focus on one of Arnold's less-studied works, [Ar2], in which he describes an interesting approach to the study of plane curves. By breaking a curve up into its "outside" contour and the remaining part - both of which are diffeomorphic to curves which are simpler than the original - and classifying the possible "outsides" and the remaining part, Arnold likely hoped to be able to classify, or at least count, all possible plane curves. Unfortunately this approach succeeded only in the simplest cases. But by this process, Arnold's work naturally leads to a ribbon graph structure on the curve, which, as Bollobás and Riordandescribed describe in [BR1] and [BR2], has an interesting three-variable polynomial invariant associated with it. We will study the properties of this invariant as it related to the original plane curve.

Part 2. **Prime Gauss Diagrams**

1. CONNECTED SUMS

In the next few sections, we will focus much of our attention on a very specific class of Gauss diagrams which we call the *prime* Gauss diagrams.

Definition 1.1. A Gauss diagram is *prime* if all of its chords form a connected and path-connected space. It is *composite* otherwise.

Since the number of chords is finite and one chord can only intersect another at a single point, connected and path-connected chord spaces are equivalent. Since its chords do not form a connected space, a composite Gauss diagram permits a natural partition on the space of its chords into *components*.

Definition 1.2. A *component* of a Gauss diagram is a maximal nonempty connected subset of the chords in a Gauss diagram. A prime Gauss diagram can be constructed for this component simply by removing the compliment of chords from the parent Gauss diagram (Fig. 3).

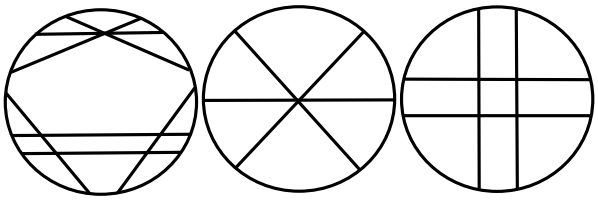


FIGURE 3. An example of a composite Gauss diagram (left) and its prime components (middle-right).

A prime Gauss diagram has at most one component. Any Gauss diagram with no components corresponds to the diffeomorphism class of the circle; both the Gauss diagram and the curve will be called *trivial* in this case. Conversely, a composite Gauss diagram has multiple components which are well-defined by maximality. As every component gives rise to a component Gauss diagram, and as every component Gauss diagram has only one component by construction, we will use these terms interchangeably.

Composite Gauss diagrams the concept of the *connected sum*, which we modify from Arnold's definition in [Ar3] for extension onto the sphere.

Definition 1.3. A *connected sum* of two generic spherical curves is defined by an embedding of a connecting bridge into the complement of the images of the two original immersions such that no double points are created or destroyed.

It is notable that Arnold's definition of the connected sum on the plane only allowed for a connecting bridge between the exteriors of plane curves, and so the lack of a well-defined exterior region on the sphere allows for a bridge between any pair of edges on two spherical curves. When on the subject of plane curves, we will be referring to Arnold's definition of connected sum.

Proposition 1.4. *The set of components of the Gauss diagram of a connected sum of two spherical curves is the union of the set of the components of the Gauss diagrams of each of the two curves.*

Proof. The Gauss diagram is the preimage of the spherical curve and so undivided arcs on the Gauss diagram correspond to edges on the spherical curve. Thus the connected sum of two spherical curves yields a curve with a Gauss diagram that is homeomorphic to the connected sum of the Gauss diagrams of the original curves. So chords from each of the original Gauss diagrams must be disjoint from each other, and as no chords were created or destroyed, the total set of components remains unchanged and are all contained in one Gauss diagram (Fig. 4). □

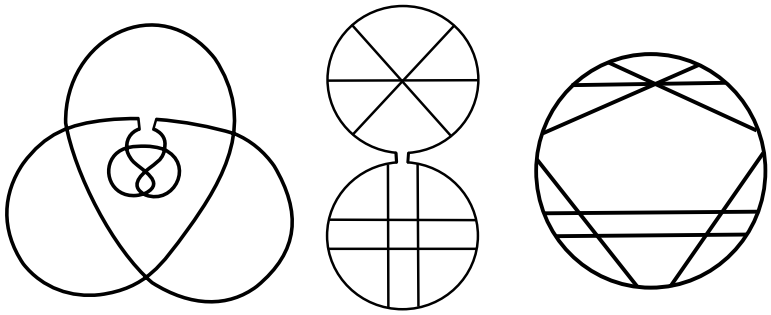


FIGURE 4. The connected sum of two spherical curves (left) can be interpreted as the connected sum of their preimages (middle) yielding a Gauss diagram whose total set of components is the union of the components from the ancestor Gauss diagrams (right).

Corollary 1.5. *If any spherical curve,* Γ*, corresponding to a prime Gauss diagram is represented as the connected sum of two spherical curves, one of them is trivial and the other is diffeomorphic to* Γ.

Proof. From proposition 1.4, components are additive under connected sum, so one of the Gauss diagrams has to have no components for the total number of components to be at most one, and thus it is trivial. Since the connected sum of any spherical curve with the trivial curve is itself, the other curve must be Γ. □

We can now state the first theorem that motivates the definition of prime Gauss diagrams.

Theorem 1.6. *A Gauss diagram is composite if and only if all corresponding spherical curves can be represented as the repeated connected sum of curves with component Gauss diagrams of the original.*

Before we proceed to prove this, we must examine some properties of Gauss diagrams.

Recall that the Gauss diagram of a spherical curve is a representation of its preimage, and thus each arc between two ends of chords on the circle, undivided by other chords corresponds to an edge on the spherical curve. We say two sets of chords are *adjacent* on the Gauss diagram if there is an undivided arc on the circle between the ends of at least one chord from each set. As each chord is affixed to the circle, we find that if a nonempty subset of chords in a Gauss diagram has no adjacent sets then it must be the entire set of chords, as all arcs on the circle must connect chords in the same set.

Suppose a Gauss diagram is composite, then the space of the chords can be partitioned into multiple components. We require the following lemma:

Lemma 1.7. *If a Gauss diagram is composite then there is at least one component that is connected to its compliment in the chord space by exactly two undivided arcs on the containing circle.*

Proof. Suppose a Gauss diagram has $n \geq 2$ components. Pick one component and define it to be the starting set, S_1, then pick a base point along an undivided arc connected to a chord in S_1. Now choose a direction and traverse the circle, moving from chord to chord along undivided arcs. The first new component encountered - that is the component containing the first encountered chord that is not in S_1 - must be adjacent to S_1, as being first requires that there is an arc between it and S_1. Take the union of this component and S_1 to create S_2 and then continue the path along the circle until we encounter the next component not in S_2. Repeat this process until S_{n-1} is constructed. We now have two disjoint chord sets, S_{n-1} and the remaining component, C. They must be adjacent and thus have at least one arc connecting one chord from each set. Now divide that arc so that the containing circle of the Gauss diagram is now homeomorphic to a line segment; if there existed no other arc between two chords in each set then one would travel from one end of the segment to the other and encounter at most one component along it, which is a contradiction. Thus there are at least two arcs connecting the two sets.

Suppose a third connecting arc existed, then, since the end of each chord corresponds to only two undivided arcs, there must be at least two points on the circle corresponding to endpoints of chords of C such that three undivided arcs connect S_{n-1} to these points. Since the chords in C form a connected space, we can make a path, $\mathcal{P}_0$, between the two noted points in C along its chords that divides the circle into two pieces and the enclosed disk into two regions, each region corresponding to each piece. If S_{n-1} is affixed to only one piece of the containing circle then the three connecting arcs must be on one side of $\mathcal{P}_0$, and thus another noted point from C must exist on that side. So another path, $\mathcal{P}_1$, can be made from one of the original noted points to the new one, dividing the containing circle into two pieces where S_{n-1} is affixed to both. Thus there must always exist some dividing path $\mathcal{P}$ in C through S_{n-1}. No part of S_{n-1} in one region of the divided disk can be connected to any part in the other region without having some chord in S_{n-1} intersect $\mathcal{P}$, violating the maximality condition on C. Thus $\mathcal{P}$ must divide S_{n-1} into two non-connected sets of components. In the $n = 2$ case this is impossible as $S_1 = S_{n-1}$ is a component and connected. In the $n > 2$ case, by construction, some part of S_{n-1} in one region must be adjacent to another part of S_{n-1} in the other region and this is impossible with $\mathcal{P}$ dividing every arc between the two regions. So we have a contradiction and the lemma is proven (Fig. 5). □

Now we can prove our theorem:

Proof of Theorem 1.6. If all corresponding spherical curves of some Gauss diagram can be represented as the connected sum of curves with component Gauss diagrams of the original, then the Gauss diagram is trivially composite as it has at least two component Gauss diagrams by proposition 1.4.

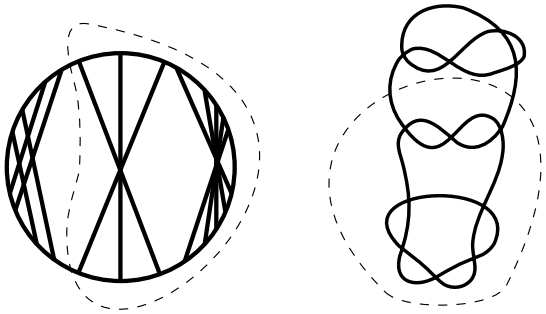


FIGURE 5. A composite Gauss diagram (left) has a chord space that can be partitioned such that one component is connected to the rest of the chord space by two undivided arcs. This corresponds to the potential for a connected sum (right).

Suppose a Gauss diagram, G, is composite with n components. Lemma 1.7 gives us that there is at least one component that is connected to the remaining set of chords by exactly two undivided arcs on the containing circle. Cut each of these arcs somewhere in the middle; this divides G into two disjoint regions. For any arbitrary corresponding spherical curve, Γ, that G is the preimage of, the cuts in G correspond to two split edges. Since these cuts divide G into two disjoint regions, G_1 and G_2, they must divide Γ into two disjoint regions, Γ_1 and Γ_2, for which G_1 and G_2 are the respective preimages. Furthermore, we know that Γ_2 was the path from the one free end of Γ_1 back to the other and intersected it nowhere else, thus between these points we can construct a connected path in the complement of Γ_1, in the space that Γ_2 used to occupy, devoid of double points. Thus we have created an edge closing Γ_1; this corresponds to joining the free ends of G_1 and effectively defines a Gauss diagram for Γ_1. The same can be done for Γ_2. We can easily see that Γ is the connected sum of the closed Γ_1 and Γ_2, as we can just recreate the connecting bridge between the two edges destroyed in the decomposition. Without loss of generality, we can say that G_1 corresponds to a component of G and thus Γ_1 has a component Gauss diagram of G. We can then repeat the process on G_2, which itself has $n-1$ remaining components until we've shown that every spherical curve corresponding to G can be represented as a connected sum of spherical curves with component Gauss diagrams of G. Thus the theorem is proven. □

Definition 1.8. A spherical curve is *composite* if it can be represented as the connected sum of nontrivial spherical curves. Otherwise, it is *prime*.

Corollary 1.9. *A spherical curve is prime if and only if its Gauss diagram is prime.*

The concept of prime and composite curves is not difficult to extend to plane curves, as we can still claim that a plane curve with a prime or composite Gauss diagram is prime or composite, respectfully. However, we see that the connected sum does not extend analogously to the case of plane curves, and thus we must give a planar analogue which we attribute to Barker and Biringer:

Definition 1.10. [BB] An *interjected sum* of two plane curves, Γ_1 and Γ_2, is formed by embedding Γ_2 into a region of Γ_1 and then embedding a connecting bridge into the complement of both, from some edge in Γ_1 to an exterior edge of Γ_2, such that no double points are created or destroyed.

Corollary 1.11. *A Gauss diagram is composite if and only if all corresponding plane curves can be represented as the interjected sum of curves with component Gauss diagrams of the original.*

Proof. Since a plane curve can be represented as an interjected sum of two other plane curves if and only if the spherical curve can be represented as a connected sum of other spherical curves, this is a restatement of theorem 1.6 for plane curve projections. □

We delve further into plane curve classification in section 8.

We can immediately see the practicality of using the definition of prime curves; they are easy to identify and they can be used to construct all spherical or plane curves through connected or interjected sums, respectively. Now, consider Arnold's definition of irreducible curves:

Definition 1.12. [Ar3] A circle immersion is called *reducible* if some double point cuts the image into two disjoint loops.

We argue preference over Arnold's concept of irreducibility as such a splitting double point will correspond to a component Gauss diagram comprised of only one chord. Thus all prime generic curves with more than two chords are irreducible. Furthermore, the connected sum of prime spherical curves - excluding the figure eight - will construct an irreducible spherical curve as no one-chord components are created. Thus, countably many of Arnold's irreducibles are composite by our definition. Conversely, consider any reducible spherical curve, Γ, where there exists a double point that separates the curve into two disjoint loops, C_1 and C_2, which we can regard as generic curves. If one of them, C_1 without loss of generality, is non-trivial (otherwise Γ is the figure eight) we make a cut along each edge connecting to this double point. We can easily close these disjoint curves to show that Γ was the connected sum of C_1 to a curve who's Gauss diagram must have at least one component corresponding to the separating double point. Components are additive under connected sums, thus the Gauss diagram of S has multiple components and so Γ is composite. Thus all primes but the figure eight are irreducible.

The connected sum in this context behaves notably similar to the composition of knots and, as an analogue, has a very interesting relationship, which we'll discuss in a later section. More immediately we focus on a bijection between prime Gauss diagrams and prime spherical curves.

2. The Signing Algorithm

In knot theory, a composite knot can be uniquely decomposed as a knot sum of prime knots. So given that a composite spherical curve can always be decomposed as the connected sum of prime curves, it is natural to ask whether this factorization is also unique. This is indeed the case and can be proven as a corollary of the following theorem.

Theorem 2.1. *Any signed chord on a Gauss diagram uniquely determines the sign of every chord that intersects it.*

On a prime Gauss diagram, every chord is connected to every other chord through a series of intersections. Thus this theorem gives the signing of the entire prime Gauss diagram by the signing of just one chord.

Remark 2.2. Suppose we have a signed Gauss diagram G and the corresponding spherical curve, Γ, then the opposite signing of G corresponds to the mirror image of Γ, and we treat these curves as equivalent on the unoriented sphere for the purposes of classification. Thus, with respect to generality, it does not matter whether any given chord is signed positively or negatively, what matters is the signing of the chord relative to the signing of the other chords in G, so we regard oppositely signed Gauss diagrams as having equivalent signings.

From this, theorem 2.1 gives rise to the following corollary:

Corollary 2.3. *A prime Gauss diagram admits only one signing.*

Recall that a signed Gauss diagram is a complete invariant for spherical curves [Ca], so that two inequivalent signings will define two different spherical curves. Thus a desired equivalent restatement of corollary 2.3 would be:

Corollary 2.3. (Restatement) *A prime Gauss diagram uniquely defines a prime spherical curve.*

Let's prove this theorem.

Proof of Theorem 2.1. Suppose G is a Gauss diagram and Γ is any spherical curve that corresponds to it. Consider a chord, B_G, on G which corresponds to a transverse double point, B_Γ, on Γ. As a preimage, the two arcs, G_1 and G_2, on the containing circle of G that begin just before B_G and end just after correspond to two separate paths, which we call the *loops* Γ_1 and Γ_2 respectively, on Γ that begin just before first crossing B_Γ and end just after the second crossing. Disregarding the shared tail ends, which can be made arbitrarily small, Γ_1 and Γ_2 can only intersect at the double points that correspond to the chords on G that are connected to both G_1 and G_2 - the chords that intersect B_G (Fig. 6).

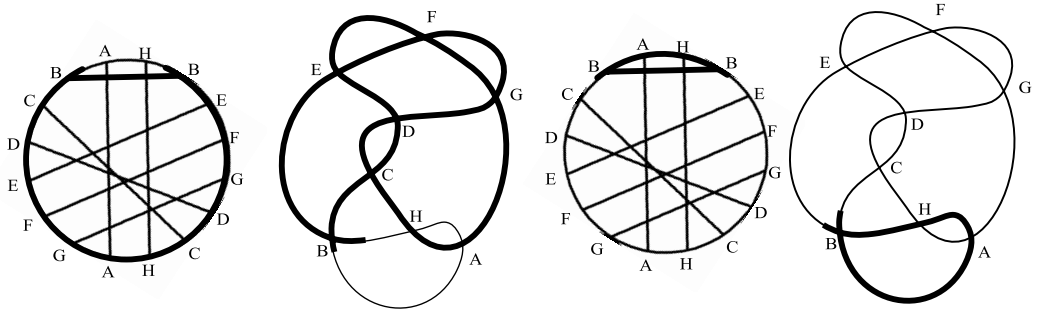


FIGURE 6. An example of a Gauss diagram, G (left) with darkened arc G_1 that corresponds to the darkened loop Γ_1 of the corresponding spherical curve, Γ (middle-left). Similarly the darkened arc G_2 (middle-right) corresponds to the darkened loop Γ_2. Note that the only non-trivial intersections between Γ_1 and Γ_2 correspond to chords that intersect B_G in G.

Suppose that the base point is placed at the start of G_1; traversing it would correspond to an oriented representation of Γ_1, orienting and ordering every edge. Note that without Γ_2, any double

point, Y_Γ, on Γ_1 corresponds to a self-intersection and so the corresponding chord, Y_G, must connect two points on G_1 and thus cannot intersect B_G. Take a checkerboard coloring of Γ_1, that is for each oriented edge the side (left or right) that a given color is on alternates with each subsequent edge. Let the color of the region with the tail ends of Γ_1 inside it be white, and let all the regions that aren't white be black. Following the orientation of Γ, numerically identify each edge in Γ_1 in the order it is encountered such that the first edge after B_Γ will be labeled 0, the next 1 and so forth until the last edge before B_Γ is enumerated. On G_1 this corresponds to a sequential identification of arcs between chords, disregarding the chords that intersect B_G (Fig. 7).

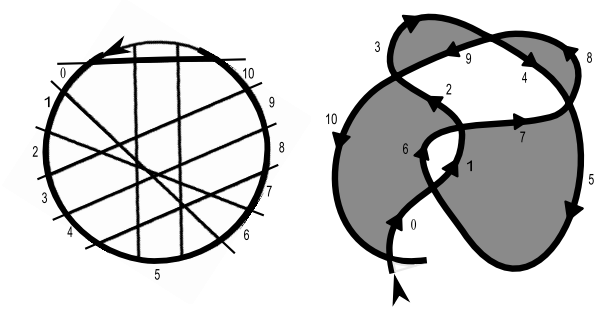


FIGURE 7. An example of the enumeration the edges of Γ_1 (right) and of the arcs of G_1 (left).

Now we must trace the rest of Γ. Any double point, X_Γ, created when Γ_2 crosses Γ_1 must be a move from white to black or black to white, alternating with each consecutive crossing. Thus we can number every such double point and, more importantly, its corresponding chord X_G on G with a function $\alpha(X_G)$ such that

$$\alpha(X_G) = \begin{cases} 1 & \text{if } X_\Gamma \text{ is a move from white to black} \\ 0 & \text{otherwise} \end{cases}$$

We see that the first such double point, H_Γ, must be a move from white to black since the remainder of Γ begins in the white region. Thus the first chord, H_G, that intersects B_G encountered after leaving G_1 and traversing G_2, will be numbered 1 and then each subsequent intersecting chord will be numbered alternatingly. We may also assign to each X_G the number $\beta(X_G)$ we earlier used to identify the arc that X_G contacts in G_1 (Fig. 8).

Given any fixed edge on Γ_1, a white to black crossing will have the opposite signing of a black to white crossing. Now consider that the sign of a white to black crossing for Γ_2 alternates with each subsequent edge in Γ_1, as every self-intersection in Γ_1 permutes the colors while fixing orientation. Since we can see that the first white to black crossing through edge 0 would be the opposite sign of B_Γ, we can deduce that the sign of any chord, X_G, intersecting B_G by the function

$$\text{sign}(X_G) = (-1)^{\alpha(X_G)+\beta(X_G)} \cdot \text{sign}(B_G).$$

Since $\alpha(X_G)$ and $\beta(X_G)$ can be derived from the Gauss diagram, we have the desired result in the case of a conveniently placed base point.

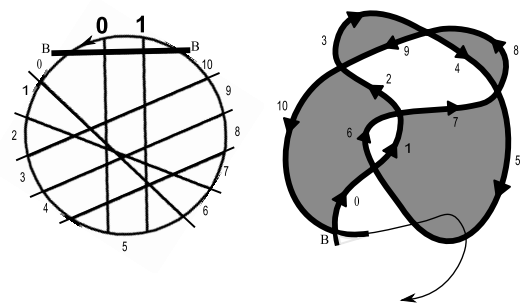


FIGURE 8. Traversing the remainder of G (left) corresponds to tracing out the remainder of Γ (right). We are able to apply the α function to the chords intersecting B_G, assigning the numbers 1 and 0 alternatingly as the corresponding curve crosses from white to black and then black to white. We also see that the β function yields 5 for both of these chords.

Suppose the base point of G started anywhere on the circle. Moving the base point through a chord corresponds to reversing the sign of that chord - as we are reversing the order of the outgoing arrows, $(1,2)$ - so we can move the base point to the default position before the start of G_1, noting for each chord the parity of the number of times we passed it along the way. Sign all the chords intersecting B_G using the signing function described above, and then reverse the sign of each intersecting chord whose parity differs from B_G. This is the signing of the intersecting chords given the original placement of the base point.

Thus given the sign of any chord we can sign every intersecting chord, regardless of the position of the base point and ignorant of the sign of the other chords. This proves our theorem. □

Following the thread of the proof of theorem 2.1, it isn't difficult to concisely describe the signing algorithm:

Proposition 2.4. The Signing Algorithm. *Suppose we have a signed chord, B_G, on an oriented Gauss diagram, G, with an oriented base point.*

(1) *Choose a side of B_G and call it G_1. Call the other side G_2.*
(2) *Move the base point so that, with the orientation on G, it is about to pass one of the endpoints of B_G and enter G_1, noting for each chord, X_G, the number of times, $|X_G|_\rho$, it is passed by the base point, and define the function*

$$\rho(X_G) = \begin{cases} 0 & \text{if } (|A_G|_\rho + |C_G|_\rho)/2 \in \mathbb{Z} \\ 1 & \text{otherwise.} \end{cases}$$

(3) *Traverse G_1 and, ignoring the chords that intersect B_G, enumerate each arc from one chord to the next such that the first arc is numbered* 0.
(4) *Traverse G_2 and for each consecutive chord, X_G, intersecting B_G we encounter, alternatingly assign the number, $\alpha(X_G)$,* 0 *or* 1, *starting with* 1.
(5) *Also, for each X_G assign the number $\beta(X_G)$, identifying the arc in G_1 that X_G contacts.*

(6) *The sign of any such X_G is*

$$\operatorname{sign}(X_G) = (-1)^{\alpha(X_G)+\beta(X_G)+\rho(X_G)} \cdot \operatorname{sign}(B_G)$$

An immediate consequence of the Signing Algorithm is that there is an upper bound on the number of unique signings imposed on composite Gauss diagrams, we discuss this along with classification of composite spherical curves in section 10.

Now that a bijection exists between unsigned prime Gauss diagrams and non-oriented prime spherical curves, and as all spherical curves can be constructed from these in terms of connected sums, classification of composite spherical curves from the prime factors can reduce generic curve classification to the planarity problem of chord diagrams. Fortunately, in [CE], Cairns and Elton found an algorithmic way to determine planarity of a chord diagram by way of its corresponding chord word, and we, in turn, can reduce the time of this algorithm by only considering chord diagrams in which the chords form a connected space.

To construct an algorithm for a chord word (which is not necessarily a Gauss word as we are only working with chord diagrams) that will determine whether or not the chords in the corresponding diagram form a connected space, we must examine how chord intersections manifest themselves in the word.

Definition 2.5. [CHB] A letter, c, in a Gauss word, W, is *between* the letter b if it appears alternatingly with c in W. We define the *betweenness function*, x_W of c to be the set of all different letters in W between c.

Example 2.6. *Consider the Gauss word* $W = (a,b,c,d,e,f,g,e,b,a,f,g,d,c)$, *then* $x_W(b) = \{c,d,f,g\}$. *Note that* $x_W(b)$ *is a set, not a word.*

Remark 2.7. A chord a intersects chord b in the chord diagram if and only if the corresponding letter a is between the letter b in the word.

This can be demonstrated by considering the two cases where chords do and do not cross and how their letters behave in the word.

Consider a set, A, of different letters in some chord word W. We define the function X_W on A as

$$X_W(A) = \left(\bigcup_{a \in A} x_W(a) \right) \cup A.$$

Given that $X_W(A)$ is itself a set of letters in W, we can compose this function with itself, $X_W^{(2)}(A)$, and since $A \subseteq X_W(A)$ and there are only n different letters in W, we find that $X_W^{(n)}(A)$ for $n \in \mathbb{N}$ strictly increases with n until $X_W^{(n)}(A) = X_W^{(n+1)}(A)$, at which point a trivial inductive step will show that $X_W^{(n)}(A) = X_W^{(m)}(A)$ for all $m \geq n$. We can now give our proposition.

Proposition 2.8. The Component Algorithm. *If the chords j and k are in the chord diagram D with corresponding word W, then j and k are in the same component if and only if there exists some $n \in \mathbb{N}$ such that $\{k\} \in X_W^{(n)}(\{j\})$.*

Proof. As the components of a chord diagram, D, partition the chord space they also define an equivalence relation, $\sim$, between chords that are in the same component. If the chords j and k intersect then we denote $j \simeq k$ and let $\simeq$ be reflexive. Note that if $j \simeq k$ then *a fortiori* $j \sim k$. Furthermore, by the construction of the components we have that

$$j \sim k \Leftrightarrow j \simeq l_1 \simeq l_2 \simeq \cdots \simeq l_{n-1} \simeq k \quad \text{for some chords } \{l_1, \ldots, l_{n-1}\} \in D.$$

Suppose $j \sim k$. Given the word, W, corresponding to D we see that $j \simeq l_1$ if and only if $\{l_1\} \in X_W(\{j\})$ by our initial properties of the betweenness function. Assume then that for some $m > 1$, $\{l_{m-1}\} \in X_W^{(m-1)}(\{j\})$, then if $l_{m-1} \simeq l_m$ we get that $\{l_m\} \in X_W^{(m)}(\{j\})$ since $\{l_m\} \in X_W(\{l_{m-1}\}) \subseteq X_W^{(m)}(\{j\})$. Thus by induction we get that $\{k\} \in X_W^{(n)}(\{j\})$. Conversely, suppose $\{k\} \in X_W^{(n)}(\{j\})$ for some $n \in \mathbb{N}$, then there exists a chord $l_{n-1} \in X_W^{(n-1)}(\{j\})$ such that $l_{n-1} \simeq k$ by construction of the function. Similarly, assume that for some $m < n$, $l_{m+1} \in X_W^{(m+1)}(\{j\})$, then there exists a chord $l_m \in X_W^{(m)}(\{j\})$ such that $l_m \simeq l_{m+1}$ and so we inductively construct the decreasing chain

$$k \simeq l_{n-1} \simeq \cdots \simeq l_1 \simeq j \Leftrightarrow j \sim k$$

which proves the proposition. □

A simple corollary of proposition 2.8 is that if for some letter b in some chord word, W, $X_W^{(n)}(\{b\}) = X_W^{(n+1)}(\{b\})$ for some $n \in \mathbb{N}$ then then $X_W^{(n)}(\{b\})$ is the complete set of chords of the component containing b. So when we want to consider chord diagrams that may potentially correspond to prime Gauss diagrams, we must consider chord words where $X_W^{(n)}(\{b\})$ gives the set of all letters in W for any $b \in W$. We incorporated this algorithm into our final program for Gauss diagram construction, which is included in the Appendix.

Furthermore, there are other methods we can employ to reduce the difficulty of the planarity problem. In the next section, we elaborate on a generalized approach to this.

Part 3. **Construction and Enumeration of Gauss Diagrams**

3. RESTRICTIONS

The magnitude of the planarity problem of chord diagrams is fairly daunting; the brute force method for constructing all Gauss diagrams with n double points involves pouring through a great number of cases, provided n is not trivially small. The planarity algorithm due to Elton and Cairns is effective, but also impractically slow for cases of more complicated chord diagrams. However, we can drastically reduce the set of chord diagrams we actually have to analyze. For this, we use the notion of prime and composite chord diagrams and corollary 1.11, as well as the following corollary:

Corollary 3.1. *A chord diagram is a Gauss diagram if and only if its component chord diagrams are Gauss diagrams.*

As we to referred to earlier, since every composite Gauss diagram can be easily constructed from its prime components, we need only construct primes. Also helpful in reducing the number of chords diagrams is a theorem attributed to Gauss.

Theorem 3.2. [Gauss] *A chord in a Gauss diagram must intersect an even number of other chords.*

This means every chord in our construction must have an even number of points between its endpoints. So if we enumerate the points on our n chord diagram from 0 to $2n-1$, every chord must join an even numbered point to an odd numbered point. We can further reduce the number of chord diagrams if we eliminate chord diagrams that are the same under rotation and reflection. This process will be explained in the following sections.

4. THE CONSTRUCTION METHOD

Before we describe the construction method, we first state our restrictions in the form of rules that we will follow as we proceed.

- Rule 1: Every chord must join an odd numbered point to an even numbered point
- Rule 2: Every chord diagram constructed must be prime
- Rule 3: Every chord diagram constructed can be equivalent to no previously constructed chord diagram (see Section 5 on cataloging).

As mentioned above, these rules allow us to drastically reduce the number of chord diagrams we have to analyze. Rule 1 is utilized to automatically eliminate chord diagrams which violate Gauss' necessary criterion for planarity. Rule 2 restricts our attention to prime chord diagrams. Rule 3 prevents us from repeating equivalent chord diagrams in the construction process.

Definition 4.1. We call the smaller endpoint of a chord the *base*. The larger endpoint of the chord is called the *pivot. We denote a chord with base b and pivot p by* $\binom{b}{p}$.

To construct all n chord diagrams, we start by drawing a circle with $2n$ enumerated points (starting with 0) evenly distributed along the circumference. First, we draw a chord by connecting the smallest numbered point to the smallest numbered rule-permitting point. We keep doing this until we cannot continue without violating a rule or the chord diagram is filled.

4.1. **Performing an Upshift.** If the we cannot continue without violating a rule, we will need to perform an upshift. To do this, we take the chord with the largest base and fix the base while moving the pivot up to the next rule-permitting point. If the pivot is already in the largest rule-permitting point, we erase the chord completely and perform an upshift on the chord that now has the largest base.

After performing an upshift, we continue by connecting the smallest available number to the smallest rule-permitting number and repeating the above process.

4.2. **When the Chord Diagram is Filled.** When every point on the circle is connected to some other point, the chord diagram is filled. Record the completed chord diagram, and then erase the chord with the largest base and perform an upshift.

Here, we outline the procedure for constructing all prime, even chord intersection chord diagrams for n double points.

Procedure 4.2.
(1) *Connect the smallest available number to the smallest rule-permitting number.*
(2) *If you can repeat without violating any rule, repeat step 1. If the smallest available number cannot be connected to any point without violating a rule, go to step 3. If every point on the circle is connected to some other point, skip to step 5.*
(3) *Perform an upshift (described in Section 4.1). If you reach the point where performing an upshift creates a chord with base 0 and a pivot greater than n, skip to Step 8.*
(4) *Go to Step 1.*
(5) *Record the filled chord diagram in the catalog.*
(6) *Erase the chord with the largest base.*
(7) *Go to step 3.*
(8) *You are done.*

4.3. **Example.** To construct all 4 chord diagrams that satisfy our rules, start by drawing a line connecting point 0 to 3. (Note: Connecting 0 to 1 would necessarily make the chord diagram composite and violate Rule 2. Likewise, connecting 0 to 2 would violate Rule 1). Next, connect 1 to 4. Now we're stuck. Connecting 2 to 5 or 7 would make the chord diagram composite. This is illustrated in Figure 9.

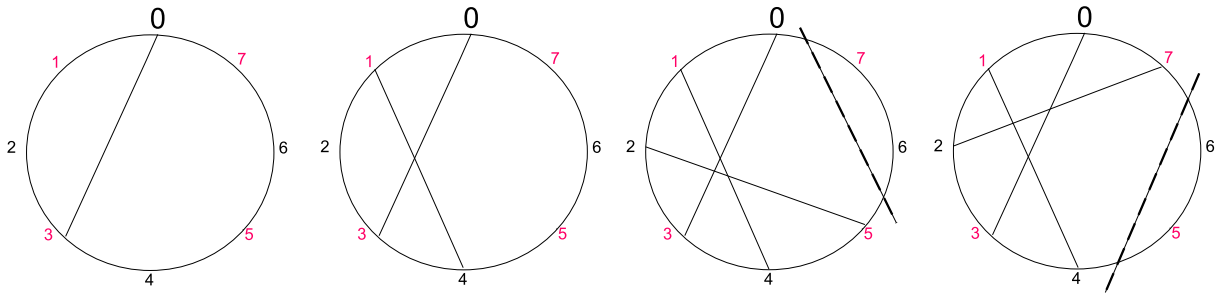


FIGURE 9. Here we illustrate the method applied to 4 chord diagrams.

So now we need to perform an upshift. Move the pivot at 4 to 6 so that we have chords connecting 0 to 3 and 1 to 6, as in Figure 10. Next, connect 2 to 5. Now connect 4 to 7 to fill the chord diagram and record, as in Figure 11.

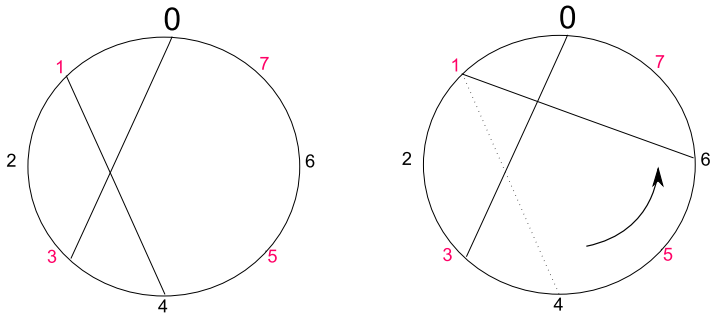


FIGURE 10. Performing an upshift

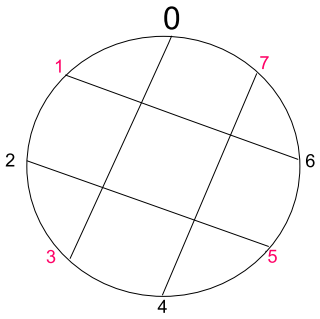


FIGURE 11. The chord diagram formed by connecting 0 to 3, 1 to 6, 2 to 5, and 4 to 7.

Now that we've found a chord diagram that fits the rules, we erase the chord with the largest base (in this case, the $\binom{4}{7}$ chord) and perform an upshift on the $\binom{2}{5}$ chord to make it a $\binom{2}{7}$ chord. Now we're in trouble because we made our diagram composite, as seen in Figure 12.

Now we erase our $\binom{2}{7}$ chord and perform an upshift on the $\binom{1}{6}$ chord. But our $\binom{1}{6}$ chord is already in the largest rule permitting point. Thus, we erase our $\binom{1}{6}$ chord and upshift our $\binom{0}{3}$ chord to a $\binom{0}{5}$ chord, as in Figure 13.

At this point, we should notice that we have gone through every possible chord diagram that has a chord like our $\binom{0}{3}$ chord. That is, we have done all chord diagrams that has a chord with length 2. Thus, we don't have to check chord diagrams with a $\binom{0}{5}$ chord because any chord diagram we get from it will be a reflection of a chord diagram we already cataloged. In general, if you're constructing all n chord diagrams, once your 0 base chord passes the nth point on the circle, you know you're done.

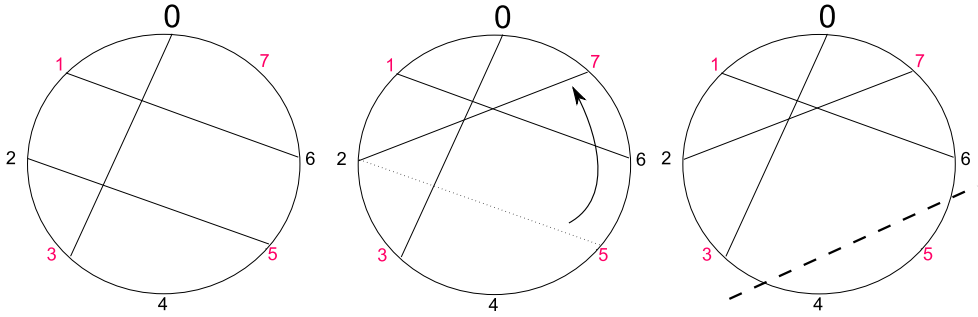


FIGURE 12. After erasing the $\binom{4}{7}$ chord, we perform an upshift on the $\binom{2}{5}$ chord to make it a $\binom{2}{7}$ chord. However, this makes our chord diagram composite.

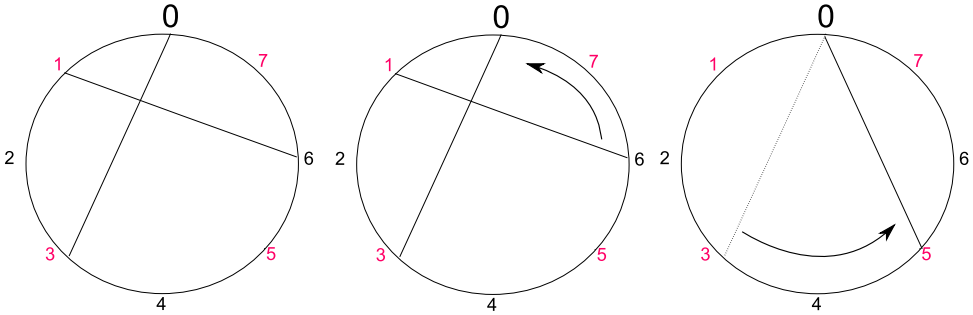


FIGURE 13. After erasing the $\binom{2}{7}$ chord, we normally would perform an upshift on the $\binom{1}{6}$ chord. However, since its already in the largest rule permitting point, we erase our $\binom{1}{6}$ chord and upshift our $\binom{0}{3}$ chord to a $\binom{0}{5}$ chord.

5. CATALOGING

For small n it will be fairly simple to tell whether Rule 3 has been violated; one can easily inspect whether he or she has already recorded some rotation and reflection of a particular chord diagram. However, for large n, it is not always easy to detect chord diagrams related by a rotation and a reflection. We would like to prevent searching through every recorded chord diagram (and there may be many!) just to see if it has already been recorded in some equivalent form. Fortunately, by proceeding by the above method, a natural order is already in place, which will make it possible to efficiently determine if an equivalent chord diagram has already been recorded.

Definition 5.1. We call the set of endpoints (in ascending order by base) of each chord in a chord diagram Γ the *fingerprint* of Γ. Informally[1], the *primary fingerprint* of a chord diagram is the fingerprint of the equivalent chord diagram that appears first in the catalog. *An example is given in Figure 14.*

Consequently, the primary fingerprint of a chord diagram is a complete invariant of that chord diagram up to rotation and reflection.

[1]A more formal definition will follow.

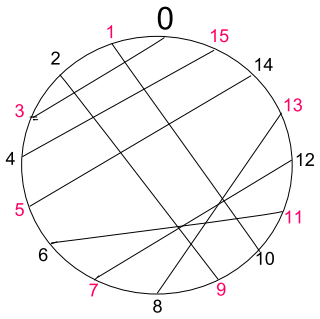


FIGURE 14. This chord diagram has fingerprint $\binom{0}{3}\binom{1}{10}\binom{2}{9}\binom{4}{15}\binom{5}{14}\binom{6}{11}\binom{7}{12}\binom{8}{13}$. This is also its primary fingerprint.

5.1. **Computing Primary Fingerprint.**

Definition 5.2. The *length* of a chord is given by the number of points contained in the smaller arc defined by that chord. *An example is given in Figure 15.*

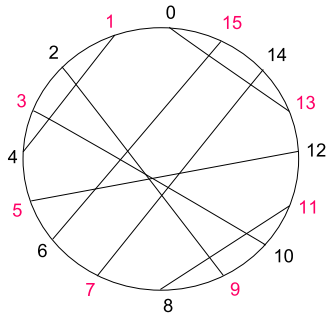


FIGURE 15. The length of the $\binom{0}{13}$ chord is 2. The length of the $\binom{2}{9}$ chord is 6.

Given any chord diagram, we can easily compute its primary fingerprint. First, we locate the chord(s) with the smallest length, l. Now if there is only one chord with length l, we must look at both of its endpoints separately. Relabel the points so that the base is 0 and the pivot is $l+1$. From this new enumeration, compute its fingerprint. Now we switch, labeling the pivot 0 and the base $l+1$, and compute its fingerprint. To determine which is the primary fingerprint we compare the fingerprints from left to right and find the leftmost pair of endpoints at which the two fingerprints are not the same. Now we compare the pivots of that chord. The fingerprint with the lesser pivot is the primary fingerprint of the chord diagram.

Example 5.3. *In Figure 16, the* $\binom{10}{13}$ *chord has the smallest length, 2, so we relabel the 10 with 0 and the 13 with 3 (middle ring of numbers). Next we relabel the 13 with 0 and the 10 with 3 (outer ring of numbers). Doing this gives fingerprints of* $\binom{0}{3}\binom{1}{10}\binom{2}{9}\binom{4}{11}\binom{5}{12}\binom{6}{15}\binom{7}{14}\binom{8}{13}$ *and* $\binom{0}{3}\binom{1}{10}\binom{2}{9}\binom{4}{13}\binom{5}{12}\binom{6}{11}\binom{7}{14}\binom{8}{15}\binom{5}{12}$*, respectively. Notice, the base 4 chord is the leftmost chord in which the two fingerprints differ. The former fingerprint has a lesser base 4 pivot, 11, than the latter fingerprint, 13. Thus,* $\binom{0}{3}\binom{1}{10}\binom{2}{9}\binom{4}{11}\binom{5}{12}\binom{6}{15}\binom{7}{14}\binom{8}{13}$ *is the primary fingerprint of this chord diagram.*

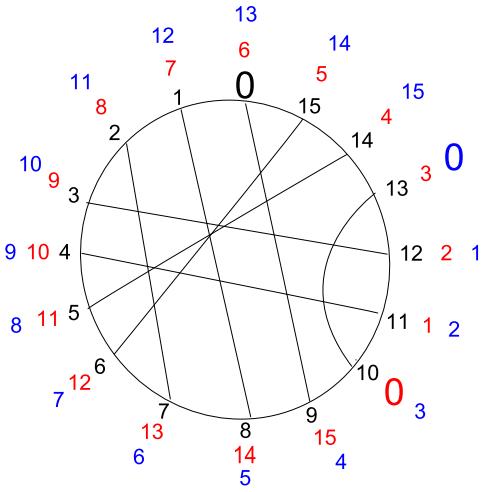


FIGURE 16. Determining primary fingerprint

If both relabelings give the same fingerprint, they both are the primary fingerprint of the chord diagram.

Example 5.4. *In Figure 17, the* $\binom{12}{15}$ *chord has the smallest length, 2, so we relabel the 12 with 0 and the 15 with 3 (middle ring of numbers). Next we relabel the 15 with 0 and the 12 with 3 (outer ring of numbers). Doing this gives fingerprints of* $\binom{0}{3}\binom{1}{8}\binom{2}{11}\binom{4}{9}\binom{5}{12}\binom{6}{13}\binom{7}{14}\binom{10}{15}$ *and* $\binom{0}{3}\binom{1}{8}\binom{2}{11}\binom{4}{9}\binom{5}{12}\binom{6}{13}\binom{7}{14}\binom{10}{15}$*, respectively. Notice, both are exactly the same. Thus, they both are the primary fingerprint of the chord diagram.*

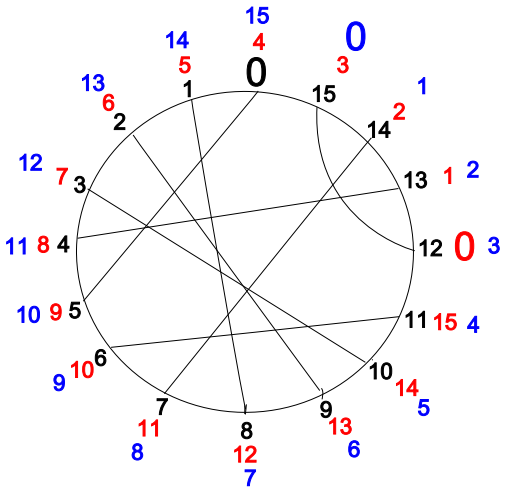


FIGURE 17. Determining primary fingerprint when two relabelings give the same fingerprint

If there are multiple chords that have the same smallest length, we must repeat the above process with each chord and compare the fingerprints we get from them in the same manner.

Example 5.5. *In Figure 15, we had three chords with smallest length, 2. Thus, to compute its primary fingerprint, we must perform 6 relabelings, as depicted in Figure 18. This gives us fingerprints of*

(1) $\binom{0}{3}\binom{1}{10}\binom{2}{9}\binom{4}{11}\binom{5}{8}\binom{6}{13}\binom{7}{14}\binom{12}{15}$
(2) $\binom{0}{3}\binom{1}{8}\binom{2}{9}\binom{4}{11}\binom{5}{14}\binom{6}{13}\binom{7}{10}\binom{12}{15}$
(3) $\binom{0}{3}\binom{1}{10}\binom{2}{11}\binom{4}{7}\binom{5}{14}\binom{6}{13}\binom{8}{15}\binom{9}{12}$
(4) $\binom{0}{3}\binom{1}{10}\binom{2}{9}\binom{4}{7}\binom{5}{12}\binom{6}{13}\binom{8}{15}\binom{11}{14}$
(5) $\binom{0}{3}\binom{1}{8}\binom{2}{9}\binom{4}{13}\binom{5}{12}\binom{6}{15}\binom{7}{10}\binom{11}{14}$
(6) $\binom{0}{3}\binom{1}{10}\binom{2}{11}\binom{4}{13}\binom{5}{8}\binom{6}{15}\binom{7}{14}\binom{9}{12}$.

Comparing left to right, we eliminate the 1^{st}, 3^{rd}, 4^{th}, *and* 6^{th} *fingerprints because of their larger base 1 pivot. This leaves us with the* 2^{nd} *and* 5^{th} *fingerprints, which are the same until the base 4 chord. Since the* 2^{nd} *fingerprint has the smaller base 4 pivot, we know it must be the primary fingerprint.*

Notice we really didn't have to compute each fingerprint thoroughly. Once we know we have a relabeling with a $\binom{1}{8}$ *chord, we can immediately rule out relabelings that give* $\binom{1}{10}$ *chords. Likewise, we can further rule out the* 5^{th} *relabeling once we encounter the* $\binom{4}{13}$ *chord.*

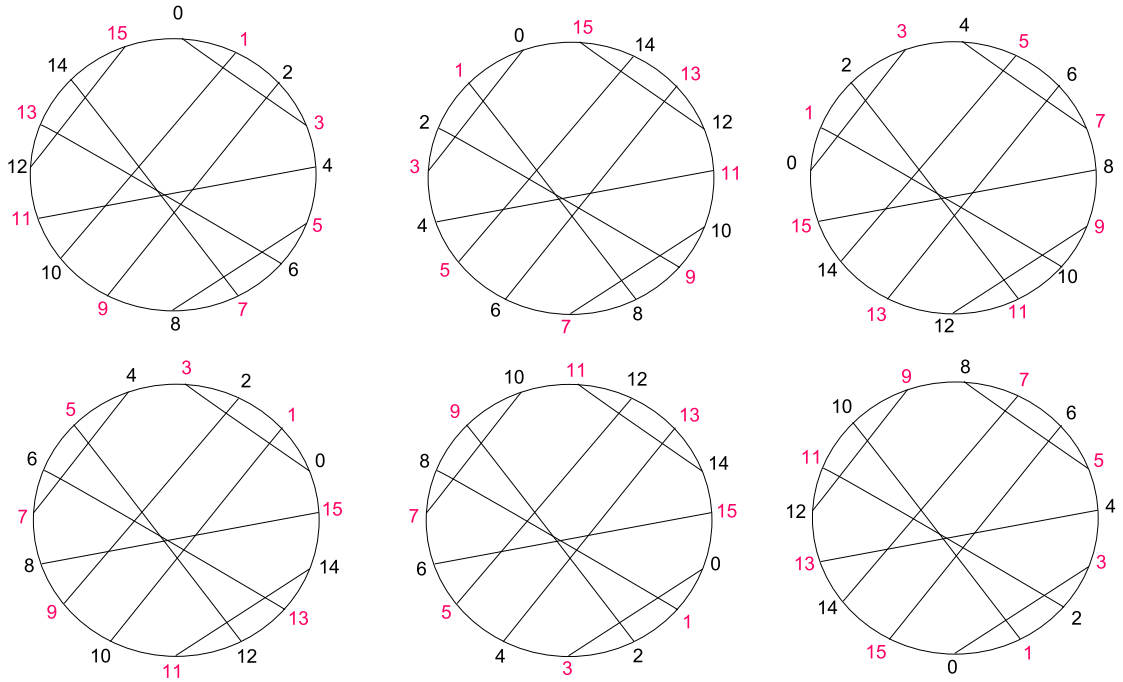


FIGURE 18. Six relabelings of the chord diagram in Figure 15

As a result, when going through the method, if the fingerprint of a chord diagram differs from its primary fingerprint, we must have already recorded an equivalent chord diagram in our catalog. Often times, we can tell quickly without even computing the primary fingerprint of a chord diagram. For instance, if we are working through the method and reach the point where we have a

$\binom{0}{5}$ chord, we know any chord diagram with a chord of length 2 must have already been cataloged since creating a chord of length 2 will necessarily mean the fingerprint of the chord diagram will differ from its primary fingerprint. In fact, we can work this into the construction process by limiting ourselves to drawing chords that have a length greater than or equal to the length of the chord with base 0. This should significantly reduce the number of chord diagrams we have to construct.

5.2. **Lexicographical Ordering.** Now that we have a good understanding of what the primary fingerprint of a chord diagram is, we can begin to define it in a more rigorous setting. First we must establish that the set of all chords for n double points is a well-ordered set. The ordering of chords are given in the following way.

$$\binom{b_i}{p_i} < \binom{b_j}{p_j} \text{ if and only if } p_i - b_i < p_j - b_j.$$
$$\binom{b_i}{p_i} = \binom{b_j}{p_j} \text{ if and only if } p_i - b_i = p_j - b_j.$$

It can be easily verified that this gives a total ordering. Because the set of chords for n double points is finite, we have that it is well-ordered. Now we consider the set of all fingerprints for n double points and define a lexicographical ordering which consequently preserves well-ordering.

$$\binom{b_{1,1}}{p_{1,1}}\binom{b_{2,1}}{p_{2,1}}\ldots\binom{b_{n,1}}{p_{n,1}} < \binom{b_{1,2}}{p_{1,2}}\binom{b_{2,2}}{p_{2,2}}\ldots\binom{b_{n,2}}{p_{n,2}} \text{ if and only if there exists an } m > 0 \text{ such that for all } i < m,$$
$$\binom{b_{i,1}}{p_{i,1}} = \binom{b_{i,2}}{p_{i,2}} \text{ and } \binom{b_{m,1}}{p_{m,1}} < \binom{b_{m,2}}{p_{m,2}}.$$

In other words, the first fingerprint is less than the second if and only if we have $\binom{b_{m,1}}{p_{m,1}} < \binom{b_{m,2}}{p_{m,2}}$ and all the preceding chords are equal. This is analogous to the alphabetical ordering of words, where the first letter is given the most weight. Equality follows in the expected way.

$$\binom{b_{1,1}}{p_{1,1}}\binom{b_{2,1}}{p_{2,1}}\ldots\binom{b_{n,1}}{p_{n,1}} = \binom{b_{1,2}}{p_{1,2}}\binom{b_{2,2}}{p_{2,2}}\ldots\binom{b_{n,2}}{p_{n,2}} \text{ if and only if for all } i, \binom{b_{i,1}}{p_{i,1}} = \binom{b_{i,2}}{p_{i,2}}.$$

Now we can redefine the primary fingerprint of a chord diagram as follows.

Definition 5.6. The *primary fingerprint* of a chord diagram Γ is the least element in the set of all equivalent fingerprints of Γ.

Remark 5.7. The least element in any well-ordering is necessarily unique.

The proof that this is equivalent to the original definition is seen from the fact that chord diagrams are constructed in lexicographical order with smaller fingerprints being constructed first.

While lexicographical ordering does give us a more rigorous definition, it also imposes a well-ordering on primary fingerprints and chord diagrams.

5.3. **Proof of the Construction Method.** Now it becomes important to prove that following Procedure 4.2 does indeed construct all prime, even chord intersection chord diagrams with no repetition.

Theorem 5.8. *The construction algorithm yields all prime, even chord intersection chord diagrams with no repetition.*

Proof. Given any prime, even chord intersection chord diagram for a given number of double points, it has a set of equivalent fingerprints. Because this set is well-ordered, it must have a least element. This least element is the primary fingerprint and corresponds to the first recorded chord diagram of its equivalence class. That no two equivalent chord diagrams are cataloged is guaranteed by Rule 3. □

6. Numerical Notation

We know that the Gauss diagram is a complete invariant for prime spherical curves. However, we'd like to take it one step further and assign to each prime spherical curve a number so that no two different numbers correspond to the same prime spherical curve and no two different prime spherical curves correspond to the same number. More formally, we want a injective function between the set of all prime spherical curves and the nonnegative integers. This leads to the the following.

Given a prime spherical curve γ with primary fingerprint $\binom{b_1}{p_1}\binom{b_2}{p_2}...\binom{b_n}{p_n}$, we define the numerical function of γ by

$$\Phi(\gamma) = (p_1 - b_1)(2n)^{n-1} + (p_2 - b_2)(2n)^{n-2} + ... + (p_n - b_n).$$

The trivial curve, K_1, is by convention assigned the number 0.

Theorem 6.1. *If two prime spherical curves correspond to the same numerical notation, they must be the same prime spherical curve.*

Proof. Suppose we have two prime spherical curves γ_1 and γ_2 with primary fingerprints $\binom{b_{1,1}}{p_{1,1}}\binom{b_{2,1}}{p_{2,1}}...\binom{b_{n,1}}{p_{n,1}}$ and $\binom{b_{1,2}}{p_{1,2}}\binom{b_{2,2}}{p_{2,2}}...\binom{b_{m,2}}{p_{m,2}}$, respectively, and suppose that

$$\Phi(\gamma_1) = \Phi(\gamma_2).$$

This means

$$(p_{1,1} - b_{1,1})(2n)^{n-1} + (p_{2,1} - b_{2,1})(2n)^{n-2} + ... + (p_{n,1} - b_{n,1}) = \\ (p_{1,2} - b_{1,2})(2m)^{m-1} + (p_{2,2} - b_{2,2})(2m)^{m-2} + ... + (p_{m,2} - b_{m,2}). \tag{1}$$

First let's assume $n \neq m$ and without loss of generality say $n < m$. Then $(2n)^n < (2m)^n$. Since $n < m$ are integers, we have

$$(2n)^n < (2m)^n \leq (2m)^{m-1}. \tag{2}$$

Now for any chord $\binom{b}{p}$ with an n double point Gauss diagram, we must have $0 < p - b < 2n$. The numerical notation then can be thought of as a number expanded in a base-$2n$ numerical system where each digit ranges from 0 to $2n-1$. So we have

$$(p_{1,1} - b_{1,1})(2n)^{n-1} + (p_{2,1} - b_{2,1})(2n)^{n-2} + ... + (p_{n,1} - b_{n,1}) < (2n)^n. \tag{3}$$

For our m double point curve, we have

$$(2m)^{m-1} < (p_{1,1} - b_{1,1})(2m)^{m-1} + (p_{2,1} - b_{2,1})(2m)^{m-2} + ... + (p_{m,1} - b_{m,1}). \tag{4}$$

Combining (2), (3), and (4) gives

$$(p_{1,1}-b_{1,1})(2n)^{n-1}+(p_{2,1}-b_{2,1})(2n)^{n-2}+\ldots+(p_{n,1}-b_{n,1})<$$
$$(p_{1,2}-b_{1,2})(2m)^{m-1}+(p_{2,2}-b_{2,2})(2m)^{m-2}+\ldots+(p_{m,2}-b_{m,2}).$$

But this is a contradiction so we must have $n=m$. So let's rewrite (1) as

$$(p_{1,1}-b_{1,1})(2n)^{n-1}+(p_{2,1}-b_{2,1})(2n)^{n-2}+\ldots+(p_{n,1}-b_{n,1})=$$
$$(p_{1,2}-b_{1,2})(2n)^{n-1}+(p_{2,2}-b_{2,2})(2n)^{n-2}+\ldots+(p_{n,2}-b_{n,2}).$$

Again, let's think of our numerical notation as a number expanded in a base-$2n$ numerical system where each digit ranges from 0 to $2n-1$. This means each $p_{i,1}-b_{i,1}=p_{i,2}-b_{i,2}$.[2] It immediately follows that the primary fingerprints are the same. Therefore γ_1 and γ_2 must be the same spherical curve. □

6.1. **Examples.** Here we give the numerical notation of all prime spherical curves up to seven double points.

0, 1, 129, 1883, 35553, 55555, 859599, 901073, 901359, 25490069, 26560621, 26565745, 26566081, 26642969, 26648039, 26648429, 41702043, 41778483, 56761117

Notice how we are able encode a lot of information in a relatively succinct manner. Moreover, by expressing our prime spherical curves in numerical notation, we have a specific and systematic classification.

6.2. **Reconstruction from Numerical Notation.** Given the numerical notation $\Phi(\gamma)$ for a prime spherical curve γ, we can build that curve. There must exist an n such that $(2n)^{n-1}<\Phi(\gamma)<(2n+2)^n$. This n is the number of double points for γ. Now we expand $\Phi(\gamma)$ into base-$2n$. From the coefficients on the $(2n)^i$ terms, we can determine the primary fingerprint of γ, and thus determine γ.

Example 6.2. *Let's reconstruct our prime spherical curve from the numerical notation 56761117. First we find our n. Since*

$$14^6<56761117<16^7,$$

we know our curve must have 7 double points. Now we expand 56761117 into base-14. This gives

$$7(14^6)+7(14^5)+7(14^4)+7(14^3)+7(14^2)+7(14)+7=56761117.$$

So our curve has primary fingerprint $\binom{0}{7}\binom{1}{8}\binom{2}{9}\binom{3}{10}\binom{4}{11}\binom{5}{12}\binom{6}{13}$*, which corresponds to the 7 point star (Figure 19).*

We have given a method for constructing and cataloging all chord diagrams with a given number of double points. Ideally, this method could be implemented into a computer program and together with an algorithm based on the result in [CE], one could efficiently generate all prime Gauss diagrams for a given number of double points. We have in fact implemented a form of this method into our algorithm for Gauss diagram construction, which is given in the Appendix. We have

[2]This is analogous to comparing two numbers in base-10 by comparing the digits.

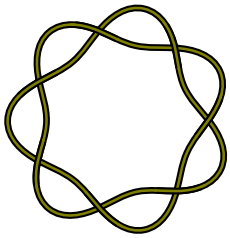


FIGURE 19. The 7 point star

also introduced a complete invariant for chord diagrams. Thus we have also defined a complete numerical invariant for spherical curves.

Part 4. **The Dowker Notation**

While still on the subject of prime Gauss diagrams, it seems worthy to take a brief aside into knot theory.

Given a Gauss word, W, with $2n$ letters that corresponds to some spherical curve, Γ, we are able to pick a starting point on the word and enumerate the letters from 1 to $2n$ in either direction and then identify the pairs of numbers corresponding to the same letter. By a theorem due to Gauss, for any letter $b \in W$ there is always an even number of letters between b and the other b, and thus every pair of numbers consists of one even and one odd number. We can then list these pairs in the increasing order of the odd numbers and then suppress the odd numbers entirely to get a sequence of n even numbers. This sequence is the *Dowker Notation* for the curve Γ, and as the process for its construction is reversible, it is an equivalent invariant to the Gauss word.

Example 6.3. *Given the Gauss word* $W = (a,b,c,d,e,f,g,e,b,a,f,g,d,c)$ *we can derive the Dowker notation*

$$\begin{pmatrix} a & b & c & d & e & f & g & e & b & a & f & g & d & c \\ 9 & 10 & 11 & 12 & 13 & 14 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 6 & 10 & 14 & 12 & 4 & 8 & 2 \\ 1 & 3 & 5 & 7 & 9 & 11 & 13 \end{pmatrix}$$
$$\Leftrightarrow (6,10,14,12,4,8,2)$$

In knot theory, the Gauss word and the Dowker notation are used for general knot classification much in the same way as they are used in alternating knot classification; in fact, any knot projection corresponding to a given Dowker notation is made by constructing a corresponding spherical curve and resolving each crossing alternatingly. The Dowker notation can also be extended beyond alternating knots: if one signs the numbers - positively or negatively - to indicate the resolution of the crossings, every knot can be represented in this way.

An interesting property of an unsigned Dowker notation with more than one even number is that it will uniquely describe a prime alternating knot (and its mirror image if the knot is not amphichiral) if the sequence of numbers cannot be broken into permutations of two consecutive sequences [Ad]. This gives rise to an interesting proposition.

Proposition 6.4. *A prime spherical curve with* $n > 2$ *double points uniquely resolves to the projection of a prime alternating knot (and its mirror image if the knot is not amphichiral) with the least number, n, of crossings (Fig. 20).*

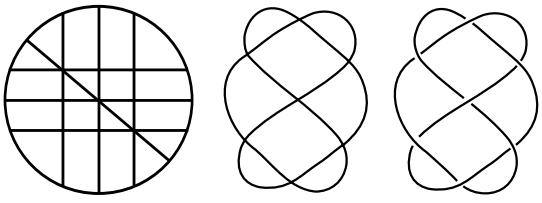


FIGURE 20. A prime Gauss diagram (left) gives rise to a prime spherical curve (middle) which gives rise to an irreducible prime alternating knot (right).

Proof. Suppose an $n > 2$ number Dowker notation can be broken into permutations of two nonempy consecutive subsequences, A and B, then there exists an even number, ω in A such that $\alpha \leq \omega$ for every $\alpha \in A$ and $\beta > \omega$ for every $\beta \in B$. If we write the corresponding numbers out, 1 through $2n$, we see that every number less than or equal to ω in the sequence is in A if it is even, or corresponds to some number in A if it is odd. Similarly, every number greater than ω is, or corresponds to, some number in B. Thus all pairs of corresponding numbers are either in $\bar{A} = \{1, \ldots, \omega\}$ or $\bar{B} = \{\omega + 1, \ldots, 2n\}$. If we label each pair of numbers with a corresponding letter, creating a Gauss word, then no letter corresponding to numbers in $\bar{A}$ can be between letters corresponding to numbers in $\bar{B}$ and thus none of the letters in $\bar{A}$ can be in the same component as $\bar{B}$, and so the corresponding Gauss diagram, G, is composite.

It follows from above that if the Gauss diagram is prime and has more than two chords, then the Dowker notation cannot be broken into two consecutive subsequences and thus it uniquely defines a prime alternating knot and its mirror image. Since the alternating resolution of every spherical curve determined by this Dowker notation is also an alternating knot determined by the notation, it follows that every prime spherical curve with at least three crossings will be resolved to this prime alternating knot or its mirror image.

The projection of an alternating knot is reducible if and only if the corresponding spherical curve is also reducible; this likely motivated Arnold's classification of such spherical curves. Thus since every prime spherical curve with $n > 2$ double points is irreducible, it follows that every corresponding alternating prime knot projection, which must also have n crossings, is also irreducible. It has been proven that a reduced alternating projection of a knot has the least number of crossings of any projection of that knot [Ad], and thus the resolution of a spherical curve with $n > 2$ double points gives a prime alternating knot projection with the minimal number, n, of crossings. Thus the proposition is proven. □

We also have the following remark.

Remark 6.5. Every projection of a prime knot is a prime spherical curve.

Thus there is a bijection between projections of prime, reduced alternating knots and prime spherical curves.

It should be noted that while every prime spherical curve can uniquely describe a prime alternating knot and its mirror image, two completely different spherical curves can be resolved into the same knot. An example is given in Fig. 21.

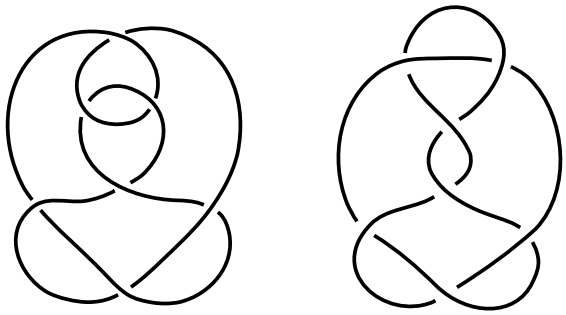


FIGURE 21. Two different reduced alternating knot projections of a prime knot, 7_5, which are the resolutions of two diffeomorphically inequivalent spherical curves.

Part 5. Invariants of Plane Curves

7. Arnold Invariants , The Polyak Function, and the Zen Invariant

When considering plane curve classification, it is often important to note the well-established invariants of Whitney and Arnold. The *Whitney index*, or *index* of an oriented plane curve is the integer number of rotations made by the vector tangent to the curve as it is traversed. It is conventional to regard a complete counterclockwise rotation as positive while a clockwise rotation would be negative. Allowing for an oriented plane curve with nonzero index to have the opposite index when we reverse this orientation. Thus we will use the absolute value of the index for either orientation as the index of the unoriented plane curve (Fig. 22).

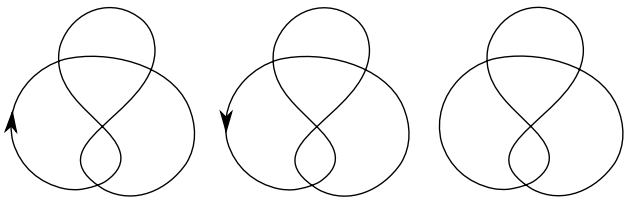


FIGURE 22. An oriented plane curve (left) and the same plane curve with opposite orientation (middle) with indices -1 and 1 respectively. The unoriented plane curve (right) has index $|\pm 1| = 1$.

Whitney proved that two generic plane curves were homotopic if and only if they have the same index [Wh]. Following from this, Arnold proved that all generic plane curves could be deformed into other, typically diffeomorphically inequivalent generic plane curves through what he called J^+, J^- and *strangeness*(Fig. 23) moves if and only if the two curves were homotopic. Assigning values to these moves and normalizing them with respect to *standard representatives* (Fig. 24) of all the homotopy classes gives rise to the *Arnold invariants*, which are respectively named for the moves that define them. For a more thorough treatment on the construction and the direct calculation of Arnold invariants, it is recommended that you consult [Ar1] or [Ar3].

Arnold was aware that these first-order invariants, along with the Whitney index, were insufficient for differentiating all diffeomorphically inequivalent plane curves. Polyak, in [Po], tackled the problem of complete plane curve classification by demonstrating that *n*-th order invariants of plane curves could be constructed by functions on the signed Gauss diagram and the index, and successfully defined the Arnold invariants in terms of these functions. We elaborate on this.

Definition 7.1. [Po] A representation, $\phi : A \to G$, of a chord diagram, A, in a signed Gauss diagram, G, is an embedding of A to G, mapping the circle of A to the circle of G (preserving orientation), each of the chords of A to a chord of G and a basepoint to a basepoint.

So given any chord, $c \in A$, the the corresponding chord, $\phi(c) \in G$, has a sign associated with the signing of G. So we define the signing function on ϕ as

$$\text{sign}(\phi) = \prod_{c \in A} \text{sign}(\phi(c)).$$

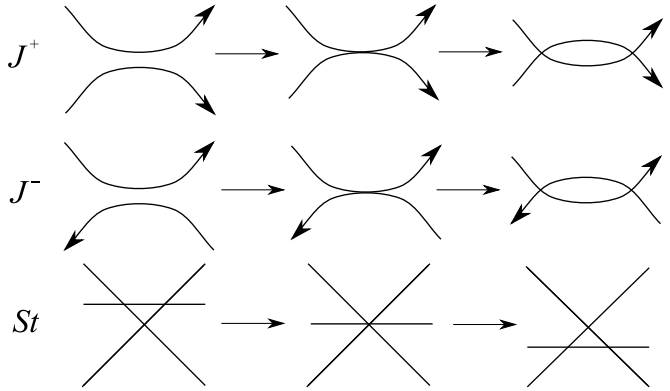


FIGURE 23. The moves that give rise to the Arnold invariants. The J^+ move (top) is the resolution of a direct self-tangency, the J^- move (middle) is the resolution of an inverse self-tangency, and the strangeness, St, move (bottom) is the resolution of a triple point.

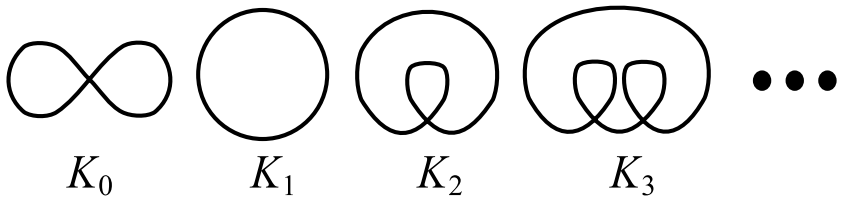


FIGURE 24. The standard representatives of all the homotopy classes of generic plane curves, corresponding to Whitney index. Curve K_n will have the index n.

This is used to define the *Polyak function:*

$$\langle A, G\rangle = \sum_{\phi: A \to G} \text{sign}(\phi).$$

Letting $\mathbf{A}$ be a vector space over $\mathbb{Q}$ generated by chord diagrams, $\langle A, G\rangle$ may be extended to $A \in \mathbf{A}$ by linearity.

Polyak gave a very useful theorem relating to this function:

Theorem 7.2. [Po] *Choose a base point on* Γ *and denote by G the corresponding signed Gauss diagram of* Γ. *Then*

$$J^+(\Gamma) = \langle B_2 - B_3 - 3B_4, G\rangle - \frac{n-1}{2} - \frac{ind(\Gamma)^2}{2}$$

$$J^-(\Gamma) = \langle B_2 - B_3 - 3B_4, G\rangle - \frac{3n-1}{2} - \frac{ind(\Gamma)^2}{2}$$

$$St(\Gamma) = \frac{1}{2}\langle -B_2 + B_3 + B_4, G\rangle - \frac{n-1}{4} - \frac{ind(\Gamma)^2}{4}.$$

Where $J^+(\Gamma)$, $J^-(\Gamma)$, *and* $St(\Gamma)$ *are the Arnold invariants of* Γ, $ind(\Gamma)$ *is the Whitney index, and* B_2, B_3 *and* B_4 *are given by Figure 25.*

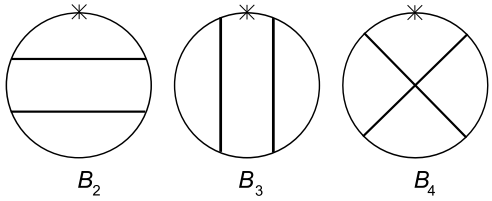


FIGURE 25. The based chord diagrams for use with the equations in theorem 7.2.

Polyak defined the *degree* of any chord diagram to be the number of chords. The degree of $A \in \mathbf{A}$ is the highest degree of the diagrams in A. This relates closely to the concept of the *degree* of a plane curve invariant:

Definition 7.3. [Po] An invariant of plane curves is said to be of *degree* less or equal m if it vanishes on any singular curve with at least $m+1$ self-tangency or triple points.

Polyak demonstrated this relation in another useful theorem.

Theorem 7.4. [Po] *Let $A \in \mathbf{A}$ be a linear combination of chord diagrams of degree less or equal to m. Suppose that the Gauss diagram invariant $\langle A, G\rangle$ is independent of the choice of base point for any plane curve Γ with Gauss diagram G. Then $\langle A, G\rangle$ is an invariant of plane curves of degree less or equal to $[\frac{m}{2}]$.*

It is clear that Arnold's invariants are of degree one by this theorem. To further demonstrate the practicality of Polyak's results, we explicitly constructed our own invariants of n-th degree, based on a relatively simple curve.

The *n-armed Zen Master*, or *n-Master* for short, is a diffeomorphic deformation of the circle on the plane such that n non-intersecting protrusions create n inverse self-tangencies (Fig 26). The n-Master can be resolved in 2^n ways, resulting in at most 2^n different generic plane curves (Fig. 27). However, of these resolutions only the complete execution of all the J^- moves can result in the maximum number, $2n$, of double points and yields a predictable Gauss diagram (Fig. 28). We call this the *enlightened* resolution of the n-Master.

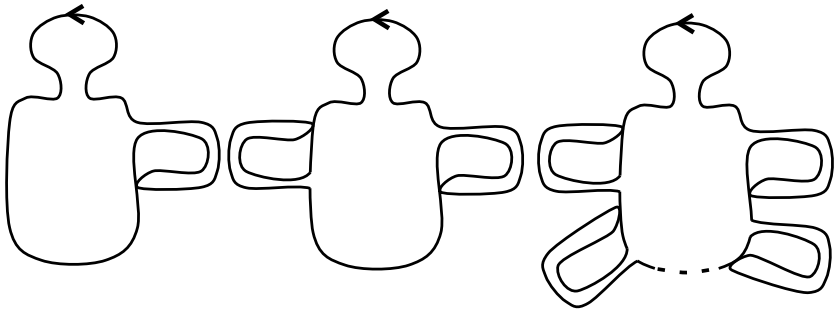


FIGURE 26. From the left to right, the 1-armed, 2-armed, and n-armed Zen Masters.

Now we define our invariant. Consider the non-based chord diagram corresponding to the enlightened resolution of the n-Master; this diagram has degree $2n$. We see that there are four possible

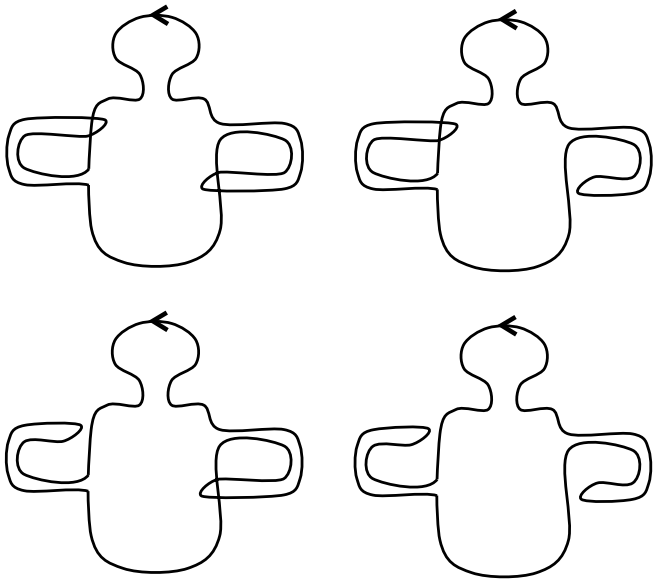


FIGURE 27. The four possible resolutions of the 2-Master. Note the top-right and bottom-left are diffeomorphic, and that the top-left resolution is enlightened.

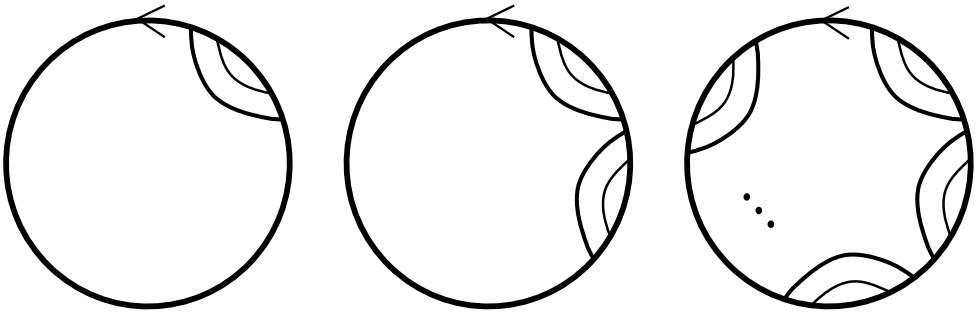


FIGURE 28. From left to right, the signed Gauss diagrams for the 1, 2, and n-Masters with enlightened resolutions. Darkened chords are positively signed.

inequivalent positions for the base point and we identify them as different chord diagrams, as is depicted in Figure 29: A_{n_1}, A_{n_2}, A_{n_3} and A_{n_4}. We then define the linear combination

$$A_n = A_{n_1} - A_{n_2} + A_{n_3} - A_{n_4},$$

which also has degree $2n$, as it is the linear combination of chord diagrams with degree $2n$.

Given a plane curve, Γ, with a Gauss diagram G, we can finally define the invariant $Z_n(\Gamma) = \langle A_n, G \rangle$ which we will refer to as the n-th *Zen invariant*.

Again, recall our unresolved n-Master which we will let correspond to the plane curve Γ_n. Let Γ'_n correspond to the enlightened resolution. We want to show that $Z_n(\Gamma_n) \neq 0$, and thus that the degree of the n-th Zen invariant is at least n from definition 7.3. Normally, to calculate the invariant of of a curve with n self-tangency points, we have to expand it to a sum of the invariants of the 2^n

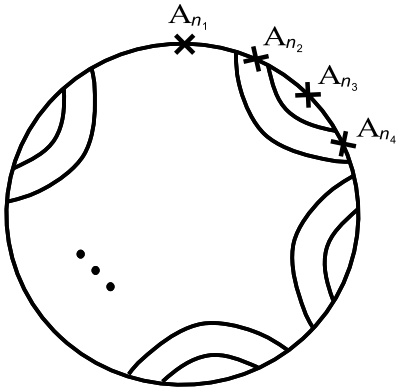


FIGURE 29. Differentiating by base point, the four different based chord diagrams of $2n$ chords corresponding to the enlightened resolution of the n-Master.

resolutions. However, all but one of these resolutions give rise to Gauss diagrams with fewer than $2n$ chords, and thus give zero values for the n-th Zen invariant, excepting exactly the enlightened resolution. So we see that $Z_n(\Gamma_n) = Z_n(\Gamma'_n) = \pm 1$, as there is only one allowed mapping for exactly one of the four A_{n_i} chord diagrams to the Gauss diagram of Γ'_n such that the base points are aligned, and so we have the desired result.

Now from theorem 7.4, if we can show that the n-th Zen invariant is independent of the choice of base point for any plane curve, Γ, then it is an invariant of plane curves of degree less than or equal to n. Since we know the n-th Zen invariant has at least degree n, this would verify that it is exactly of degree n.

We see that moving the basepoint only matters when it passes through some chord in some $\phi(A_{n_i})$, which changes the sign of each mapping, but also cyclically interchanges the terms corresponding to A_{n_1}, A_{n_2}, A_{n_3} and A_{n_4} so $Z_n(\Gamma) = \langle A_{n_1} - A_{n_2} + A_{n_3} - A_{n_4}, G_\Gamma \rangle$ is preserved (see Fig. 30). Thus the n-th Zen invariant is an invariant of degree n. Similarly defined n-th order invariants can also be constructed, perhaps even ones of practical value.

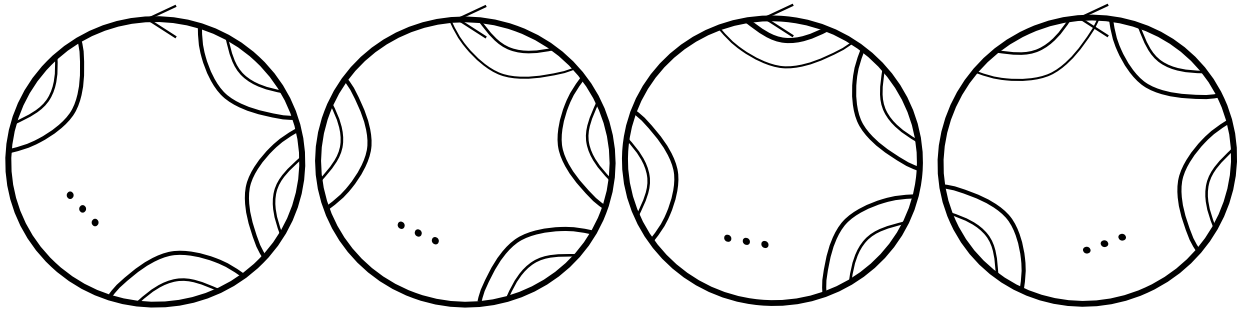


FIGURE 30. In this example, given the simple signed Gauss diagram with at least $2n$ chords corresponding to some A_{n_i}, we see that moving the base point corresponds to changing the sign of the mapping, but also cycles the corresponding chord diagrams, preserving the n-invariant. Darkened chords are positively signed.

8. Plane Structures

Lacking a sufficient complete invariant for plane curves, we are motivated to construct our own. Recall that a signed Gauss diagram is a complete invariant for all spherical curves, also recall that the choice of the unbounded region on a spherical curve will uniquely define its projection as a plane curve and that all plane curves can be represented in this way. From this definition we derive the following:

Lemma 8.1. *Every edge of a spherical curve corresponds to the boundary of exactly two regions.*

Proof. Any edge, d, of a spherical curve, Γ has two sides and every region bordered by this edge must be on at least one of these sides. If d bordered more than two regions then two would correspond to the same side and would thus be path-connected, a contradiction. If one region corresponded to both sides then there would exist a path in the compliment of Γ from one side of d to the other, another contradiction as Γ is closed. The result follows. □

Proposition 8.2. *A spherical curve with n double points has $n+2$ regions.*

Proof. A plane curve will always have the same number of regions as its spherical counterpart. Consider all of the standard representatives of all of the homotopy classes of all of the plane curves (Fig. 24) and note that the lemma is inductively true for all of them. Strangeness moves do not create double points and J^+ and J^- moves always create or destroy two regions and two double points, which we demonstrate in Figure 31. Since the proposition is true for every generic plane curve that is homotopic to the standard representatives, the result follows. □

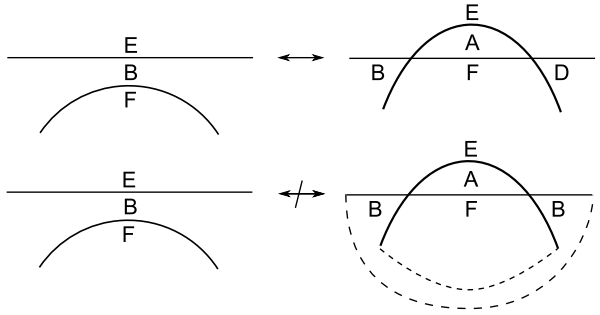


FIGURE 31. The outcome of a J^+ or J^- move (top) on some plane curve Γ results in the creation or destruction of a pair of regions for every pair of double points created or destroyed, respectively. Regions are identified by letters and as every edge corresponds to two different regions, only E and F may be the same. If B is left intact by the creation of two double points (bottom) then Γ is a multi-component curve as the two networks of edges represented by the two dotted lines are disjoint.

Theorem 8.3. *Every region of a nontrivial spherical curve is uniquely defined by the edges that comprise its boundary.*

Proof. Let A be a region in a spherical curve, Γ, with $n \geq 1$ double points. Every double point corresponds to at least three different neighboring regions, each of which shares an edge with one

of the others (Fig. 32). The boundary of A, Γ^A, is comprised of at least one edge and must have at least one double point, p, otherwise it would be trivial, making Γ trivial or multi-component, a contradiction. Suppose Γ^A is also the complete boundary for a different region, B, which is to say $\Gamma^A = \Gamma^B$. Since, thus far, the only described neighbors of p are A and B, there exists a third neighboring region, C. C must share an edge, ε, with A or B, so without loss of generality it shares ε with B. So $\varepsilon \in \Gamma^B, \Gamma^C$ but $\varepsilon \notin \Gamma^A$, lest ε should correspond to more than two regions. So $\Gamma^A \neq \Gamma^B$, a contradiction. Thus the theorem is proven. □

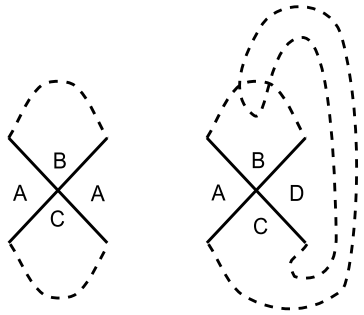


FIGURE 32. A double point can either split a spherical curve into two disjoint parts (left), or it can't (right). In either case the double point neighbors at least three regions.

With theorem 8.3 in mind, we define the following terms:

Definition 8.4. A Gauss diagram, G, is *structured* if some of the undivided arcs between chords on the circles are emphasized. This set of emphasized arcs is a *structure* on G.

Each arc on a signed Gauss diagram uniquely corresponds to an edge on the corresponding spherical curve. Thus we deduce the following corollary.

Corollary 8.5. *If a structure on a signed Gauss diagram corresponds exactly to the boundary of a region on a spherical curve, then that structure is a complete invariant for a plane curve and will be called a* plane structure *(Fig. 33).*

Proof. Theorem 8.3 gives that the boundary uniquely defines a region, and every region of a spherical curve uniquely defines a plane curve. □

Lemma 8.6. *The boundary of any region on a spherical curve can be represented as an oriented path of edges joined by immediate right turns.*

Proof. Let A be any region in some nontrivial spherical curve Γ, and let Γ^A be the set of edges that comprise its boundary. We note that Γ^A itself corresponds to a closed path where every edge connects to some other edge at double points. Pick any edge, ε_0, in Γ^A; we can define an orientation on it - independent of any orientation on Γ - so that A is locally on its right-hand side. Now observe that when we come to the end of ε_0 we can make an immediate right turn onto some edge, ε_1, which must also be in Γ^A and be oriented with A to its right. There are a finite number of edges in Γ, and so we can repeat this a finite number of times until $\varepsilon_n = \varepsilon_m$ for some $n \geq m \geq 0$. Thus we have a

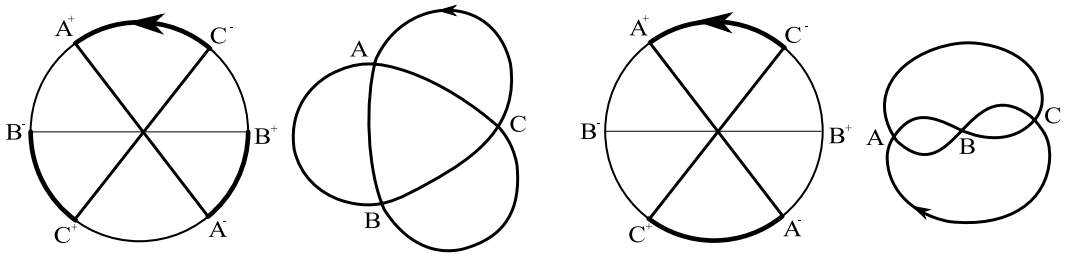


FIGURE 33. A pair of plane-structured, signed Gauss diagrams and their corresponding plane curves. Darkened arcs are corresponnd to plane structures and darkened chords are positively signed.

closed path, Γ^{ε}, from ε_m to ε_n that is joined by immediate right turns and oriented such that A is always kept on one side. Since each right turn along Γ^{ε} prevents the existence of some other edge in Γ^A that is not in Γ^{ε} (Fig. 34), it follows that $m = 0$ and $\Gamma^{\varepsilon} = \Gamma^A$, giving the desired result. □

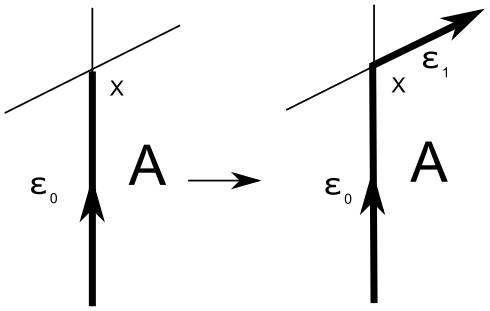


FIGURE 34. Given an oriented edge, ε_0, with the region A to its right, we are capable of making an immediate right turn at the double point, X, such that the subsequent edge, ε_1, is also part of the border of A, Γ^A; we go on in this way to construct Γ^{ε}. The other edges cannot be in Γ^A unless they are also eventually in Γ^{ε}.

This lemma gives rise to a powerful proposition:

Proposition 8.7. *Every plane structure on a signed Gauss diagram corresponds to a walk, defined by motions between adjacent letters, in some direction on the corresponding signed Gauss word if and only if*

(1) *encountering some letter, $a^{\pm}$, will take us to its opposite, $a^{\mp}$*
(2) *encountering positively signed letters maintains the direction of travel*
(3) *encountering negatively signed letters reverses the direction of travel.*

Example 8.8. *Consider the signed Gauss word*

$$\left(a^+\ b^-\ c^+\ d^+\ e^-\ a^-\ f^+\ e^+\ d^-\ g^-\ b^+\ c^-\ g^+\ f^-\right).$$

If we choose d^+ as a starting point and move left we get the walk

$$\left(a^+ \overset{3}{\leftarrow} b^-\ c^+ \overset{1}{\leftarrow} d^+\ e^- \overset{4}{\leftarrow} a^-\ f^+\ e^+ \overset{5}{\rightarrow} d^-\ g^-\ b^+ \overset{2}{\leftarrow} c^-\ g^+\ f^-\right).$$

This corresponds to the plane structure and plane curve in Figure 35.

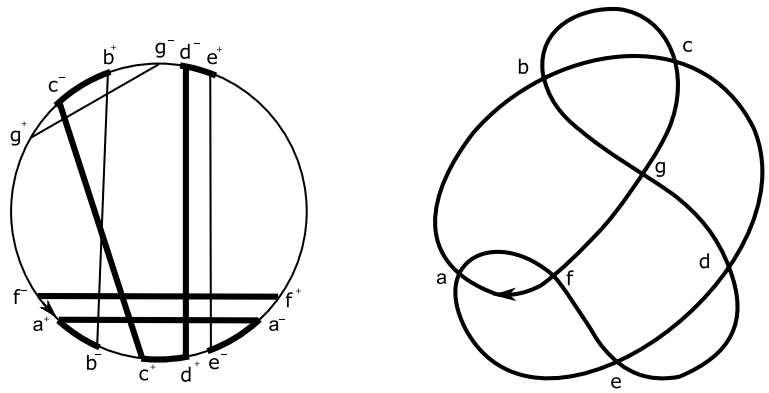


FIGURE 35. A plane-structured, signed Gauss diagram corresponding to the walk in example 8.8 (left) and the identified plane curve (right).

We see that proposition 8.7 is rather useful as any signed Gauss word, and thus also any signed Gauss diagram, can easily classify every corresponding plane curve by finding $n+2$ different walks. As we've already demonstrated that any unsigned Gauss diagram has a limited and knowable number of signings, we have successfully reduced the problem of plane curve classification to the planarity problem of chord diagrams.

of proposition 8.7. : Suppose we have a signed Gauss diagram, G, and Γ is the spherical curve it defines. Let W be a signed Gauss word corresponding to G and let A be a region in a nontrivial spherical curve, Γ, with boundary Γ^A. Note that changing the base point or reversing the orientation on G will not change the signing of any of the letters in W, it merely alters how we would traverse it - where we start and which direction we travel, respectively - if it was to correspond to an oriented journey around the containing circle of G.

Lemma 8.6 tells us that we can orient Γ^A without regard for the orientation of Γ such that Γ^A is made up of right turns and A is always to the right of every edge in Γ^A. Consider any given right turn on Γ^A; it is always a move from the ingoing to the outgoing crossing, or outgoing to ingoing, of a double point. Since Γ^A gives rise to a plane structure on G, we see that switching crossings allows our structure to describe a walk around the circle of G where subsequent arcs are joined by the opposite ends of chords. Each arc can be represented as a movement from one chord endpoint to an adjacent chord endpoint, and so we can represent this walk on G as a unique walk on W where at least condition (1) of the proposition is satisfied.

Now consider a right turn in Γ^A with respect to every possible based orientation of Γ, this is explicitly depicted in Figure 36. We can see that the orientation of subsequent edges due to the orientation on Γ is unchanged with respect to the orientation on Γ^A if and only if the move corresponds to a jump from some positively signed letter, a^+, to some negatively signed letter, a^-, on W. Similarly the orientation of subsequent edges due to the orientation of Γ is reversed with respect to the orientation of Γ^A if and only if the move corresponds to a jump from a negatively

signed letter, a^-, to a positively signed letter, a^+, on W. Thus we have represented the route taken by Γ^A in terms of the signing of W and the orientation of Γ, the later of which is arbitrary and independent of W. Thus pick any edge, ε, in Γ^A and orient Γ such that the orientation of of Γ coincides with the orientation of Γ^A on ε. This corresponds to choosing a directed path between two adjacent points in W. Since the orientation of Γ is fixed, we can now use the signing of W to determine the direction of travel for each move between letters to completely define Γ^A by following rules (2) and (3) of the proposition. Thus every region corresponds to one of these walks. Conversely, since every edge corresponds to two different regions, each "to the right" depending on its orientation, we get that every walk satisfying the rules of the proposition will correspond to a plane structure regardless of the starting point and direction. Thus the proposition is proven. □

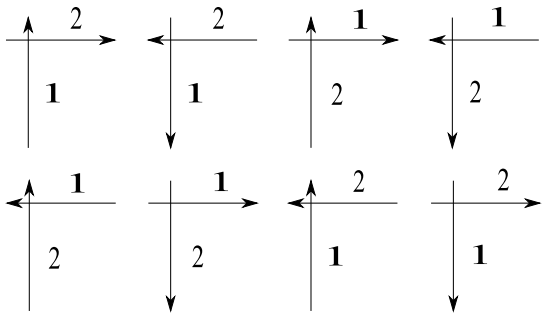


FIGURE 36. A right turn in Γ^A with respect to every possible based orientation of Γ. In some, the orientation of the edges due to the orientation on Γ is unchanged with respect to the orientation on Γ^A (top). In the others the orientation of the edges due to the orientation of Γ is reversed with respect to the orientation of Γ^A (bottom). All correspond with the rules of proposition 8.7.

9. SEIFERT STRUCTURES AND WHITNEY INDEX

The plane structure is not the only useful structure we can impose on a signed Gauss diagram. Consider first the following definition:

Definition 9.1. [Luo] On an oriented spherical curve, Γ, the *Seifert splitting* at a double point is an orientation-preserving cut (Fig. 37). If Γ is completely decomposed in this way, the *Seifert cycles* are the disjoint oriented loops that comprise the decomposition (Fig. 38).

Each Seifert cycle is composed of edges from the parent spherical curve, and each edge corresponds uniquely to some cycle which we can determine by taking immediate orientation preserving turns. So we are compelled to introduce the *Seifert structure*:

Definition 9.2. The *Seifert structure* is a structure on a signed Gauss diagram that corresponds uniquely to a Seifert cycle in the corresponding spherical curve.

Proposition 9.3. *A structure is a Seifert structure if and only if it corresponds to a walk defined by motions between adjacent letters in some direction on a signed Gauss word where encountering some letter, $a^\pm$, will take us to its opposite, $a^\mp$ and the direction of travel is maintained.*

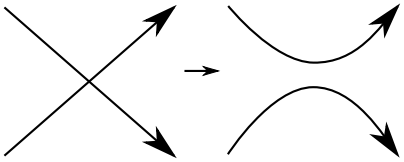


FIGURE 37. The definition of Seifert splitting for spherical curves.

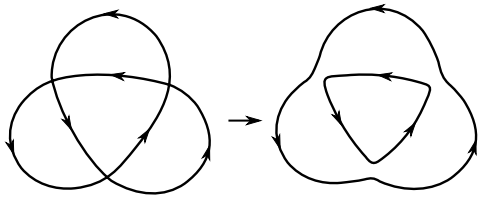


FIGURE 38. The decomposition of the trefoil into Seifert cycles.

Proof. Any structure on a signed Gauss diagram, G, can be uniquely represented as a series of motions between adjacent letters on the corresponding signed Gauss word, W and the converse is also true. As with plane structures, every resolution of a double point by a Seifert splitting corresponds to a crossing-switch on the corresponding Gauss diagram, so the series of motions on W that corresponds to some Seifert cycle comprises a walk where encountering some letter, $a^{\pm}$, will take us to its opposite, $a^{\mp}$. By definition 9.2, orientation is maintained with each splitting, so the direction of the walk must be maintained from letter to letter. Thus the Seifert cycle always corresponds to a walk described by the proposition. Since every motion between letters on W corresponds to an edge on the spherical curve, and any direction-preserving jump between opposite letters corresponds to an immediate left or right turn that preserves orientation, we have that the walk of this form must correspond to a Seifert structure. Thus the proposition is proven. □

Note that the regions of spherical curves compose these *provinces* partitioned by the complement of the Seifert cycles. Each Seifert cycle borders two provinces, the provinces that *belong* to it. To better describe the relationship between regions and provinces, we require the following definitions illustrated in Figure 39.

Definition 9.4. Two regions in a spherical curve are *joined* if they are in the same province.

Definition 9.5. Two regions in a spherical curve are *adjacent* if they share a common double point.

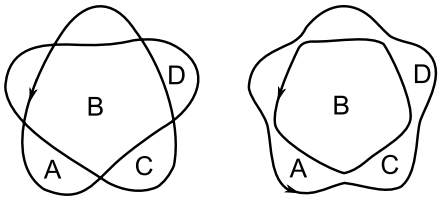


FIGURE 39. A spherical curve with the regions A, B, C, and D identified (left) and the corresponding Seifert decomposition (right). We see that A and C are joined and adjacent, A and B are adjacent, and A and D are joined.

The provinces are path connected, and thus the term *joined* describes an equivalence relation on the regions. Before we give a proposition describing a useful relationship between joined and adjacent regions, we require the following lemma.

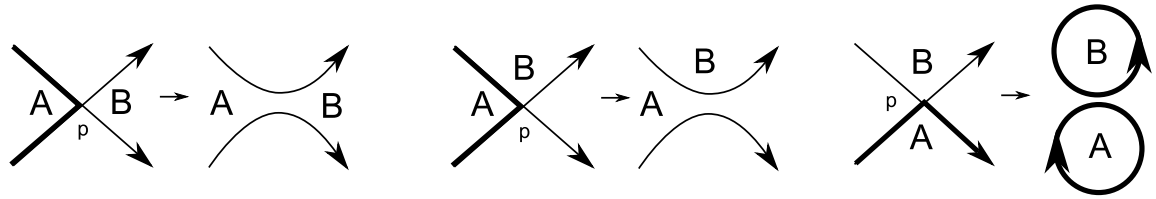


FIGURE 40. If the shared double point, p, is orientation reversing for the boundary of both regions, A and B, the regions are joined (left). If p is orientation reversing for just the boundary of A, then A and B are not joined (middle). If p is orientation preserving for the boundary of both regions, the regions are not joined (right).

Lemma 9.6. *On any nontrivial oriented spherical curve, every double point corresponds to exactly two Seifert cycles.*

Proof. Consider an oriented, nontrivial spherical curve, Γ; each double point used for a Seifert cycle corresponds to two edges in that cycle, and each edge will be used for only one Seifert cycle. Thus any double point must only correspond to one or two Seifert cycles. Suppose the double point corresponds to only one Seifert cycle, then Seifert split every crossing but this one. The the cycle corresponding to this last crossing must be diffeomorphic to an oriented spherical curve with only one double point, of which there is one: the figure eight. The figure eight decomposes into two disjoint Seifert cycles, contradicting the property of the crossing. Thus the lemma is proven. □

Now the proposition:

Proposition 9.7. *On any nontrivial oriented spherical curve, adjacent regions are joined if and only if every shared double point is orientation-reversing on the corresponding plane structure for both regions.*

Proof. Given an oriented spherical curve, Γ, with a signed Gauss diagram, G, consider two adjacent regions, A and B, and a shared double point, p. If the orientation of the the boundary of A, Γ^A, reverses with respect to the orientation of Γ at p, then Γ^A corresponds to both Seifert cycles at p and A is thus joined to the other adjacent region for which p holds the same condition. We see that both of these borders correspond to a plane structure in which p is orientation-reversing, in that, for each, the corresponding arcs connected to the p-chord on G are oppositely oriented on the circle. If B is this other adjacent region, then A and B are joined, if not then A and B cannot be joined because they share an edge which must correspond to a Seifert cycle boundary between them. Conversely, if p is not orientation-reversing for either Γ^A or Γ^B, then, by lemma 9.6, each boundary resolves to a separate Seifert cycle at this point and thus Γ^A and Γ^B are not joined, as is illustrated in Figure 40. This proves the proposition. □

Two regions can only be joined if there is a series of joined, adjacent regions between them, so it follows from proposition 9.7 that the set of regions comprising a province can be completely determined and represented by the corresponding set of plane-structured signed Gauss diagrams. By comparing the arcs of the plane structures to the arcs of the Seifert structures, we are capable of determining which provinces belong to which Seifert cycles, and thus whether or not any two Seifert

cycles share a common province on the sphere. This motivates a definition of distance between cycles and provinces in the Seifert decomposition, given by the least number of Seifert cycles that must be crossed to get from one to the other. When distances are inequivalent, we have the comparative terms, *closer* and *farther*, defined in the obvious way. Also Seifert cycles are *connected*, meaning they have a shared double point, if and only if their corresponding Seifert structures correspond to at least one of the same chords on the Gauss diagram. Since every spherical curve is a connected space, it follows that every Seifert cycle can be related to another via a series of connections. Thus we are able to completely reconstruct an unoriented Seifert decomposition on the surface of the sphere.

When a Seifert decomposition is projected onto the plane, we see that the unbounded province of the decomposition always contains the unbounded region of the corresponding plane curve. This defines a concept of *interior* and *exterior* provinces belonging to a Seifert cycle, where the exterior province is closer to the unbounded province, and the interior province is further. A Seifert cycle, $\mathcal{A}$, is *inside* another Seifert cycle, $\mathcal{B}$, if $\mathcal{A}$ is closer to the interior province of $\mathcal{B}$ than the exterior. $\mathcal{A}$ is *outside* $\mathcal{B}$ if $\mathcal{A}$ is closer to the exterior province.

Remark 9.8. Every cycle is either inside or outside any other cycle and no two cycles are inside each other.

Each cycle, c, can be assigned an index: $+1$ if c is counterclockwise and -1 if c is clockwise. Barker and Biringer[BB] noted that the index of the plane curve could be calculated by the formula

$$\text{index}(\Gamma) = \sum_{\text{Seifert cycles } c\in\Gamma} \text{index}(c).$$

Note that this is the index of the oriented curve, so it can be positive or negative; we must take the absolute value of this number to get the index of the unoriented curve.

If a Seifert cycle, $\mathcal{A}$, is connected to another Seifert cycle $\mathcal{B}$, then they are oriented oppositely if and only if they are outside one another. Using our plane and Seifert structures to reconstruct an unoriented Seifert decomposition on the surface of the sphere, we can choose an unbounded province and project it onto the plane, noting when cycles are inside and outside one another. Once this is done, we can orient any one cycle arbitrarily - as reversing orientation on the original curve will reverse the orientation of this cycle - and use the relations due to connections described above to orient every other cycle. We then use Barker and Biringer's formula to calculate the index, which will in turn give the index of every plane curve whose exterior region comprises this province. Moving the unbounded region across a Seifert cycle, $\mathcal{A}$, of a plane curve will also move the unbounded province, swapping interior and exterior provinces of $\mathcal{A}$ so that all cycles inside of $\mathcal{A}$ will be outside and vice-versa. Thus the orientation of $\mathcal{A}$ must reverse and there will be a ± 2 change to the index of the oriented curve, depending on the original orientation of the cycle. Proceeding in this manor, we can move all around a Seifert decomposition and assign indices to the provinces and thus find the index of every plane curve corresponding to every plane structure on our original signed Gauss diagram. From this we can use Polyak's equations [Po] in theorem 7.2 to quickly calculate the Arnold invariants for every plane curve in this family, as well as calculate any invariant desribed in terms of Polyak's function.

Example 9.9. *In Figures 41 and 42 we have a worked example of an application of propositions 8.7, 9.7, and 9.3, Barker and Biringer's formula, and other understood relations of Seifert cycles, which are discussed above.*

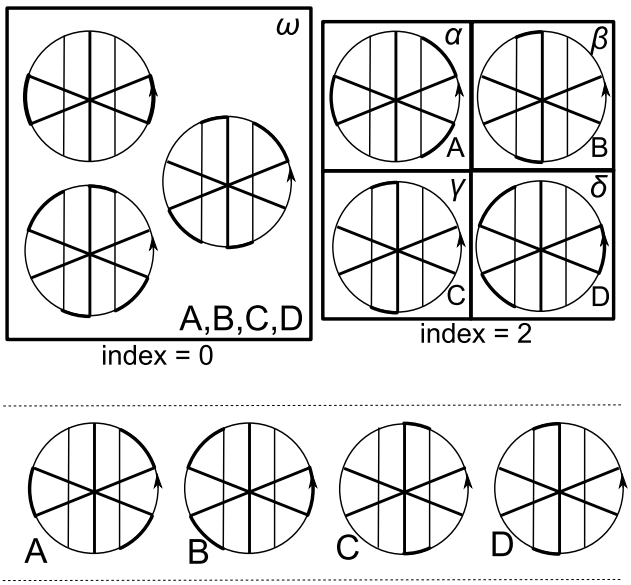


FIGURE 41. Boxed at the top are all of the plane structures for a signed Gauss diagram derived from proposition 8.7. From proposition 9.7 we were able to partition the set of structures to their corresponding provinces, $\alpha, \beta, \gamma, \delta, \omega$, which are indicated by the labeled boxes. On the bottom are the four Seifert structures corresponding to cycles $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$. The roman letters in the partitioning boxes indicate which cycles border which province, and the Seifert structures can be used to determine which cycles are connected. We use this to construct Figure 42 and from that we can use Barker and Biringer's formula to calculate the index associated with each province. This gives us the index of the unoriented curves associated with each plane structure, labeled below the boxes.

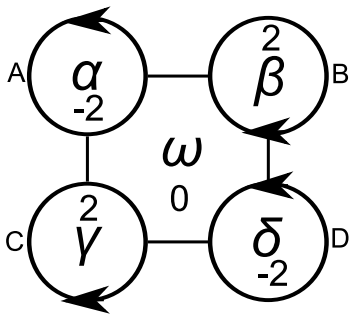


FIGURE 42. The Seifert cycles derived from Figure 41.

Part 6. **Symmetries**

10. COMPOSITE CURVE CLASSIFICATION AND INVARIANT EQUIVALENCE

Considering only unoriented spherical curves and maintaining our claim that signed Gauss diagrams are equivalent under *symmetric moves* - opposite signing, orientation reversal, base point movement and reflection - we must still consider the signings of relative components. Given a composite Gauss diagram with n components, the maximum number of unique signings, and thus unique spherical curves, is 2^{n-1}. Allowing for certain symmetries, this number may be reduced.

Remark 10.1. Signings are equivalent if and only if, through any combination of the symmetric moves, one signed Gauss diagram can be changed into another (Fig. 43).

The truth of this remark follows trivially from how we defined equivalent unoriented curves.

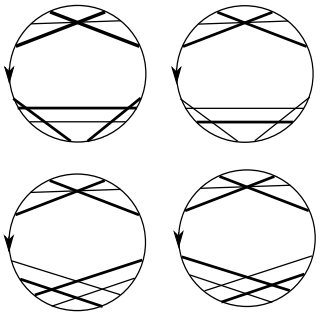


FIGURE 43. Two different Gauss diagrams with the same components (top and bottom) and the two different potential signings of each of them (left to right). Darkened and undarkened lines are oppositely signed. We see that for the top Gauss diagram the signings are in fact inequivalent, as one cannot be transformed into the other via the symmetric moves, and so the spherical curves are also inequivalent. For the bottom Gauss diagram, the signings can be transformed into each other by the symmetric moves, giving mirror reflections of the same curve.

Along similar lines, we have until now ignored that "different" projections of spherical curves - by choosing "different" unbounded regions - occasionally yield equivalent plane curves. We say such regions are "different" in that we can make them distinct with some sort of arbitrary labeling, or perhaps differentiate then by the orientation of the curve itself, but they have a *symmetric equivalence* in that they cannot be differentiated by the curve.

Proposition 10.2. *Regions are symmetrically equivalent if and only if their corresponding plane structured, signed Gauss diagrams can be made equivalent through symmetric moves.*

Proof. This is a trivial observation of the fact that equivalent plane curves are equivalent if and only if they have equivalent plane structured, signed Gauss diagrams, and symmetric moves extend this notion of equivalence. □

Corollary 10.3. *If a plane structure can be reflected or rotated into another plane structure in a prime Gauss diagram, then those structures correspond to symmetrically equivalent regions (Fig. 44).*

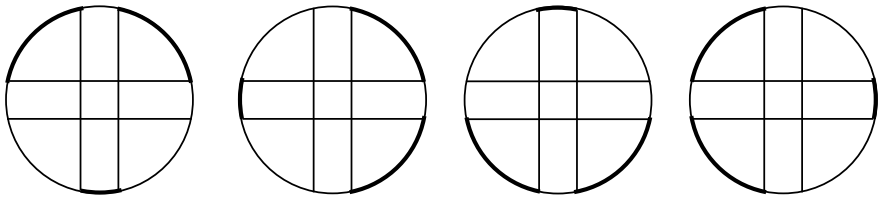


FIGURE 44. Coexisting plane structures on a prime Gauss diagram that are equivalent by proposition 10.3. Chord signing is suppressed as only one signing is permitted.

.

Proof. This is an observed result of corollary 2.3 and proposition 10.2 as only one allowed signing on some Gauss diagram G guarantees equivalence of every signing G for all plane structures. □

11. SYMMETRY CLASSES

We now utilize a matrix form of the Gauss word to examine the transformations of spherical curves induced by symmetry transformations on the sphere.

Definition 11.1. The Gauss matrix W_b of a curve with n double points is a $2n$ dimensional square matrix with entries corresponding to the Gauss word of the curve with base point b. The set of Gauss matrices of curves with n double points is denoted by G^{2n}.

More specifically, each column corresponds to an occurrence of a double point of the curve. The outgoing branches of the double points of the curve are labeled x and y such that $x \times y$ is positive. If the j-th branch intersects the ith branch, it is entered in the i-th row of the jth column, as either $+1$ or -1, $+1$ to indicate it is a y branch, or -1 to indicate it is an x branch. The result is an anti-symmetric matrix with only one non-zero entry in each row and each column.

Definition 11.2. The **base point transformation**, $\mathcal{P} : W \longrightarrow W$ shifts the base point a Gauss matrix by one letter in the direction of orientation, sending W_b to W_{b+1}. Let

$$B := \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

and let e_q is the q-th standard basis vector in $\mathbb{R}^{2n}$. The transformation is given by

$$\mathcal{P} \circ W_b = [e_{2n}, e_1, \cdots, e_{2n-1}]\,[W]_b\,[e_2, e_3, \cdots, e_{2n}, e_1] \tag{5}$$

and the inverse, which shifts the base point by one letter opposite the direction of orientation, is given by

$$\mathcal{P}^{-1} \circ W_b = [e_2, e_3, \cdots, e_{2n}, e_1]\,[W]_b\,[e_{2n}, e_1, \cdots, e_{2n-1}]\,. \tag{6}$$

Definition 11.3. The **folding transformation**, $\mathcal{F} : G^{2n} \longrightarrow G^{2n}$ is the transformation given by

$$\mathcal{F} \circ W_b := B\,[W]_b\,B \tag{7}$$

Note that the transformation is "base point dependent", meaning that $\mathcal{F}$ does not commute with $\mathcal{P}$. Also, note that since

$$B^2 = [I]\,, \tag{8}$$

it follows that, for a Gauss matrix W_b,

$$\mathcal{F} \circ \mathcal{F} \circ W_b = B^2\,[W]_b\,B^2 = [I]\,[W]_b\,[I] = [W]_b \tag{9}$$

$\mathcal{F} \circ \mathcal{F} \circ W_b = W_b$, so, $\mathcal{F}^{-1} = \mathcal{F}$.

Definition 11.4. The orientation transformation ,$\mathcal{T}$, is given by

$$\mathcal{T} \circ W := -\,[W]\,. \tag{10}$$

Note

$$\mathcal{P}^{-1} \circ \mathcal{T} \circ \mathcal{P} \circ W_b = \mathcal{T} \circ W_b, \tag{11}$$

i.e., $\mathcal{T}$ is *not* base point dependent. Also note that, as expected, $\mathcal{T}^{-1}$=$\mathcal{T}$.

$\mathcal{T}$ does not correspond to the standard notion of "orientation change" which is usually used to refer to reversing the direction of the parametrization. The transformation which reverses the orientation of S^1 in the domain of the immersion but leaves the standard orientation of the plane unchanged is given by the product $\mathcal{T} \circ \mathcal{F}$.

Definition 11.5. Let γ be a parametrization from the circle to the sphere and consider the associated set of immersions $\{\gamma\}$ generated by γ under different orientations of the circle and sphere,

$$\begin{aligned} S^1_+ &\xrightarrow{\gamma} S^2_+ \\ S^1_- &\xrightarrow{\gamma} S^2_+ \\ S^1_+ &\xrightarrow{\gamma} S^2_- \\ S^1_- &\xrightarrow{\gamma} S^2_-\,. \end{aligned} \tag{12}$$

$\widetilde{W}_{\{\gamma\}} \subset G^{2n}$ is the set of signed Gauss matrices of the set of immersions $\{\gamma\}$.

Remark 11.6. The set of transformations $\mathcal{P}$, $\mathcal{F}$, $\mathcal{T}$, and I form a group under composition of action on $\widetilde{W}$.

Proof. It has already been shown that the inverse of each transformation is a power of itself. To see that $\widetilde{W}$ is closed under the transformations, consider the implications for the spherical curve corresponding to each transformation. As was discussed in the previous sections, Polyak [Po] has shown that, given a fixed orientation of the sphere, the signed Gauss word uniquely determines a spherical curve. Clearly, the curve determined by the word does not depend on the choice of a base point, so any cyclic permutation ($\mathcal{P}$) of the letters in the word preserves the curve. Reversing the ordering of the letters of the word without changing their respective signs ($\mathcal{F}$) corresponds to the immersion from the circle to the curve with the same image on the sphere but with opposite orientation. On the other hand, switching the sign of each letter while leaving the ordering unchanged ($\mathcal{T}$) does not correspond to the immersion with the same image on the sphere. Rather, the resulting

curve is the image on the oppositely oriented sphere, that is, the mirror image of the original curve reflected in the plane perpendicular to the orienting pole of the sphere. □

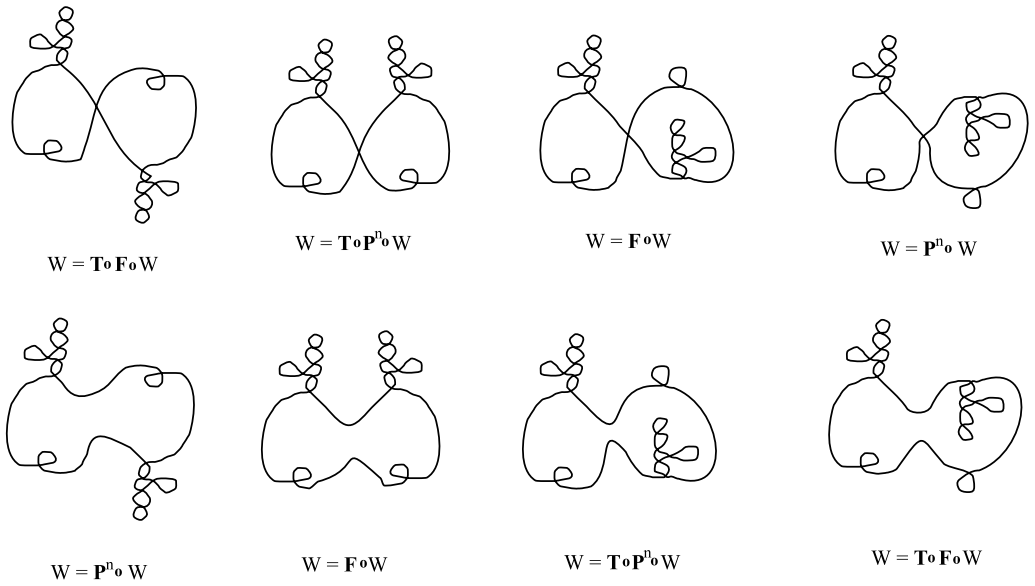


FIGURE 45. Symmetric curves generated from an asymmetric Gauss matrix with fixed base point. The top row shows the action of the transformations when the base point is a double point. The bottom row shows the action when the base point is a connected sum.

12. CONDITIONS FOR ROTATIONAL SYMMETRY

Identifying the symmetry type of a curve relies heavily on locating specific base points of the Gauss word. Consequently, as the number of double points of curves increases, the guess and check method becomes increasingly inefficient. The process is greatly simplified by exploiting the conditions for rotational symmetry. Rotational symmetry is the coarsest type of symmetry—that is, if a curve with rotational symmetry possesses an additional symmetry, the other symmetry is contained within the segment which generates the curve under rotation. By identifying the rotational symmetry first, the generator can be projected to a lower dimensional Gauss matrix without destroying the other symmetries. Identifying the other symmetries from the reduced matrix expediates the process by a factor of the order of symmetry m.

Procedure 12.1. *The set of plane curves obtained from a spherical curve with Gauss word W contains a curve with rotational symmetry of order m if and only if the word is composed of m substrings, $L_1,\ldots,L_m$, of equal length, which satisfy the condition*

$$(L_j^i, L_p^k) \Longrightarrow (L^i_{[j+n]mod(m)}, L^k_{[p+n]mod(m)}). \tag{13}$$

Proof. Let γ be a plane curve with rotational symmetry of order m. Pick a starting point $\gamma(0)$ and consider the set of points.

$$\{\gamma(0), R\circ\gamma(0), R^2\circ\gamma(0), ..., R^{m-1}\circ\gamma(0)\},$$

that is, the m distinct images of $\gamma(0)$ under consecutive powers of R, where R is the rotation of the plane which maps the curve to itself. Since the curve has rotational symmetry, these points all lie on the curve γ. Starting from $\gamma(0)$, trace along the curve until it arrives at one of the image points $R^i\circ\gamma(0)$, and denote by L_1 the segment of the curve between $\gamma(0)$ and $R^i\circ\gamma(0)$. Continue following the curve until encountering the next image point $R^{2i}\circ\gamma(0)$, and call this second leg of the journey L_2. Since the starting point of L_2 is $R^i\circ\gamma(0)$, and since R^i is continous and conformal, we have $L_2 = R^i(L_1)$. Similarly, for L_3 (the segment following L_2), we have $L_3 = R^i(L_2) = R^{2i}(L_1)$. Continuing in this manner, γ is divided into m sub-segments, $\gamma = L_1 \cup L_2 .. \cup L_m$, and each is the image of L_1 under a rotation.

The rotations preserve the structure of L_1, so each segment contains the same number of double points as are contained in L_1. Consequently, the Gauss word for the curve γ contains m substrings of equal length—one substring for each segment. To form the word, first, trace each segment of the curve in order, labeling the i-th double point encountered on the j-th segment by L^i_j. After labeling every double point twice (once for each of the intersecting branches), construct the Gauss word by re-tracing the curve, and, at each double point, listing both labels as an ordered pair, (L^i_j, L^k_p), with the first coordinate designated for the segment currently underfoot. The resulting Gauss word is a string of ordered pairs $(L^1_1, L)(L^2_1, L)(L^3_1, L)...(L^1_2, L)(L^2_2, L)....$ For any segment L_i, the rotation which maps L_1 onto L_i also maps L_2 onto L_{i+1}, L_3 onto L_{i+2} and so on. It follows that the sequence formed from the second coordinates in the substring corresponding to L_i is isomorphic (with respect to order modulo m) to the sequence of second coordinates in the first substring, L_1. □

Corollary 12.2. *The hypotheses of condition 12.1 are satisfied if and only if the word satisfies the equivalent condition.*

$$(L^i_j, L^k_{j+n}) \Longrightarrow (L^k_j, L^i_{[j-n]mod(m)}). \tag{14}$$

Procedure 12.3. *For any curve which is mapped under an ambient diffeomorphism to a curve with rotational symmetry of order m, the number of double points is a multiple of m.*

Proof. Suppose not. Since the number of double points is invariant under diffeomorphism, the hypothesis implies the existence of a curve which is invariant under rotation and contains a number of double points which is not a multiple of m. This implies the existence of a non-empty point set for which rotation is not a one-to-one transformation. □

Procedure 12.4. *For any curve with non-zero rotation index K which is mapped under an ambient diffeomorphism to a curve with rotational symmetry of order m, the following condition must hold.*

$$gcd\,(K, m) = 1. \tag{15}$$

Proof. If a curve has rotational symmetry of order m, it is composed of m identical legs $L_1, ..., L_m$, and each leg contributes equally to the total rotation index of the curve. Since rotation is conformal, the endpoints of the legs are connected such that the tangent vector remains continuous when passing from one leg to the next. It follows that $\int_{L_i}\ddot{\gamma}$, the total rotation of the tangent on each leg, occurs on the interiors of the legs, and the total index of the curve is the sum of the contributions

from each leg.

$$K = m \int_{L_i} \ddot{\gamma} ds. \tag{16}$$

It follows that the total rotation of the tangent which occurs on each leg is $\frac{K}{m}$. Now, suppose the index shares a common factor c with m. This implies that $c\frac{K}{m}$ is an integer multiple of a full rotation and that the endpoint L_c return to the starting point of L_0, thus, closing the curve prematurely. By contradiction, the index of the symmetric curve is coprime to the order of symmetry m. Since rotation index is preserved under diffeomorphism, it follows that the condition extends to all diffeomorphisms of a symmetric curve. □

Theorem 12.5. *The class of plane curves with non-zero index K and n double points contains curves with rotational symmetry of order m if and only if m, K and n satisfy conditions 12.1 and 12.3, and*

$$n \geq m \left| \left(|K| - m + [|K|-1]_{mod(m)} \right) \right| - [|K|-1]_{mod(m)} \tag{17}$$

The proof of the lower bound condition makes use of the following arguments which we establish first.

Lemma 12.6. *Define $R_{p,\phi}$ as the rotation of the plane about point p by angle ϕ and let γ be a curve in the plane which is mapped to itself by $R_{p,\phi}$. γ contains the point p if and only if γ has index $K = 0$.*

Proof. Suppose the fixed point p is contained in γ. First assume that the fixed point does not lie on a double point of γ. Under a rotation of order greater than 2, the segment passing through the center point is not mapped to itself, so we only need to consider a rotation of order 2. Let l by the line passing through the fixed point p which intersects γ perpendicularly. l must either intersect γ at an even number of points or it intersects the curve tangent to the curve or through a double point. If the other points of intersection of l with γ are at points of tangency, there is a local diffeomorphism of the curve which removes these intersection points and preserves the rotational symmetry of γ, so, without loss of generality, we may assume that the points of intersection of the line with the γ are not points of tangency. In addition to the fixed point p, the two halves of γ must contain the same number of intersection points with γ, for each half is mapped to the other by $R_{p,\phi}$. Then, including the fixed point, the total number of points of intersection of l with γ must be odd. This implies that the set of points of intersection must contain an odd number of double points and, thus, the two half rays of l cannot contain the same number of points of intersection, and γ cannot possess order 2 rotational symmetry.

Next, assume that the fixed point, p, lies on a double point of γ. The possible rotations which map the branches of the double point to each other are rotations of order 4 or order 2. The outgoing branches of the double point are uniquely determined by the orientation γ, but under a rotation of order 4, the outgoing branches of the double point are mapped to four distinct configurations, which would imply γ had four distinct orientations. The contradiction shows that γ cannot posess rotational symmetry of order four. This only leaves the possibility of γ having rotational symmetry for $1/2$ of a full rotation. The half rotation of the plane maps the outgoing branches to the incoming

branches, so the image of γ under the rotation has reversed orientation. Since the index of γ, K, cannot change under the rotation, this implies $K = -K$, and, thus, $K = 0$. □

The next argument refers to the 'center region index' of a curve with rotational symmetry. The center region index is the winding number of the curve about the fixed point of rotation. As was shown above, the center region index is well defined when $K \neq 0$

Lemma 12.7. *Let γ be a curve in the plane with rotational symmetry of order m, rotation index $K \neq 0$, and center region index k_c. γ is smoothly homotopic to a curve $\widetilde{\gamma}$ which satisfies*

(1) *$\widetilde{\gamma}$ has order m rotational symmetry*
(2) *$\widetilde{\gamma}$ has center region index k_c*
(3) *There is a set of m rays emanating from the center point and spaced at equal angles for which, at every point of intersection with $\widetilde{\gamma}$, the orientation of γ is the same*
(4) *$\widetilde{\gamma}$ has the same number of double points as γ or fewer.*

Proof. Send out a set of m rays from the fixed point of the rotation which maps γ to itself. If the rays satisfy (3), we are done. Otherwise, consider the wedge of the plane bounded by two rays (the rotational symmetry of γ ensures that each wedge is identical), and label the rays x and y so that $x \times y$ agrees with the sign of k_c, the region index of the region containing the fixed point of the rotation. By lemma 12.6, this region exists, and the curve does not intersect the vertex of the wedge. There are at least $|k_c|$ segments of the curve which enter the wedge from x, cross the wedge, and exit the wedge through y. Any additional segments which enter the wedge on one side and leave on the opposite side must come in oppositely oriented pairs. Moreover, the segments which enter and leave the wedge on the same side must also come in pairs. Since the curve is continuous, a segment which enters and leaves on the same ray must contain a point where the curve is tangent to a ray through the center point (a ray between x and y).

Without loss of generality, assume this segment enters and leaves the ray on x. Sweeping the ray in the direction of the normal to the curve at the point of intersection (in this case, towards x) introduces two points of intersection which trace out two branches of the curve. One branch must eventually lead to another point of tangency with the ray which belongs to a segment which leaves and enters the wedge on y. The segment of the curve between the two points of tangency has winding number about the center region equal to zero, so a homotopy h which fixes the two points of tangency but maps the segment between them to segment contained within a single wedge need not pass through the center point and preserves the center region index.

Applying h to the equivalent points of each wedge preserves rotational symmetry, so h satisfies (1). h^{-1} can only increase the number of double points, so $\widetilde{\gamma} := h \circ \gamma$ has fewer or the same number of double points as γ, and h satisfies (4). □

Lemma 12.8. *Let γ be a curve with order m rotational symmetry, total rotation index K, and center region index k_c. If $|K| > 0$, and $k_c = K$, then γ has at least $m(|K| - 1)$ double points.*

Proof. By lemma 12.7, the the curve is homotopic to a curve for which the wedges have k_c segments which enter and leave the wedge on opposite rays. Since $k_c = K$, the number of wedge traversing segments is $|K|$. An arrangement of these segments in which $|K| - 1$ segments cross the wedge in parallel and the remaining segment intersects each one diagonally creates $|K| - 1$ double points. If there was a way for $|K|$ segments to cross the wedge with fewer intersections, identifying

the points on the edges of the ray would yield a curve of index K with fewer than $|K|-1$ double points, which would be a contradiction. □

Lemma 12.9. *The number of double points n achieved by the curves with order m rotational symmetry, index K, and center region index k_c is bounded below by*

$$n \geq m(|k_c|-1)+|K|-|k_c| \tag{18}$$

Proof. By lemma 12.7, any curve with order m rotational symmetry, index K and center index k_c is homotopic to a curve with wedges traversed by only $|k_c|$ segments and all other components of the rotation index K accounted for by $l := (K-k_c)/m$ loops contained between the bounding rays of the wedge. By lemma 12.7, the wedge crossing segments add at least $|k_c|-1$ double points to each wedge. One cannot increase or decrease the index of the curve while preserving the center index without adding a loop, and adding a loop adds at least one double point. So, the wedge configuration has a minimum of $|k_c|-1+l$ double points, making the minimum number of double points contained by the curve $m(|k_c|-1+l)$. Substituting in $l := (K-k_c)/m$ gives the lower bound. □

Proof of Theorem 12.5. The wedge subtends $1/m$ of a full rotation about the center region, so each strand of the wedge contributes one mth of a rotation to the total center index k_c. Adding a loop to a wedge contributes ± 1 full rotation to the index of the curve and does not alter the center region index. Globally, this corresponds to an increment of either $+m$ or $-m$ to the total index of the curve. From this we obtain

$$lm+s=|K|\,, \tag{19}$$

where l is the number of loops per wedge, m is the order of symmetry, s is the number of segments per wedge and K is the total index of the curve. By lemma 12.9, we also have that the minimal number of double points for a particular combination of loops and segments is $m(|k_c|-1)+|K|-|k_c|$. Since minimizing the number segments per wedge minimizes the number of double points, the choice of l which achieves this minimum is $l = m-(|K|-1)_{mod(m)}$. Inserting this expression into the equation for the number of double points for a specific combination of s, m, and l gives

$$\begin{aligned} n &= m(s-1)+l \\ &= m(\big|K-m+[|K|-1]_{mod(m)}\big|-1)+m-(|K|-1)_{mod(m)} \\ &= m\big||K|-m+[|K|-1]_{mod(m)}\big|-(|K|-1)_{mod(m)}. \end{aligned} \tag{20}$$

□

The lower bound condition greatly reduces the number of potential symmetry types that need to be checked using condition 12.1. In some cases the possibility of the curve possessing rotational symmetry is ruled out entirety without even needing to consider this condition.

Part 7. Different Approaches

13. Pendants and Hangability

In [Ar2], Arnold describes an interesting approach to the study of plane curves which involves considering a curve as a combination of two "simpler" curves. We give a slight extension of his construction which provides the motivation for introducing ribbon graphs in section 14. We will mostly deal with equivalence under *homeomorphism* as opposed to under *diffeomorphism*, so, for example, the two sets in figure 46 are considered equivalent. Also, in this section we will deal with *multi-component* curves, i.e. images of several copies of S^1 into $\mathbb{R}^2$.

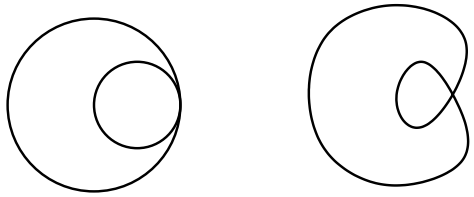


FIGURE 46. Two sets that are equivalent under *homeomorphism* but not under *diffeomorphism*.

13.1. Introduction.

Definition 13.1. [Ar2] Let Γ be a plane curve. The *cactus* of Γ is its exterior contour, that is the part of the curve adjacent to the exterior region (Fig. 47).

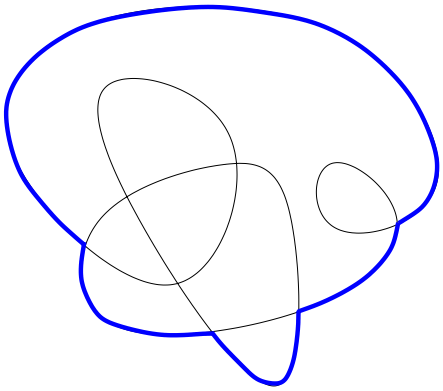


FIGURE 47. A *cactus* of a curve.

Notice that the cactus of a curve is homeomorphic to the boundary of several tangent disks.

On the inside of each disk of the cactus remains a set which is homeomorphic to a (possibly multi-component) curve with fewer crossings than the original curve (Fig. 48). We designate as distinguished the points on each of these curves where there were crossings.

Definition 13.2. [Ar2] A *pendant* is a multi-component curve with n distinguished points on its exterior contour (Fig. 48).

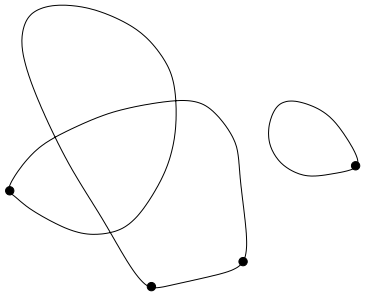


FIGURE 48. A *pendant*.

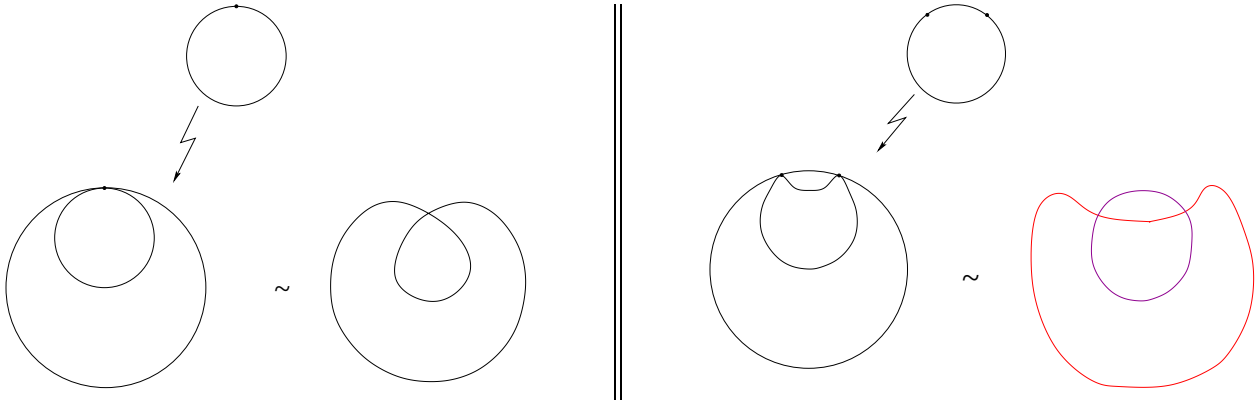


FIGURE 49. An example of hangable and non-hangable curves.

FIGURE 50. The curve A_n consisting of n twists.

Although Arnold did not deal with multi-component curves, this is a straightforward generalization.

Not every pendant can be inside a disk of a cactus of a plane curve. This notion is captured by the following definition.

Definition 13.3. A pendant is *hangable* if it can be placed inside a circle so that the only common poinuts are the distinguished ones and so that the resulting set is homeomorphic to the image of a regular plane curve.

Figure 49 shows examples of hangable and non-hangable pendants.

Arnold tried to classify all hangable pendants, probably hoping to use it to come up with a method to enumerate (or list) all plane curves. However, he succeeded only in the simplest cases, when the curve consists of n twists (Fig. 50). Refer to [Ar2] for an in-depth discussion of these curves.

13.2. **Peeling.** Let Γ be a plane curve. As mentioned above, removing a cactus from Γ leaves a multi-component curve. We can take its own cactus, which will consist of several components,

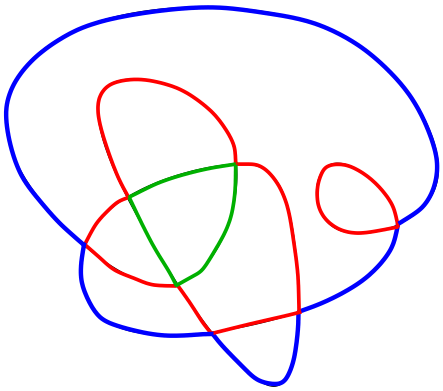


FIGURE 51. A *peeling* of a curve. Level 0 cactus is blue, level 1 is red, and level 2 is green.

each of which is homeomorphic to the boundary of several tangent disks. Repeating this process breaks our curve into "layers". We call this construction a *peeling* of Γ. We define the *level n cactus* of Γ as the n-th layer in the process (Fig. 51).

Notice that each level cactus is homeomorphic to the boundary of a union of several collections of tangent disks. Thusa peeling of Γ puts a structure on Γ as a "connection of several simpler curves." Each double point in Γ corresponds to either a connection between circles withing a cactus of a certain level, or to a connection between cacti. In accordance with this, we have the following definition.

Definition 13.4. A double point of a plane curve Γ is called a *hanging* point if it connects two levels of cacti of the peeling. It is called a *joining point* otherwise.

For example, the curve in figure 51 has 7 hanging points and no joining points.

The peeling construction generalizes the separation of a cactus of a curve from the pendant which was the original motivation for its introduction. Instead of separating the curve into a pendant hanging on the level 0 cactus, we can view it as a pendant formed by all cacti of level $\geqslant k$ hanging inside a region of the curve formed by all cacti of level $< k$, for each k.

13.3. **Unsolved Problems.**

- Classify hangable pendants for some larger set of curves, for example on tree-like curves.
- Can this theory be used to put some sort of an upper bound on the number of plane curves? There should be some kind of a recursion relation since a curve with n double points is formed by putting hangable pendants with $\leqslant n$ distinguished/double points inside a circle. Is there a nice expression for it?

14. RIBBON GRAPH STRUCTURE ON CURVES

The "peeling" construction described in section 13.2 puts a structure on the plane curve by breaking it into several disks and keeping track of the order in which these disks are connected. This structure is precisely captured by the notion of a ribbon graph.

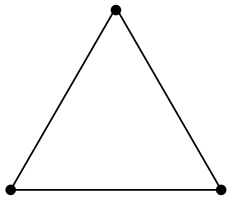


FIGURE 52. A graph.

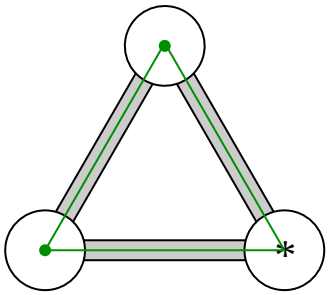


FIGURE 53. A ribbon graph (black) with its underlying regular graph (green).

14.1. **Basic Graph Theory.** Before introducing ribbon graphs we need some notions from classical graph theory. A graph is a pair (V,E) of sets with $E \subseteq V \times V$. Informally, a graph is a set of points (vertices) and some of these vertices are connected by edges (Fig. 52)

An *n-cycle* is a graph with with *n* vertices and *n* edges such that every vertex has 2 edges coming from it (it looks like an *n*-gon). A *tree* is a graph that contains no cycles of any length.

An edge in a graph is called a *loop* if it goes from a vertex to itself. It is called a *bridge* if its removal increases the number of connected components of the graph. It is called *ordinary* if it is neither a bridge nor a loop.

In the following sections, these types of graphs will be called regular to distinguish them from ribbon graphs.

14.2. **Ribbon Graphs.** Informally, a ribbon graph is a "fattened" version of a regular graph. That is the vertices are represented by disks instead of points while the edges are represented by ribbons instead of lines (see figure 53). This construction is equivalent to a regular graph with a fixed cyclic order of edges at every vertex. More formally,

Definition 14.1. A ribbon graph G is a pair (V,E) of collections of topological disks (representing vertices and edges, respectively) such that

- Edge-disks and vertex-disks intersect in disjoint line segments.
- Each such segment lies on the boundary of exactly one edge-disk and exactly one vertex-disk.
- Each edge-disk contains two such segments.

In some cases it may be useful to consider "based" ribbon graphs, i.e. ribbon graphs with one vertex designated as distinguished. We will denote this vertex by a "*" (Fig. 53).

A large part of the combinatorial structure of a ribbon graph is captured by the following functions.

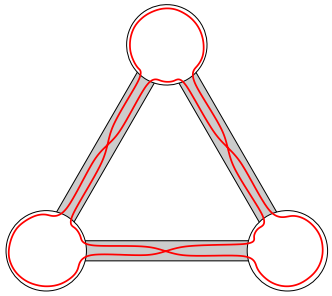


FIGURE 54. A ribbon graph and its medial curve.

Definition 14.2. Let $G = (E,V)$ be a ribbon graph. Define

- $v(G) := |V|$,
- $e(G) := |E|$,
- $k(G) :=$ number of connected components of G,
- $r(G) := v(G) - k(G)$ (this function is called the *rank* of G),
- $n(G) := e(G) - r(g)$ (this function is called the *nullity* of G),
- $bc(G) :=$ number of connected components of the boundary of G.

Notice that the last function is the one that concerns with the ribbon structure whereas the other ones work just as well for regular graphs.

Definition 14.3. A *medial curve* of a ribbon graph is a curve on the ribbon graph which has a crossing on each edge and follows the side of the disk when inside a vertex. This is illustrated in figure 54.

We will need medial curves when studying the relationship between plane curves and ribbon graphs.

14.3. **The Bollobás-Riordan Polynomial.** A *spanning subgraph* of a ribbon graph G is a subgraph consisting of all the vertices of G and a subset of edges of G. Let $\mathcal{F}(G)$ be the set of all spanning subgraphs of G.

Definition 14.4. Let G be a ribbon graph. Then the Bollobás-Riordan polynomial of G is

$$R_G(x,y,z) := \sum_{F \in \mathcal{F}(G)} x^{r(G)-r(F)} y^{n(F)} z^{k(F)-bc(F)+n(F)}.$$

This polynomial was introduced (with a simple change of variables) by B. Bollobás and O. Riordan as an analogue of the Tutte polynomial of regular graphs. A specialization $z = 1$ gives the Tutte polynomial of the underlying graph (i.e. the graph obtained by forgetting the cyclic order of edges at each vertex). In particular, if G is planar, we have for each $F \in \mathcal{F}(G)$, $k(F) - bc(F) + n(F) = 0$, so the Bollobás-Riordan polynomial reduces to the Tutte polynomial. For more information about the Bollobás-Riordan polynomial as well as proofs of the above statements see [BR1, BR2].

14.3.1. *Signed Bollobás-Riordan Polynomial.* A signed ribbon graph $\widehat{G}$ is a ribbon graph $G = (V,E)$ endowed with a sign function $\varepsilon : E \to \{\pm 1\}$. Given a spanning subgraph $\widehat{F} = (V,E',\varepsilon|_{E'})$

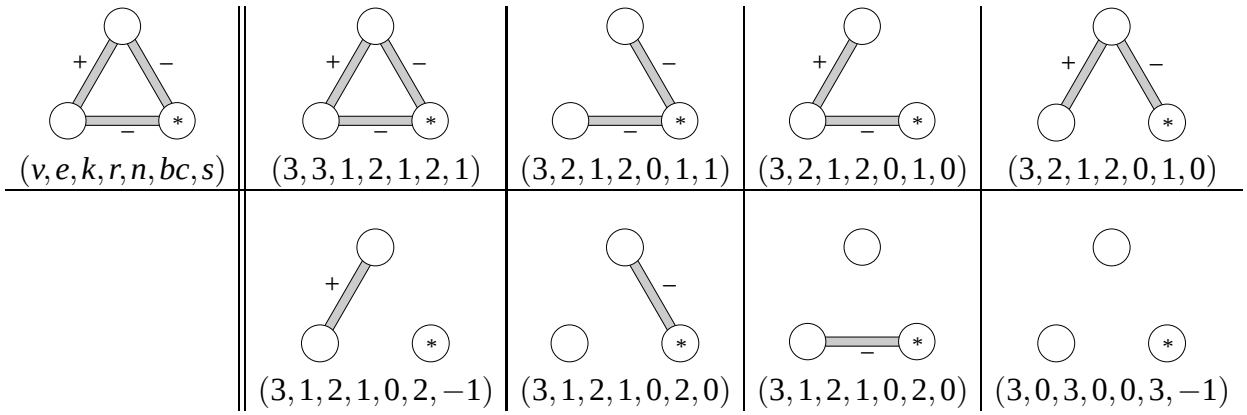


FIGURE 55. An example of the calculation of the Bollobás-Riordan polynomial. The resulting polynomial is $x+xy^{-1}+1+1+y+x+x+xy = xy+3x+y+xy^{-1}+2$.

of $\widehat{G}$ define $e_-(\widehat{F}) := |\{e \in E' : \varepsilon(e) = -1\}|$. Also, define $\widehat{\overline{F}} := (V, E \setminus E', \varepsilon|_{E\setminus E'})$. Notice that $\widehat{\overline{F}}$ is also a spanning subgraph of $\widehat{G}$. Define

$$s(\widehat{F};\widehat{G}) := \frac{e_-(\widehat{F}) - e_-(\widehat{\overline{F}})}{2}.$$

Definition 14.5. [CP] Let $G = (E,V,\varepsilon)$ be a signed ribbon graph. Then the signed Bollobás-Riordan polynomial of G is

$$R_{\widehat{G}}(x,y,z) := \sum_{\widehat{F}\in\mathcal{F}(\widehat{G})} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G})} y^{n(\widehat{F})-s(\widehat{F};\widehat{G})} z^{k(\widehat{F})-bc(\widehat{F})+n(\widehat{F})}.$$

An example of the calculation of the signed Bollobás-Riordan polynomial is given in figure 55.

Remark 14.6. The signed Bollobás-Riordan polynomial is not strictly a polynomial since it may contain half-integer and negative powers of the variables.

Remark 14.7. If all edges of $\widehat{G}$ are positive, then the signed Bollobás-Riordan polynomial is the same as the unsigned one.

14.3.2. *Properties of the Bollobás-Riordan Polynomial.* Given two signed ribbon graphs $\widehat{G_1}$ and $\widehat{G_2}$ we can consider their disjoint union $\widehat{G_1} \cup \widehat{G_2}$ and their one-point joint $\widehat{G_1} \cdot \widehat{G_2}$. The disjoint union is formed by considering the two graphs next to each other while the one-point joint is formed by connecting a vertex of $\widehat{G_1}$ to a vertex of $\widehat{G_2}$ by a bridge and then merging the two vertices (see figure 56).

It turns out that we only need to form one-point involving a based vertex of one graph with any vertex of the other. In this case the resulting vertex is not distinguished. The reason for this rule will become clear once we introduce the ribbon graphs coming from plane curves.

Theorem 14.8. *Let $\widehat{G_1}$ and $\widehat{G_2}$ be two signed ribbon graphs. Then*

$$R_{\widehat{G_1}\cup\widehat{G_2}} = R_{\widehat{G_1}} R_{\widehat{G_2}}, \quad R_{\widehat{G_1}\cdot\widehat{G_2}} = R_{\widehat{G_1}} R_{\widehat{G_2}}. \tag{21}$$

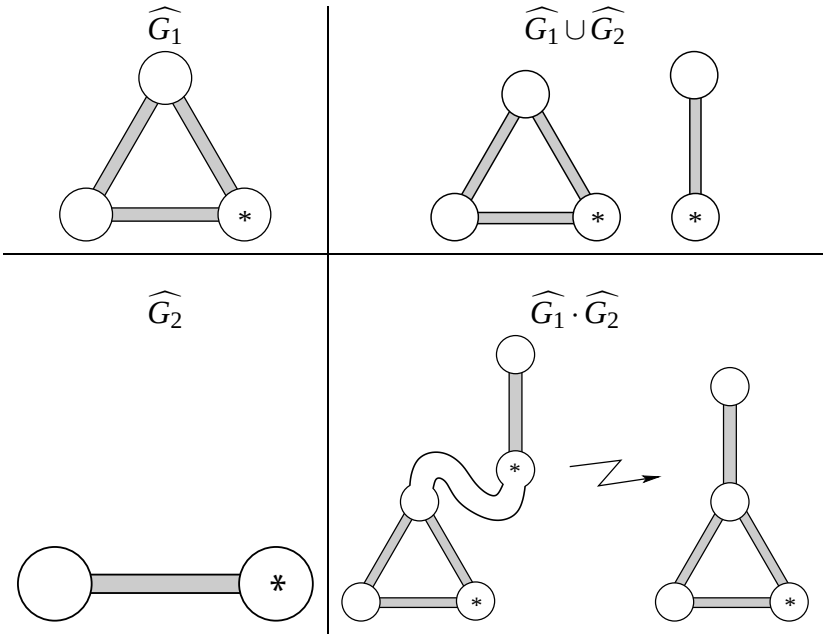


FIGURE 56. Construction of disjoint union and one-point joint of two ribbon graphs.

Theorem 14.9. *Let $\widehat{G}$ be a signed ribbon graph. Then for each positive edge e of $\widehat{G}$ we have*

$$R_{\widehat{G}} = R_{\widehat{G/e}} + R_{\widehat{G-e}}, \tag{22}$$

if e is ordinary, and

$$R_{\widehat{G}} = (x+1)R_{\widehat{G/e}}, \tag{23}$$

if e is a bridge.

Also, for each negative edge e of $\widehat{G}$ we have

$$x^{-1/2}y^{1/2}R_{\widehat{G-e}} + x^{1/2}y^{-1/2}R_{\widehat{G/e}} \tag{24}$$

if e is ordinary, and

$$R_{\widehat{G}} = x^{1/2}(y^{1/2} + y^{-1/2})R_{\widehat{G/e}}, \tag{25}$$

if e is a bridge.

The proofs for these theorems are modifications of the ones in [BR2] to include signing. They are not particularly illuminating and are postponed until section 14.5.

We will also need the following observations.

If $\widehat{F}$ is a spanning subgraph of $\widehat{G}$, then $v(\widehat{F}) = v(\widehat{G})$. So we may rewrite the definition as follows.

$$R_{\widehat{G}}(x,y,z) := \sum_{\widehat{F} \in \mathcal{F}(\widehat{G})} x^{k(\widehat{F})-k(\widehat{G})+s(\widehat{F};\widehat{G})} y^{e(\widehat{F})+k(\widehat{F})-v(\widehat{G})-s(\widehat{F};\widehat{G})} z^{2k(\widehat{F})-bc(\widehat{F})+e(\widehat{F})-v(\widehat{G})}. \tag{26}$$

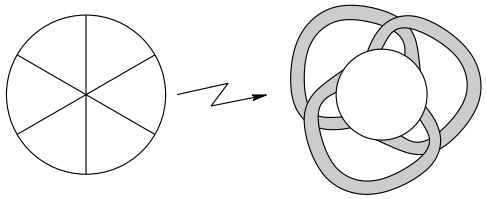


FIGURE 57. A Gauss diagram viewed as a ribbon graph.

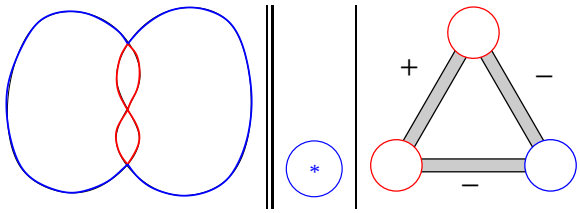


FIGURE 58. Construction of a ribbon graph from a peeling.

Remark 14.10. From this equation we can see a nice specialization of the polynomial:

$$R_{\widehat{G}}\left(x,x,\frac{1}{x}\right) = x^{-k(\widehat{G})} \sum_{\widehat{F}\in\mathcal{F}(\widehat{G})} x^{bc(\widehat{F})}.$$

This specialization does not see the signing of $\widehat{G}$.

14.4. **Ribbon Graphs from Plane and Spherical Curves.** The structure of a curve can be captured by a ribbon graph in several ways. Here we will discuss the ones originating from the Gauss diagram and from the peeling of a plane curve.

Remark 14.11. In both cases, the edges correspond to the double points of the curve. Notice that for any (signed) ribbon graph $\widehat{G} = (E,V,\varepsilon)$ we have $R_{\widehat{G}}(1,1,1) = |\mathcal{F}(\widehat{G})| = 2^{|E|}$. So, if $\widehat{G}$ is a ribbon graph (either Gauss diagram or peeling) of a curve Γ, then the number of double points of Γ is $\log_2(R_{\widehat{G}}(1,1,1))$.

14.4.1. *Gauss Diagram Ribbon Graph.* A signed Gauss diagram of a plane curve may be considered as a signed ribbon graph on its own as shown in figure 57.

Since a signed Gauss diagram is a complete invariant of spherical curves (see [Ca] for the proof), the signed ribbon graph is also a complete invariant of spherical curves.

Unfortunately, we did not have enough time to investigate the properties of this ribbon graph.

14.4.2. *Peeling Ribbon Graph.* Let Γ be a plane curve and consider its peeling. Construct a signed ribbon graph $\widehat{G_\Gamma}$ inductively as follows.

- Take a vertex for each circle of the level 0 cactus and connect them with "+" edges at tangency points. Mark one of these vertices as distinguished.
- Take a vertex for each circle of the level n cactus, connect them with "+" edges at tangency points, and connect them with "−" edges at hanging points to the $(n-1)$-st set of vertices.

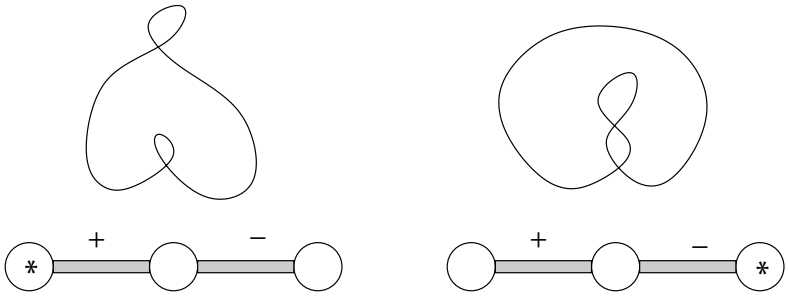


FIGURE 59. Two non-diffeomorphic plane curves whose ribbon graphs differ by the choice of the base point.

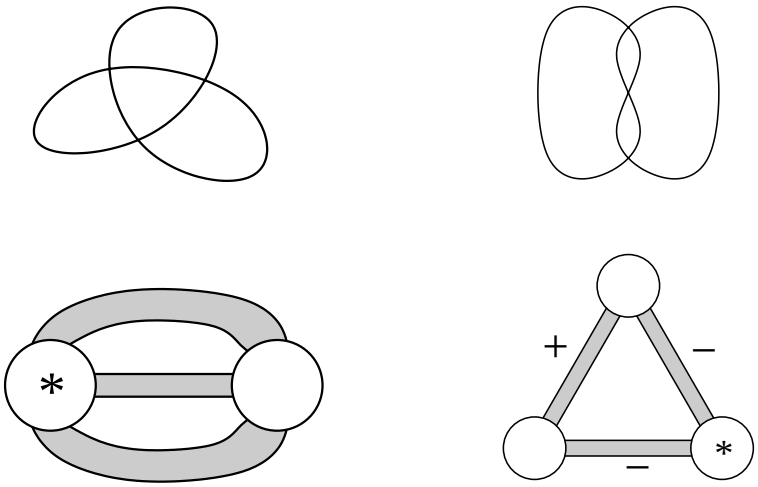


FIGURE 60. Two peeling ribbon graphs corresponding to the two projections of the spherical trefoil. Clearly, the two are not the same.

The procedure is illustrated in figure 58.

Define the peeling Bollobás-Riordan polynomial of Γ by $R_\Gamma(x,y,z) := R_{G_\Gamma}(x,y,z)$.

Remark 14.12. Distinguished points are necessary to be able to uniquely reconstruct a plane curve from its ribbon graph. Otherwise the two curves in figure 59 would have the same ribbon graphs.

Remark 14.13. Since peelings are only defined for plane curves, the peeling ribbon graph can only be constructed from a plane curve. In fact, different projections of a spherical curve may correspond to completely different ribbon graphs (Fig. 60).

Remark 14.14. The edges of the peeling ribbon graph correspond to the double points of the curve and the cyclic order of the edges is the same as the cyclic order of double points. This means that there is a homeomorphism between the original curve and the medial curve of its ribbon graph.

14.4.3. *Properties of the Peeling Ribbon Graph.*

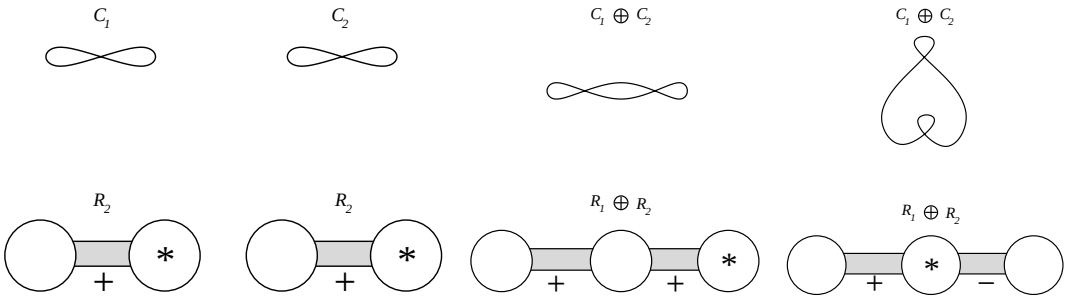


FIGURE 61. The effect of interjected sum on signed ribbon graphs. If the interjected sum is performed from the outside of a cactus, the resulting ribbon graph is just a one point joint (1). If the interjected sum is formed from the inside of a cactus, the resulting ribbon graph is a one-point joint except the positive edges coming out of the base vertex being joined become negative (2).

Proposition 14.15.

(1) *The unsigned Bollobás-Riordan polynomial is multiplicative under interjected sum.*
(2) *The Bollobás-Riordan polynomial is multiplicative under connected sum.*

Proof.

(1) Interjected sum of curves corresponds to a one-point joint of unsigned ribbon graphs. The result follows immediately from (22) and (24).
(2) Connected sum of curves corresponds to a one-point joint of ribbon graphs. The result follows immediately from (22) and (24).

□

Remark 14.16. We only need to consider one-point joints with one of the merging vertices being distinguished since an interjected sum always involves one of the exterior components.

Remark 14.17. The absence of signs in the first part of the proposition is necessary since interjected sums may affect the signings as seen in figure 61.

Proposition 14.18. *Suppose* Γ *is a tree-like curve whose peeling contains* n_1 *hanging points and* n_2 *joining points. Then*

$$R_\Gamma(x,y,z) = (x^{1/2}(y^{1/2}+y^{-1/2}))^{n_1}(x+1)^{n_2}.$$

Proof. Notice that G_Γ is a tree with n_1 negative edges and n_2 positive edges. The result follows immediately from (23) and (25). □

14.4.4. *Explicit Formula for the Bollobás-Riordan Polynomial of a Curve.* The Bollobás-Riordan polynomial of a curve was introduced using the ribbon graph construction. In some cases it is useful to rewrite the polynomial in terms of the properties of the curve and avoid the ribbon graph step. We will deal with unsigned Bollobás-Riordan polynomial. Suppose Γ is a plane curve. Take a double point x of Γ. There are two ways of "splitting" x as shown in figure 62.

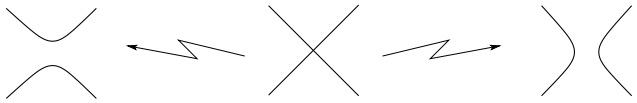


FIGURE 62. Two ways of splitting a crossing.

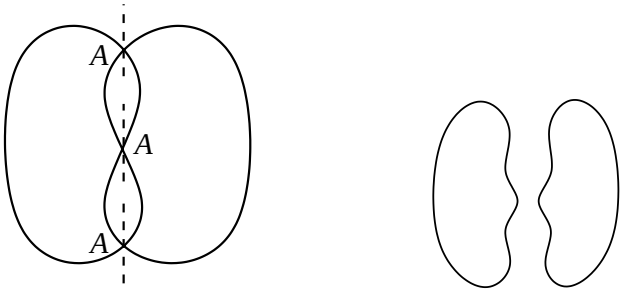


FIGURE 63. A smoothing.

Definition 14.19. An *A-splitting* of a double point x of a curve Γ is a splitting where either

i. the two regions being joined are both inside a cactus of a given level, or
ii. the two regions being joined are the outside of a cactus with the inside of a cactus of the next level.

A splitting is called a *B-splitting* if it does not satisfy the above condition.

Of course, an A-splitting corresponds to a splitting of the medial along the edge of the peeling ribbon graph, while a B-splitting corresponds to a splitting of the medial across the edge. This means that the homeomorphism between the two curves is preserved under splittings.

Definition 14.20. A *smoothing* of a plane curve Γ is a choice of either A- or B- splitting at each double point of Γ. Define $\mathcal{S}(\Gamma)$ to be the set of all smoothings of Γ. Define S_A to be the smoothing with all A-splittings and S_B to be the smoothing with all B-splittings.

Take $S \in \mathcal{S}(\Gamma)$. Performing all the splittings of S on Γ results in a disjoint union of circles (Fig. 63). Break these circles into groups according to the following rule. Two circles belong to the same group if they can be joined by undoing several A-splittings.

Definition 14.21. Let Γ be a plane curve and let $S \in \mathcal{S}(\Gamma)$. Define

$$\begin{aligned} Ce(S) &:= \text{number of } A\text{-splittings in } S, \\ Ck(S) &:= \text{number of groups resulting after breaking } \Gamma \text{ according to } S, \\ Cc(S) &:= \text{number of circles resulting after breaking } \Gamma \text{ according to } S. \end{aligned}$$

Proposition 14.22. *Suppose Γ is a plane curve with unsigned ribbon graph G_Γ. Suppose F is a spanning subgraph of G_Γ. Let S be a smoothing of Γ with A-splittings at double points corresponding to the edges present in F and B-splittings at other double points. Then*

$$e(F) = Ce(S), \qquad bc(F) = Cc(S), \qquad k(F) = Ck(S).$$

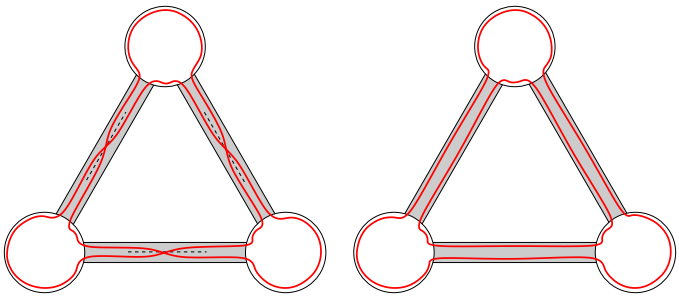


FIGURE 64. The splitting S' contains a circle for each boundary component of F. Here F consists of the bottom and right edges of the triangle.

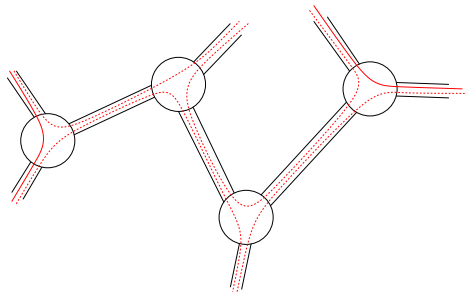


FIGURE 65. A path of vertices between two boundary components of a connected component of a ribbon graph.

Proof. The first statement is obvious since the smoothing S was constructed so that the A-splittings correspond to edges present in F.

Now prove the second statement. Take a smoothing S' of the medial curve of G_Γ which corresponds to S. That is crossings on edges present in F are split along the edge, while crossings on edges not in F are split across the edge. As mentioned before there is a homeomorphism between the two smoothings, so they must have the same number of circles. However, there is a circle of S' along each boundary component of S (Fig. 64), so the number of circles in S' is $bc(F)$. Hence $bc(F) = Cc(S)$.

Finally prove the last statement. We would like to show that two circles of S' correspond to boundary components of the same connected component of F if and only if they can be joined by undoing several splittings along the edges. One of the directions is trivial, namely if there is such a sequence of splittings then the two boundary components correspond to the same connected component.

Suppose two circles of S' correspond to the same connected component. Notice that all boundary circles which pass through a given vertex belong to the same group. But the two circles are inside a connected component of a graph, so there is a path of vertices between them (Fig. 65). Hence they have to be in the same group.

Thus we have shown that the number of groups of circles of the smoothing S' (which is the same as the number of groups of circles of S) is just $k(F)$. Hence $Ck(S) = k(F)$.

This completes the proof. □

Notice that spanning subgraphs of G_Γ are in one-to-one correspondence with the smoothings of Γ. Also, the number of vertices of G_Γ is the same as the number of connected components of the spanning subgraph with no edges, and hence, according to the proposition the same as $Ck(S_B)$. Finally, the graph G_Γ is connected, so $k(G_\Gamma) = 1$. Using these, we may rewrite the unsigned version of equation (26) as

$$R_\Gamma(x,y,z) := \sum_{S\in\mathcal{S}(\Gamma)} x^{Ck(S)-1} y^{Ce(S)+Ck(S)-Ck(S_B)} z^{2Ck(S)-Cc(S)+Ce(S)-Ck(S_B)}. \tag{27}$$

The analogue of remark 14.10 is

Remark 14.23.

$$R_\Gamma(x,x,1/x) = x^{-1} \sum_{S\in\mathcal{S}(\Gamma)} x^{Cc(S)}.$$

The definition of $Cc(S)$ does not use any properties of the peeling and hence is defined on the sphere. Thus we have

Theorem 14.24. *Let Γ be a plane curve. Then $R_\Gamma(x,x,1/x)$ depends only on the spherical type of Γ, i.e. it is a polynomial invariant of spherical curve.*

We thus call $R_\Gamma(x,x,1/x)$ the spherical Bollobás-Riordan polynomial of Γ.

Conjecture 14.25. *If Γ, a plane curve, is not a circle, then $R_\Gamma(-1,-1,-1) = 0$.*

14.5. **Proofs of Theorems 14.8 and 14.9.**

Proof of Theorem 14.8. Notice that for each $\widehat{F_1} \in \mathcal{F}(\widehat{G_1})$, $\widehat{F_2} \in \mathcal{F}(\widehat{G_2})$, we have

$$\begin{array}{rcl|rcl}
v(\widehat{F_1}\cup\widehat{F_2}) &=& v(\widehat{F_1})+v(\widehat{F_2}) & v(\widehat{F_1}\cdot\widehat{F_2}) &=& v(\widehat{F_1})+v(\widehat{F_2})-1\\
e(\widehat{F_1}\cup\widehat{F_2}) &=& e(\widehat{F_1})+e(\widehat{F_2}) & e(\widehat{F_1}\cdot\widehat{F_2}) &=& e(\widehat{F_1})+e(\widehat{F_2})\\
k(\widehat{F_1}\cup\widehat{F_2}) &=& k(\widehat{F_1})+k(\widehat{F_2}) & k(\widehat{F_1}\cdot\widehat{F_2}) &=& k(\widehat{F_1})+k(\widehat{F_2})-1\\
r(\widehat{F_1}\cup\widehat{F_2}) &=& r(\widehat{F_1})+r(\widehat{F_2}) & r(\widehat{F_1}\cdot\widehat{F_2}) &=& r(\widehat{F_1})+r(\widehat{F_2})\\
n(\widehat{F_1}\cup\widehat{F_2}) &=& n(\widehat{F_1})+n(\widehat{F_2}) & n(\widehat{F_1}\cdot\widehat{F_2}) &=& n(\widehat{F_1})+n(\widehat{F_2})\\
bc(\widehat{F_1}\cup\widehat{F_2}) &=& bc(\widehat{F_1})+bc(\widehat{F_2}) & bc(\widehat{F_1}\cdot\widehat{F_2}) &=& bc(\widehat{F_1})+bc(\widehat{F_2})-1\\
s(\widehat{F_1}\cup\widehat{F_2};\widehat{G_1}\cup\widehat{G_2}) &=& s(\widehat{F_1};\widehat{G_1})+s(\widehat{F_2};\widehat{G_2}) & s(\widehat{F_1}\cdot\widehat{F_2};\widehat{G_1}\cdot\widehat{G_2}) &=& s(\widehat{F_1};\widehat{G_1})+s(\widehat{F_2};\widehat{G_1})
\end{array}$$

Hence

$$\begin{aligned} x^{r(\widehat{G_1}\cup\widehat{G_2})-r(\widehat{F_1}\cup\widehat{F_2})+s(\widehat{F_1}\cup\widehat{F_2};\widehat{G_1}\cup\widehat{G_2})} y^{n(\widehat{F_1}\cup\widehat{F_2})-s(\widehat{F_1}\cup\widehat{F_2};\widehat{G_1}\cup\widehat{G_2})} z^{k(\widehat{F_1}\cup\widehat{F_2})-bc(\widehat{F_1}\cup\widehat{F_2})+n(\widehat{F_1}\cup\widehat{F_2})} = \\ x^{r(\widehat{G_1})-r(\widehat{F_1})+s(\widehat{F_1};\widehat{G_1})} y^{n(\widehat{F_1})-s(\widehat{F_1};\widehat{G_1})} z^{k(\widehat{F_1})-bc(\widehat{F_1})+n(\widehat{F_1})} \times \\ \times x^{r(\widehat{G_2})-r(\widehat{F_2})+s(\widehat{F_2};\widehat{G_2})} y^{n(\widehat{F_2})-s(\widehat{F_2};\widehat{G_2})} z^{k(\widehat{F_2})-bc(\widehat{F_2})+n(\widehat{F_2})}, \end{aligned} \tag{28}$$

and

$$\begin{aligned} x^{r(\widehat{G_1}\cdot\widehat{G_2})-r(\widehat{F_1}\cdot\widehat{F_2})+s(\widehat{F_1}\cdot\widehat{F_2};\widehat{G_1}\cdot\widehat{G_2})} y^{n(\widehat{F_1}\cdot\widehat{F_2})-s(\widehat{F_1}\cdot\widehat{F_2};\widehat{G_1}\cdot\widehat{G_2})} z^{k(\widehat{F_1}\cdot\widehat{F_2})-bc(\widehat{F_1}\cdot\widehat{F_2})+n(\widehat{F_1}\cdot\widehat{F_2})} = \\ x^{r(\widehat{G_1})-r(\widehat{F_1})+s(\widehat{F_1};\widehat{G_1})} y^{n(\widehat{F_1})-s(\widehat{F_1};\widehat{G_1})} z^{k(\widehat{F_1})-bc(\widehat{F_1})+n(\widehat{F_1})} \times \\ \times x^{r(\widehat{G_2})-r(\widehat{F_2})+s(\widehat{F_2};\widehat{G_1})} y^{n(\widehat{F_2})-s(\widehat{F_2};\widehat{G_1})} z^{k(\widehat{F_2})-bc(\widehat{F_2})+n(\widehat{F_2})}. \end{aligned} \tag{29}$$

However pairs of spanning subgraphs of $\widehat{G_1}$ and $\widehat{G_2}$ are in one-to-one correspondence with spanning subgraphs of $\widehat{G_1 \cup G_2}$ or $\widehat{G_1 \cdot G_2}$. Together with (28) and (29) this gives

$$R_{\widehat{G_1 \cup G_2}} = R_{\widehat{G_1 \cdot G_2}} = R_{\widehat{G_1}} R_{\widehat{G_2}}.$$

□

Proof of Theorem 14.9. Suppose e is positive.

Notice that the spanning subgraphs of $\widehat{G}$ which do not contain e are in one-to-one correspondence with the subgraphs of $\widehat{G-e}$, while the spanning subgraphs of $\widehat{G}$ which contain e are in one-to-one correspondence with the subgraphs of $\widehat{G/e}$.

If $\widehat{F}$ is a spanning subgraph of $\widehat{G}$ which does not contain e then $s(\widehat{F};\widehat{G}) = s(\widehat{F};\widehat{G-e})$ since e is positive. All other functions obviously do not depend on whether $\widehat{F}$ is a subgraph of $\widehat{G}$ or $\widehat{G-e}$.

If $\widehat{F}$ is a spanning subgraph of $\widehat{G}$ which contains e then we have (since e is not a loop)

$$\begin{array}{lcllcl} v(\widehat{F/e}) & = & v(\widehat{F})-1, & e(\widehat{F/e}) & = & e(\widehat{F})-1, \\ k(\widehat{F/e}) & = & k(\widehat{F}), & r(\widehat{F/e}) & = & r(\widehat{F})-1, \\ n(\widehat{F/e}) & = & n(\widehat{F}), & bc(\widehat{F/e}) & = & bc(\widehat{F}), \\ s(\widehat{F/e};\widehat{G/e}) & = & s(\widehat{F};\widehat{G}). & & & \end{array}$$

Again, s does not change for the same reason.

Suppose e is ordinary. So, $k(\widehat{G}) = k(\widehat{G-e}) = k(\widehat{G/e})$, and hence $r(\widehat{G}) = r(\widehat{G-e}) = r(\widehat{G/e})+1$. So, we have

$$\begin{aligned} R_{\widehat{G}} &= \sum_{\substack{\widehat{F}\in\mathcal{F}(\widehat{G}) \\ e\notin E(\widehat{F})}} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G})} y^{n(\widehat{F})-s(\widehat{F};\widehat{G})} z^{k(F)-bc(F)+n(F)} + \\ &\quad + \sum_{\substack{\widehat{F}\in\mathcal{F}(\widehat{G}) \\ e\in E(\widehat{F})}} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G})} y^{n(\widehat{F})-s(\widehat{F};\widehat{G})} z^{k(\widehat{F})-bc(\widehat{F})+n(\widehat{F})} \\ &= \sum_{\widehat{F}\in\mathcal{F}(\widehat{G-e})} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G})} y^{n(\widehat{F})-s(\widehat{F};\widehat{G})} z^{k(\widehat{F})-bc(\widehat{F})+n(\widehat{F})} + \\ &\quad + \sum_{\substack{\widehat{F}\in\mathcal{F}(\widehat{G}) \\ e\in E(\widehat{F})}} x^{r(\widehat{G/e})+1-r(\widehat{F/e})-1+s(\widehat{F/e};\widehat{G/e})} y^{n(\widehat{F/e})-s(\widehat{F/e};\widehat{G/e})} z^{k(\widehat{F/e})-bc(\widehat{F/e})+n(\widehat{F/e})} \\ &= R_{\widehat{G-e}} + R_{\widehat{G/e}}. \end{aligned}$$

Now suppose e is a bridge. Then $k(\widehat{G/e}) = k(\widehat{G}) = k(\widehat{G-e})-1$. So, we will have

$$R_{\widehat{G}} = xR_{\widehat{G-e}} + R_{\widehat{G/e}}.$$

However, since e was a bridge, $\widehat{G-e}$ is a disjoint union of $\widehat{G_1}$ and $\widehat{G_2}$ with $\widehat{G/e} = \widehat{G_1} \cdot \widehat{G_2}$. So, by (21), $R_{\widehat{G-e}} = R_{\widehat{G/e}}$. Hence,

$$R_{\widehat{G}} = (x+1)R_{\widehat{G/e}}.$$

Now suppose e is negative. In this case the function s will change depending on whether we consider subgraphs of $\widehat{G}$, $\widehat{G/e}$, or $\widehat{G-e}$. In particular, if $\widehat{F}$ is a spanning subgraph of $\widehat{G}$ which does not contain e then $s(\widehat{F};\widehat{G}) = s(\widehat{F};\widehat{G-e}) - 1/2$, and if $\widehat{F}$ is a spanning subgraph of $\widehat{G}$ which contains e then $s(\widehat{F};\widehat{G}) = s(\widehat{F};\widehat{G/e}) + 1/2$.

Thus, if e is ordinary, we have

$$\begin{aligned}
R_{\widehat{G}} &= \sum_{\substack{\widehat{F}\in\mathcal{F}(\widehat{G})\\ e\notin E(\widehat{F})}} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G})} y^{n(\widehat{F})-s(\widehat{F};\widehat{G})} z^{k(\widehat{F})-bc(\widehat{F})+n(\widehat{F})} + \\
&\qquad + \sum_{\substack{\widehat{F}\in\mathcal{F}(\widehat{G})\\ e\in E(\widehat{F})}} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G})} y^{n(\widehat{F})-s(\widehat{F};\widehat{G})} z^{k(\widehat{F})-bc(\widehat{F})+n(\widehat{F})} \\
&= \sum_{\widehat{F}\in\mathcal{F}(\widehat{G-e})} x^{r(\widehat{G})-r(\widehat{F})+s(\widehat{F};\widehat{G-e})-1/2} y^{n(\widehat{F})-s(\widehat{F};\widehat{G-e})+1/2} z^{k(\widehat{F})-bc(\widehat{F})+n(\widehat{F})} + \\
&\qquad + \sum_{\substack{\widehat{F}\in\mathcal{F}(\widehat{G})\\ e\in E(\widehat{F})}} x^{r(\widehat{G/e})+1-r(\widehat{F/e})-1+s(\widehat{F/e};\widehat{G/e})+1/2} y^{n(\widehat{F/e})-s(\widehat{F/e};\widehat{G/e})-1/2} z^{k(\widehat{F/e})-bc(\widehat{F/e})+n(\widehat{F/e})} \\
&= x^{-1/2}y^{1/2}R_{\widehat{G-e}} + x^{1/2}y^{-1/2}R_{\widehat{G/e}},
\end{aligned}$$

while if e is a bridge we have

$$R_{\widehat{G}} = x\cdot x^{-1/2}y^{1/2}R_{\widehat{G-e}} + x^{1/2}y^{-1/2}R_{\widehat{G/e}} = x^{1/2}(y^{1/2}+y^{-1/2})R_{\widehat{G/e}}.$$

□

14.6. **Unsolved Problems.**

- Prove or disprove conjecture 14.25.
- Interpret the coefficients of the spherical Bollobás-Riordan polynomial it terms of the properties of the Gauss diagram. This should be possible since the polynomial is fully determined by the Gauss diagram.
- What is the meaning of the coefficients of the general Bollobás-Riordan polynomial?
- Currently the Bollobás-Riordan polynomial does not see the distinguished point so it has no hope of being a complete invariant of plane curves. Can the definition be altered for the purpose?
- The distinguished vertex construct is not very nice, in the sense that the choice of distinguished vertex is not unique for a given curve. Is there a better one?
- Find the relationships, if any, of the Bollobás-Riordan polynomial with the more common invariants. It is impossible to obtain Arnold's invariants from it, however it might be possible to express, say, the *defect*, defined as $J^+ + 2St$. It is known that the defect, a spherical

invariant, is additive under interjected sums. On the other hand, the spherical Bollobás-Riordan polynomial is multiplicative under interjected sum (since this substitution does not see the signs). This suggests that by taking a logarithm of some specialization we might be able to obtain the defect.

Part 8. **Conclusion**

The results of the research done in this work are fairly self evident. With prime Gauss diagrams, we have significantly reduced the problem of curve classification to the planarity problem of chord diagrams. Conversely, via the construction method, we have successfully turned the planarity problem onto a smaller and more substantial group of chord diagrams for more attainable results. A complete invariant has been found for plane curves, in the form of plane structures, that behaves coherently with signed Gauss diagrams, demonstrates relationships between regions on spherical curves, and can be systematically employed to find more conventional invariants. Our study of plane curve symmetries has added an interesting algebraic element to our identification of equivalent and similar curves. And the employment of ribbon graphs and the Bollobás-Riordan polynomial has opened up a new avenue for research in the subject of plane and spherical curve geometry by tying it into combinatorics and has already demonstrated some promising results.

Nevertheless, much remains to be investigated, and if you've read this far you are likely in a future REU that could use a little inspiration in the area of open problems. There is research to be done, as we've stated, in the way of ribbon graphs and polynomials as an invariant of plane and spherical curves, we have yet to construct a more methodical approach to weeding out the combinatorial symmetries arising from the construction of composite Gauss diagrams to prevent the construction of equivalent invariants, and no work has been done to understand how our new plane structures behave under Arnold moves or other homotopic deformations of plane curves. Finally, the discovery of a constructive solution to the planarity problem, one that would allow us to quickly build all prime Gauss diagrams without having to check the planarity of non-planar chord diagrams, would have enormous implications for the work done in this paper while being a praiseworthy accomplishment in its own right.

Part 9. Appendices

Appendix A. Programs.

A.1. **Generationg Prime Gauss Diagrams.** Connected curves are building blocks from which all other curves are constructed. Since the Gauss diagram (or word) is a complete invariant of spherical curves (Corollary 2.3), enumerating prime Gauss words corresponding to spherical curves is equivalent to enumerating the curves themselves. Cairns and Elton ([CE]) invented an algorithm to check the planarity of a Gauss word. The GAP code (`http://www-gap.mcs.st-and.ac.uk/`) given below implements this algorithm to generate all prime Gauss diagrams with a given number of chords.

Based on this data we give a conjecture

Conjecture A.1. *Let s_n be the number of prime gauss diagrams. Then*

$$\lim_{n\to\infty} \frac{s_n}{3^n} = \gamma,$$

where γ is a constant.

For other interesting methods of enumerating curves see [DGZ] and [JZJ]. For a discussion of the assymptotics of the number of plane curves see [SZJ].

```
## Gauss word - [letter1, letter2, ...]
## letter - [k, sign(k)]

bet := function(G, a)
  local pp, pm, result, l, dbl, p;
  pp:=Position(G, [a,1]);
  pm:=Position(G, [a,-1]);
  if (pp < pm) then result := G{[pp..pm]};
  else result := G{[pm..pp]};
  fi;
  l := Length(result);
  p:= 1;
  while p in [1..l] do
    dbl := [result[p][1],-1*result[p][2]];
    if (dbl in result) then
      Remove(result, p);
      Remove(result, Position(result, dbl));
      l:=l-2;
      p:=p-1;
    fi;
    p:=p+1;
  od;
  return result;
end;

f := function(G, A)
```

```
  local result, l;
  result:=[];
  for l in A do
    result := Union(result, bet(G,l[1]));
  od;
  return Union(A, result);
end;

check_connected := function(G, k)
  local f_prev, f_new, Betweenness, i;
  Betweenness := [];
  f_prev := f(G, [G[1]]);
  f_new := f(G, f_prev);
  while (f_new <> f_prev) do
    f_prev := f_new;
    f_new := f(G, f_prev);
  od;
  for i in f_new do Add(Betweenness, i[1]); od;
  return Set(Betweenness) = [1..k];
end;

check_planar := function(G, k)
  local S, S_bar, S_inv, i, j, pp, pm, check;
  check := true;
  S := []; S_bar := []; S_inv := [];
  for i in [1..k] do
    pp := Position(G, [i, 1]);
    pm := Position(G, [i, -1]);
    if (pp < pm) then
      S_bar[i] := G{[pp..pm]};
    else
      S_bar[i] := Union(G{[1..pm]}, G{[pp..2*k]});
    fi;
    S[i] := ShallowCopy(S_bar[i]);
    Remove(S[i], Position(S[i], [i, 1]));
    Remove(S[i], Position(S[i], [i, -1]));
    S_inv[i] := List(S[i], x -> [x[1],-1*x[2]]);
  od;
  for i in [1..k] do
    for j in [1..k] do
      if (Sum(Intersection(S_bar[i], S_inv[j]), x -> x[2]) <> 0) then
        check := false;
      fi;
    od;
```

```
  od;
  return check;
end;

build_check_word := function(G,k,n,grand_result)
  local i, j, G_next, pp, pm, betw, signs_cancel, dm_lst, t, conn;
  if (n = k+1) then
    signs_cancel := true;
    for i in [1..k] do
      pp:=Position(G,[i,1]);
      pm:=Position(G,[i,-1]);
      if (pp < pm) then betw := G{[pp..pm]};
      else betw := G{[pm..pp]};
      fi;
      if (not (Sum(betw,x -> x[2]) = 0)) then signs_cancel := false; fi;
    od;
    if (signs_cancel) then
      if(check_connected(G,k)) then
        if (check_planar(G,k)) then
          dm_lst := List(G,x -> x[1]);
          Add(grand_result, ShallowCopy(dm_lst));
        fi;
      fi;
    fi;
  else
    G_next:=ShallowCopy(G);
    for i in [2..(2*k-1)] do
      if(not IsBound(G_next[i])) then
        j:= i+1;
        while j <= 2*k do
          if(not IsBound(G_next[j])) then
            conn := false;
            G_next[i] := [n,1];
            G_next[j] := [n,-1];
            build_check_word(G_next, k, n+1, grand_result);
            G_next[i] := [n,-1];
            G_next[j] := [n,1];
            build_check_word(G_next, k, n+1, grand_result);
            Unbind(G_next[j]);
          fi;
          j := j+2;
        od;
        return;
      fi;
```

```
    od;
  fi;
end;

rotate:=function(w,r)
  local i, result, permut, l, cur_ind;
  l := Length(w);
  result := ShallowCopy(w);
  cur_ind := 1;
  permut:=[];
  for i in [1+r..l] do
    if(IsBound(permut[result[i]])) then
      result[i] := permut[result[i]];
    else
      permut[result[i]] := cur_ind;
      result[i] := cur_ind;
      cur_ind := cur_ind + 1;
    fi;
  od;
  for i in [1..r] do
    if(IsBound(permut[result[i]])) then
      result[i] := permut[result[i]];
    else
      permut[result[i]] := cur_ind;
      result[i] := cur_ind;
      cur_ind := cur_ind + 1;
    fi;
  od;
  return Concatenation(result{[1+r..l]}, result{[1..r]});
end;

clean_up:=function(result,k,fout)
  local i,j,trash_list, clean_list, piece1, piece2,
        piece1r, piece2r, m1, m2;
  clean_list := [];
  i := 1;
  while (Length(result) > 0) do
    clean_list[i] := ShallowCopy(result[1]);
    AppendTo(fout, result[1], "\n");
    trash_list:=[];
    for j in [1..2*k] do
      piece1 := rotate(result[1],j);
      m1:= Position(piece1{[2..2*k]}, 1);
      piece1r:=
```

```
        rotate(Concatenation(piece1{[m1+1..2*k]}, piece1{[1..m1]}), 2*k);
      piece2 := rotate(Reversed(result[1]),j);
      m2:= Position(piece2{[2..2*k]}, 1);
      piece2r:=
        rotate(Concatenation(piece2{[m2+1..2*k]}, piece2{[1..m2]}), 2*k);
      Add(trash_list, ShallowCopy(piece1));
      Add(trash_list, ShallowCopy(piece1r));
      Add(trash_list, ShallowCopy(piece2));
      Add(trash_list, ShallowCopy(piece2r));
    od;
    result:= Difference(result,trash_list);
    i:= i+1;
  od;
  return clean_list;
end;

list_diagrams := function(k,fout)
  local G, i, grand_result;
  AppendTo(fout, "\n!!!!!!!", k, " chords!!!!!!!!\n");
  grand_result := [];
  G := [];
  G[1] := [1,1];
  for i in [1.. k] do
    G[2*i] := [1,-1];
    build_check_word(G,k,2,grand_result);
    Unbind(G[2*i]);
  od;
  return clean_up(grand_result,k,fout);
end;
```

A.2. **Drawing Gauss Diagrams.** The following Maple code generates Gauss diagram pictures in `eps` format from a file containing one Gauss word per line. The Gauss word is entered as a Maple list, so that the Gauss diagram in figure 66 would be entered as `[1, 2, 3, 1, 2, 3]` or equivalent.

```
> with(linalg): with(plots):
 file:="7x.log";
 line := parse(readline(file)):
 k := nops(line)/2:
 cur_ln_num := 1:
 pts := [seq([cos(i/k*Pi), sin(i/k*Pi), 0], i=1..2*k)]:
 while line <> 0 do
  plotsetup(ps,plotoutput =
    (‘gauss‘ || k || ‘_‘ || cur_ln_num || ‘.eps‘),
```

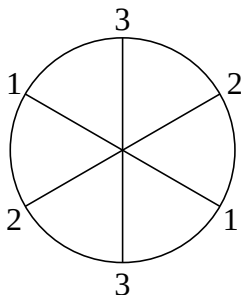


FIGURE 66. A Gauss diagram.

```
    plotoptions = 'portrait, noborder, width=3in, height=3in'):
  endpts := [seq([], i=1..k)]:
  for i from 1 to 2*k do
    endpts[line[i]] := [op(endpts[line[i]]), pts[i]]:
  end do;
  diagram := spacecurve({[cos(2*Pi*t), sin(2*Pi*t), 0],
    seq(convert(
                evalm(endpts[i][1] + (endpts[i][2] - endpts[i][1])*t),
                list), i=1..k)},
              t = 0 .. 1,
              color = black,
              axes = none,
              orientation = [-90, 0],
              thickness = 3):
  display(diagram);
  cur_ln_num := cur_ln_num + 1:
  line := readline(file):
  if (line <> 0) then line := parse(line): end if;
end do;
plotsetup(default);
```

APPENDIX B. TABLES OF PRIME GAUSS DIAGRAMS.

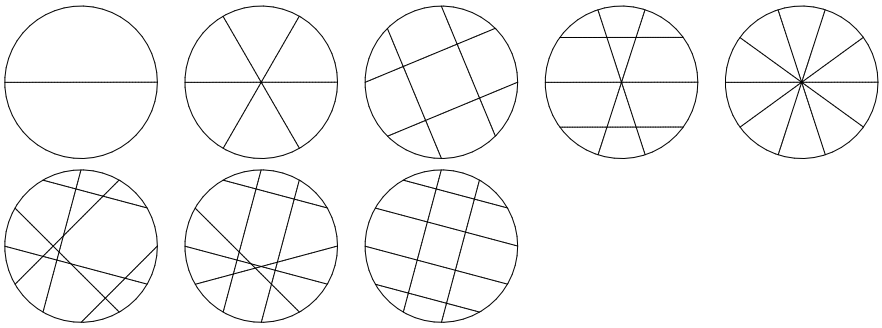
The following computations were conducted on a Dual Pentium 4 (3GHz) running GAP 2.4r4 with code loaded as a dynamic module. Both the curve and the surface were considered unoriented. The following table summarizes the results.

n	time	number of diagrams
1	0	1
2	0	0
3	0.002s	1
4	0.011s	1
5	0.136s	2
6	1.738s	3
7	46.310s	10
8		27
9		101
10	60h 33m 45.377s	364

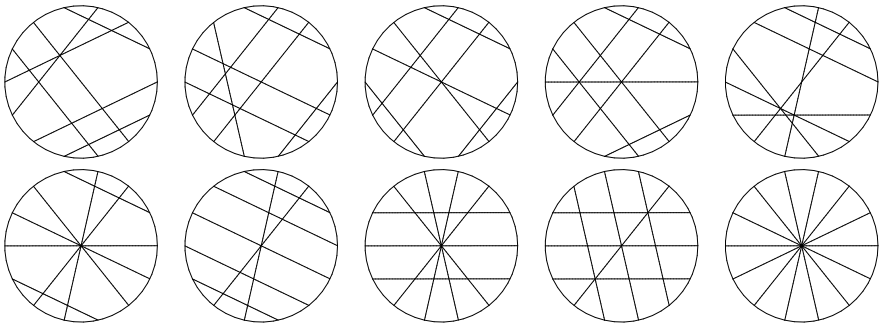
This is the same as the number of projections of reduced, prime, alternating knots with n crossings.

Below we give all these Gauss diagrams.

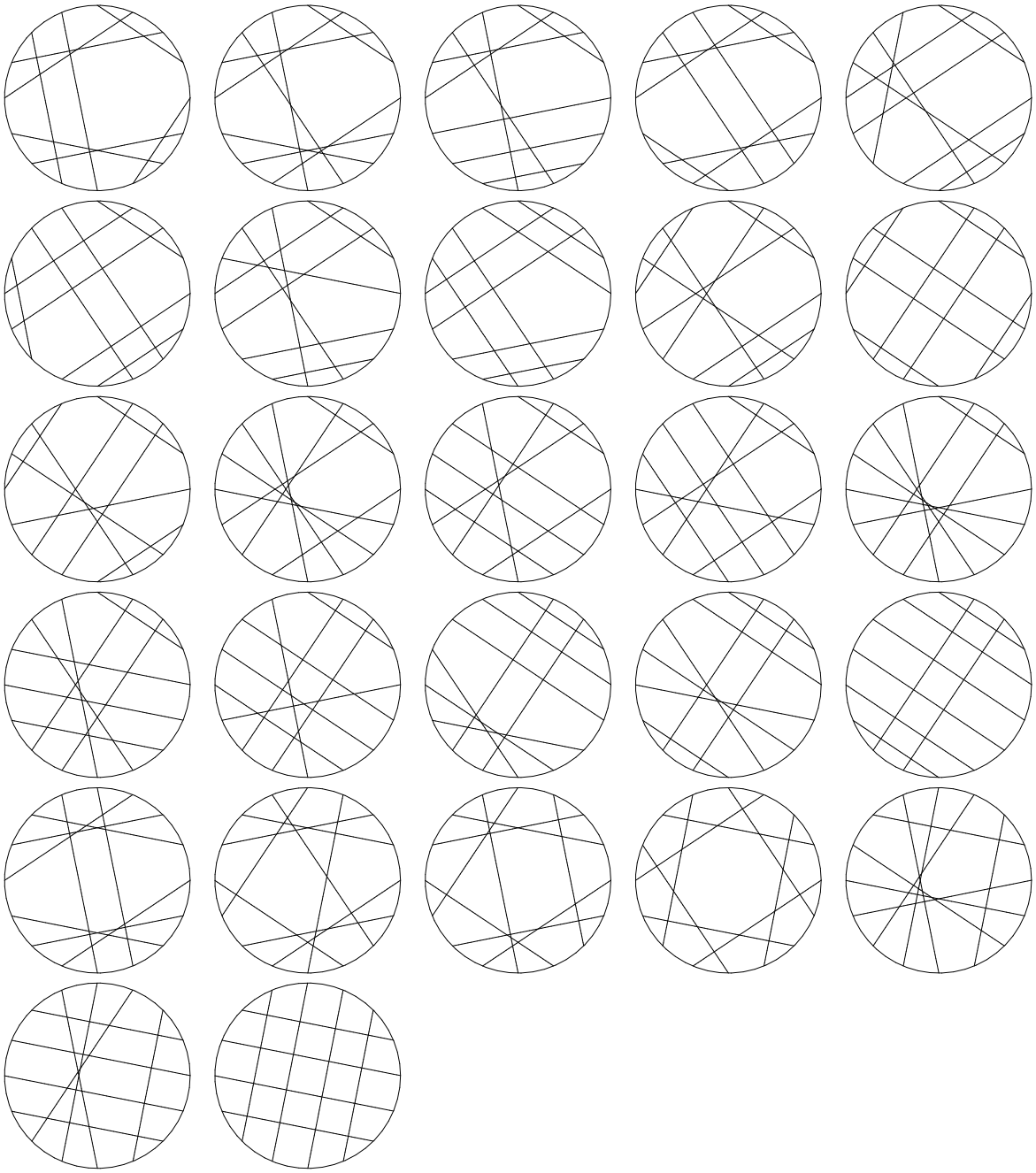
Prime Gauss diagrams with $1-6$ crossings.



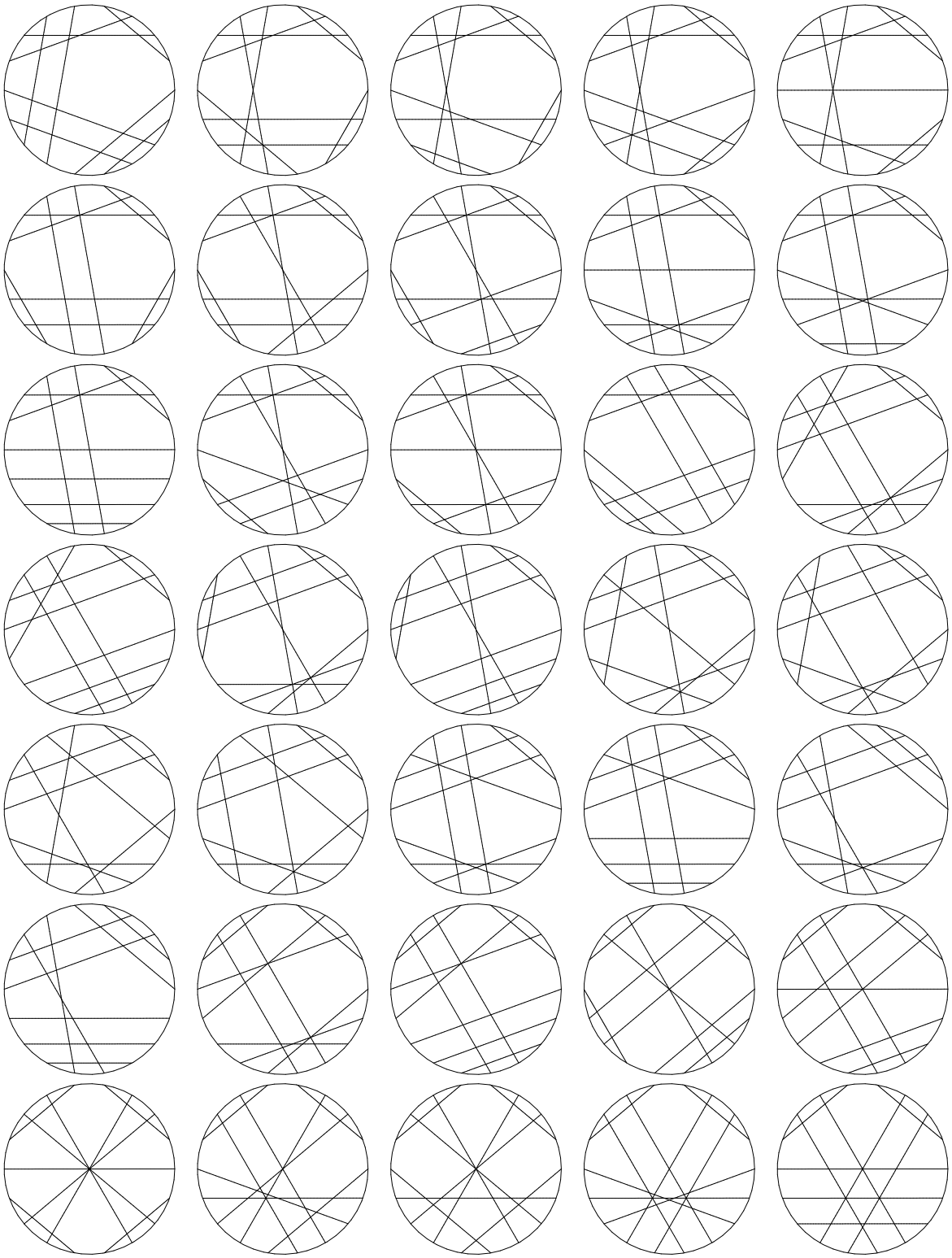
Prime Gauss diagrams with 7 crossings.

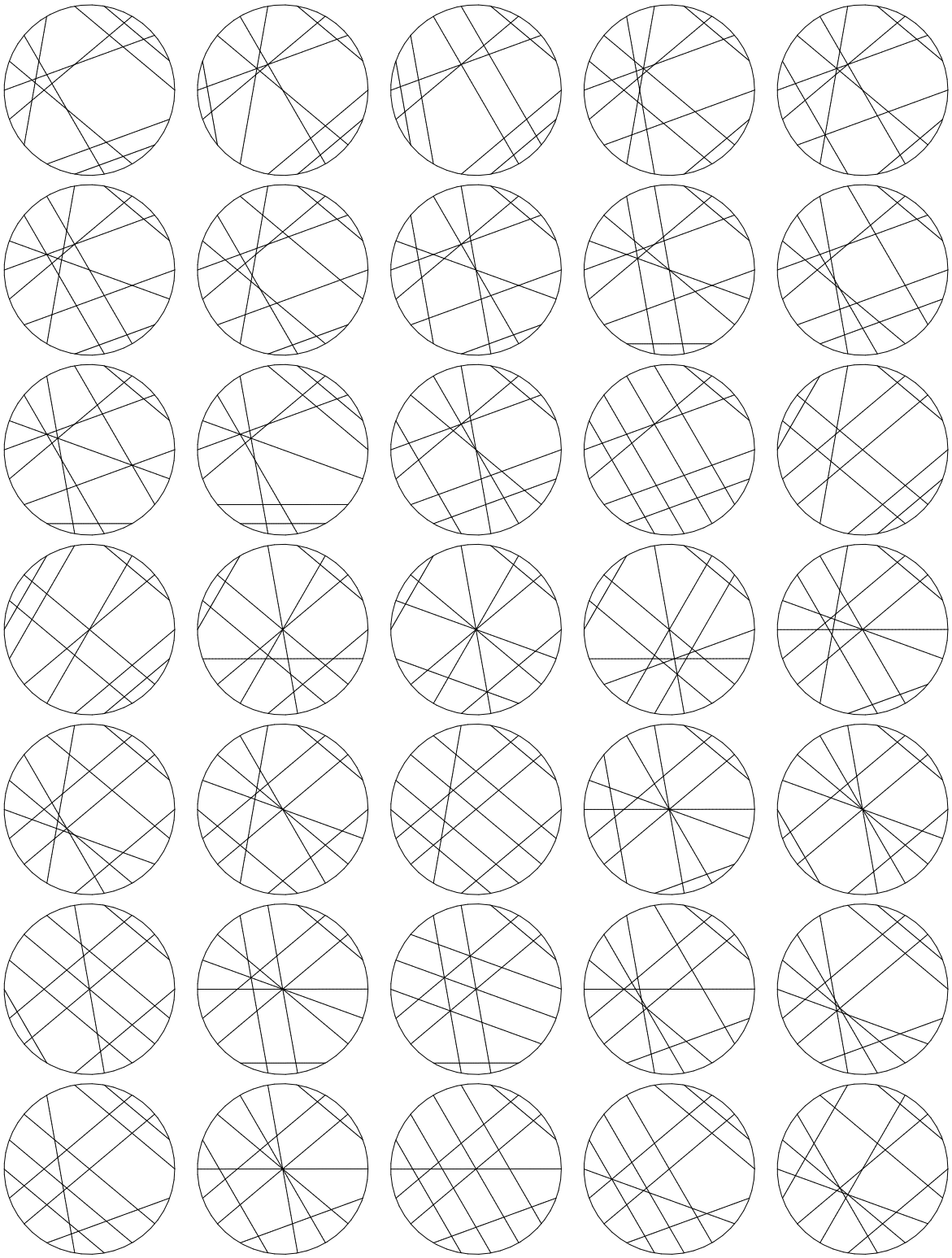


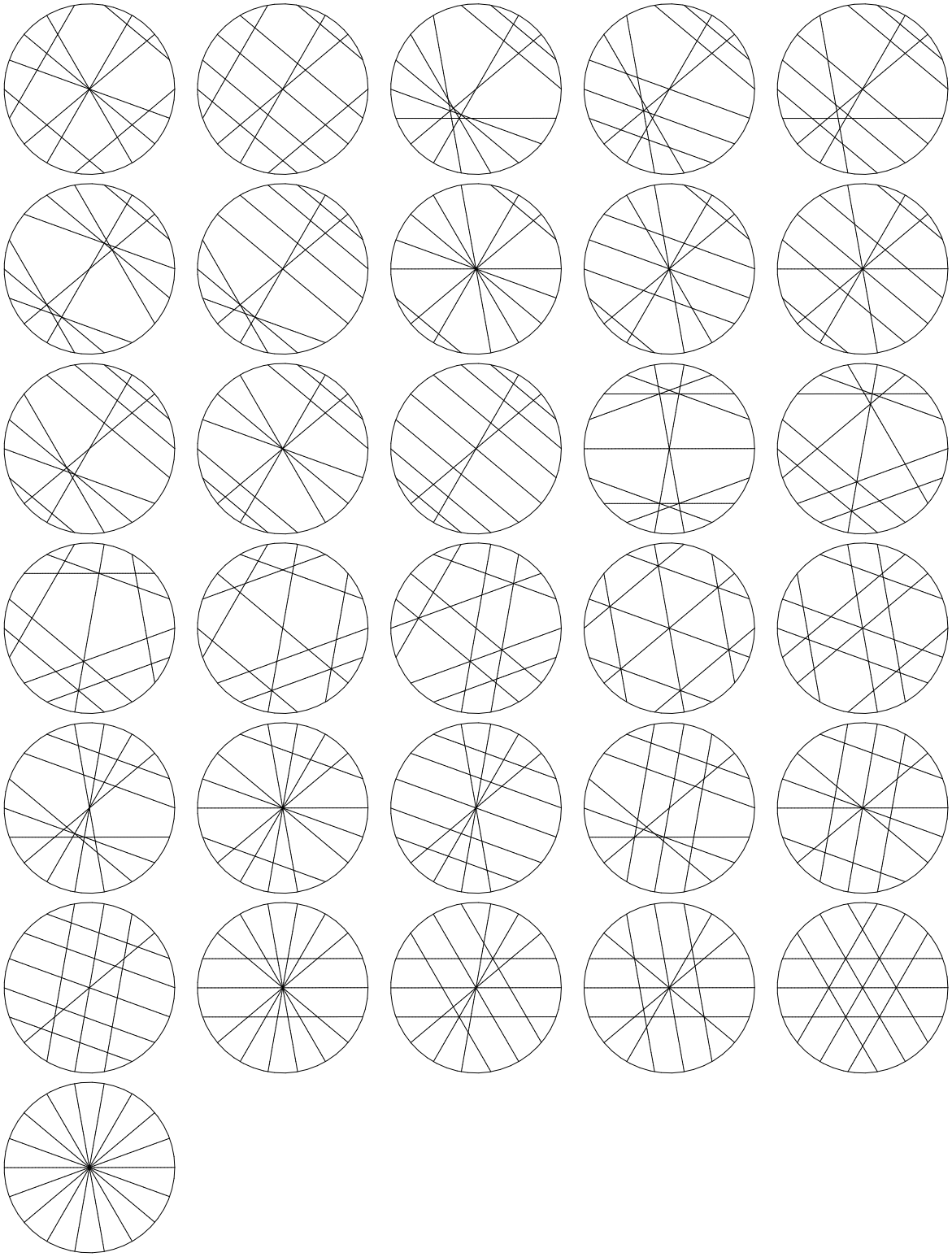
Prime Gauss diagrams with 8 **crossings.**



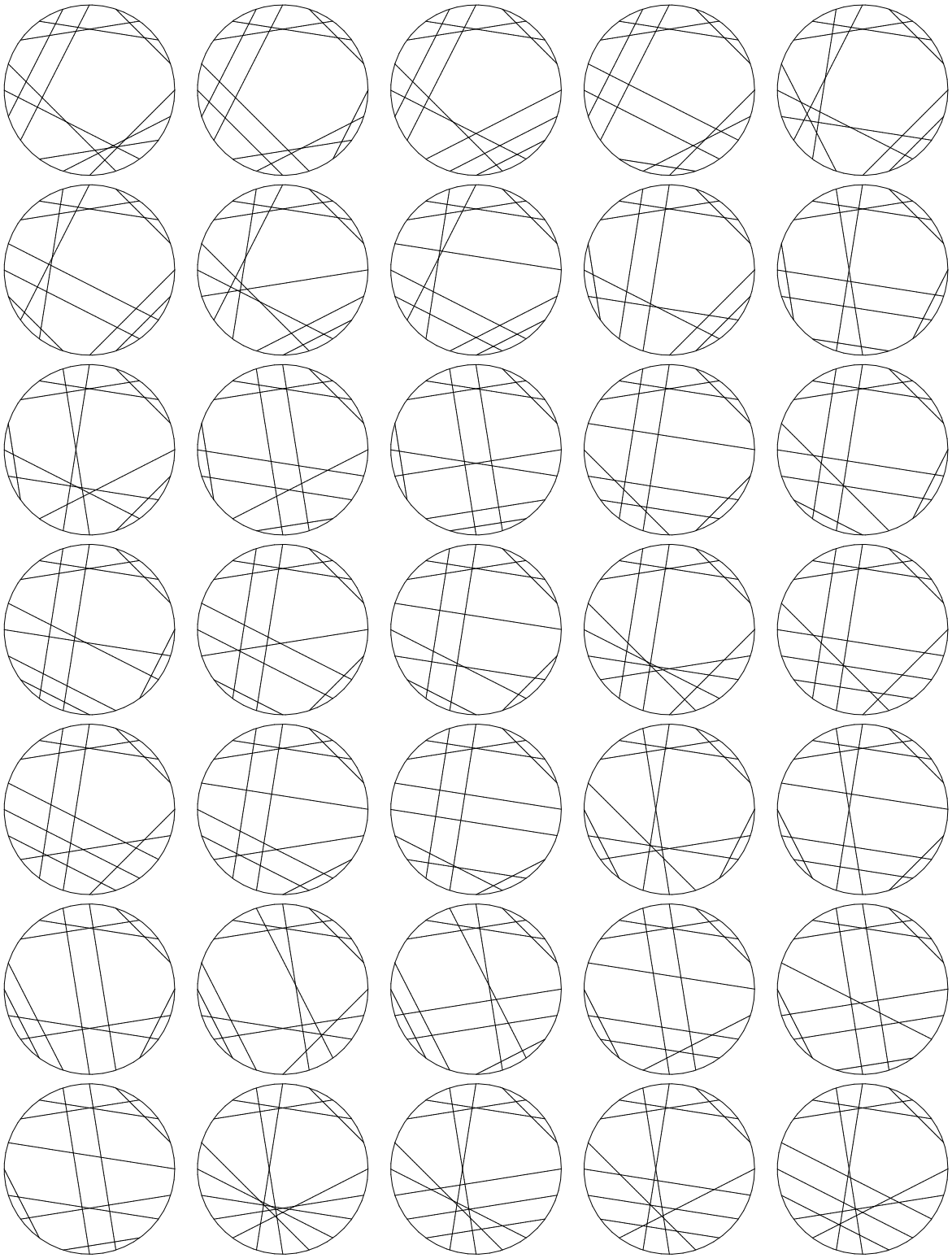
Prime Gauss diagrams with 9 crossings

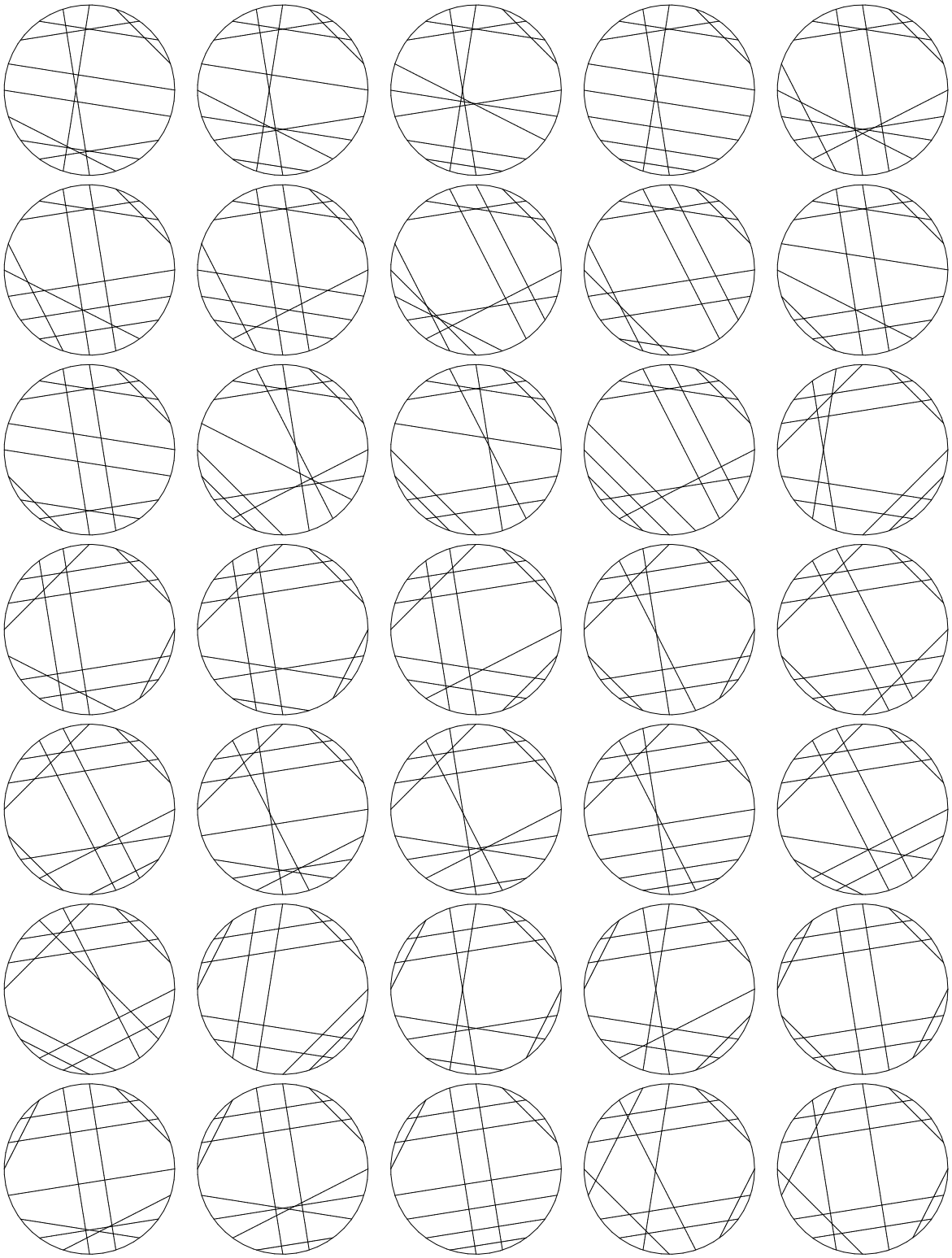


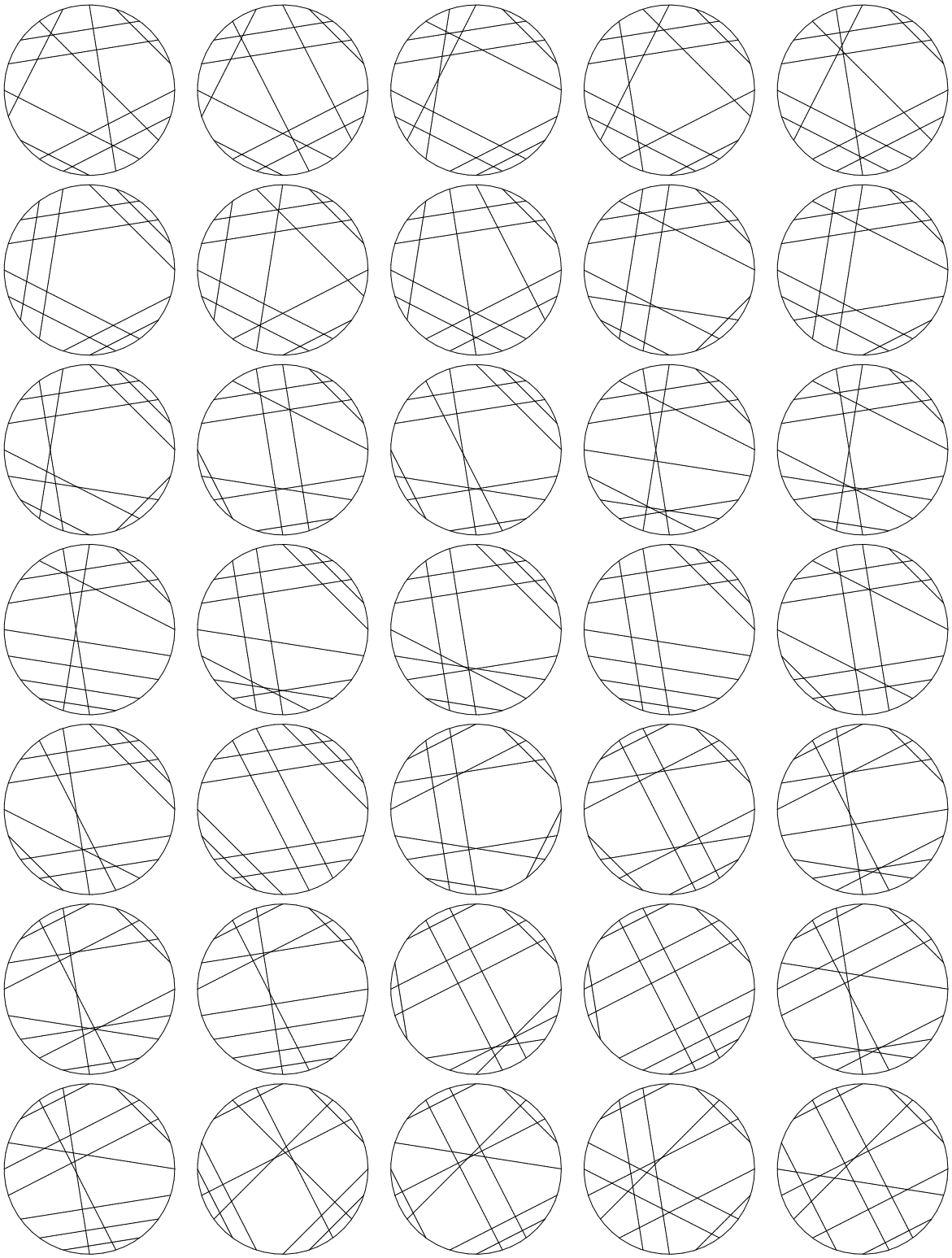


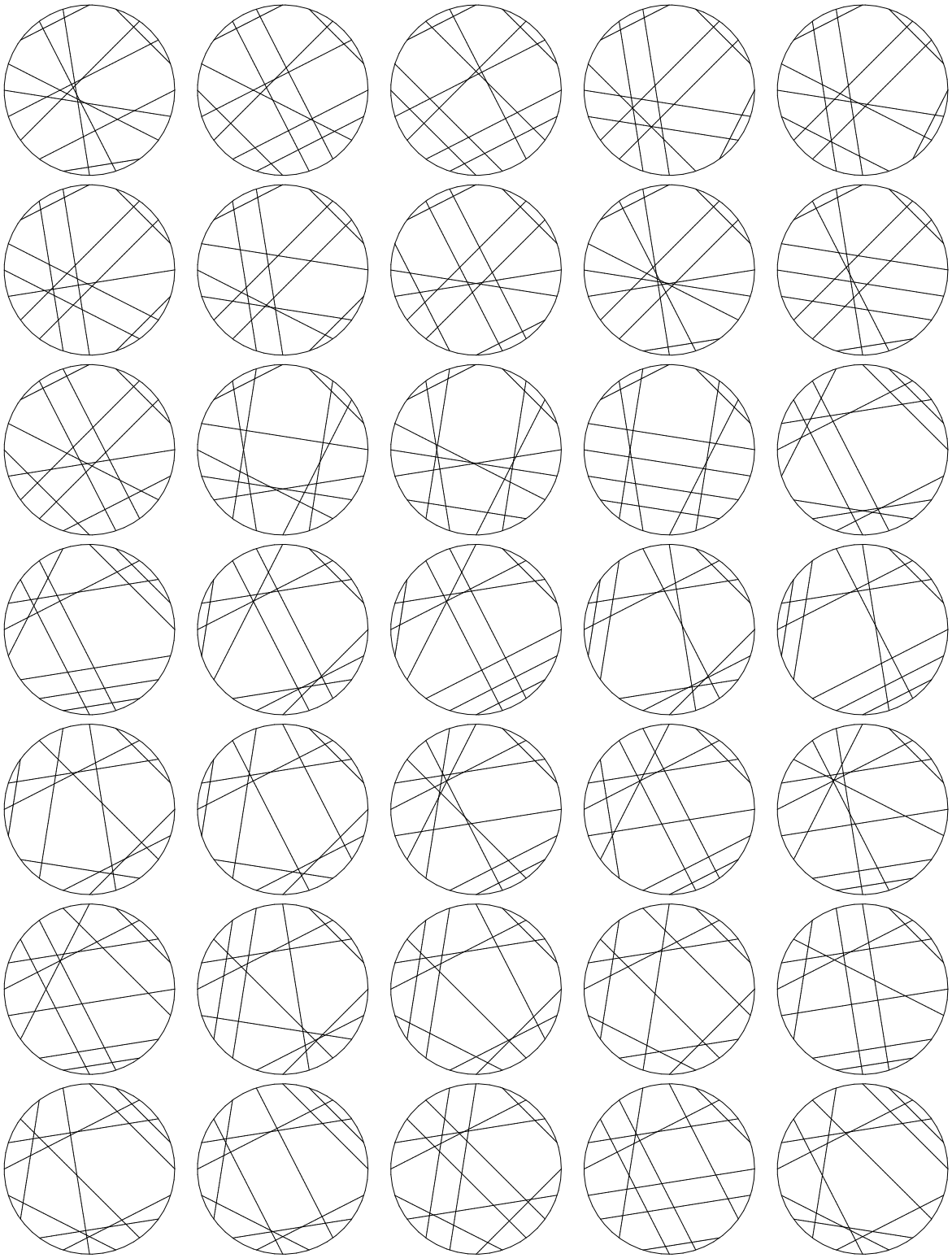


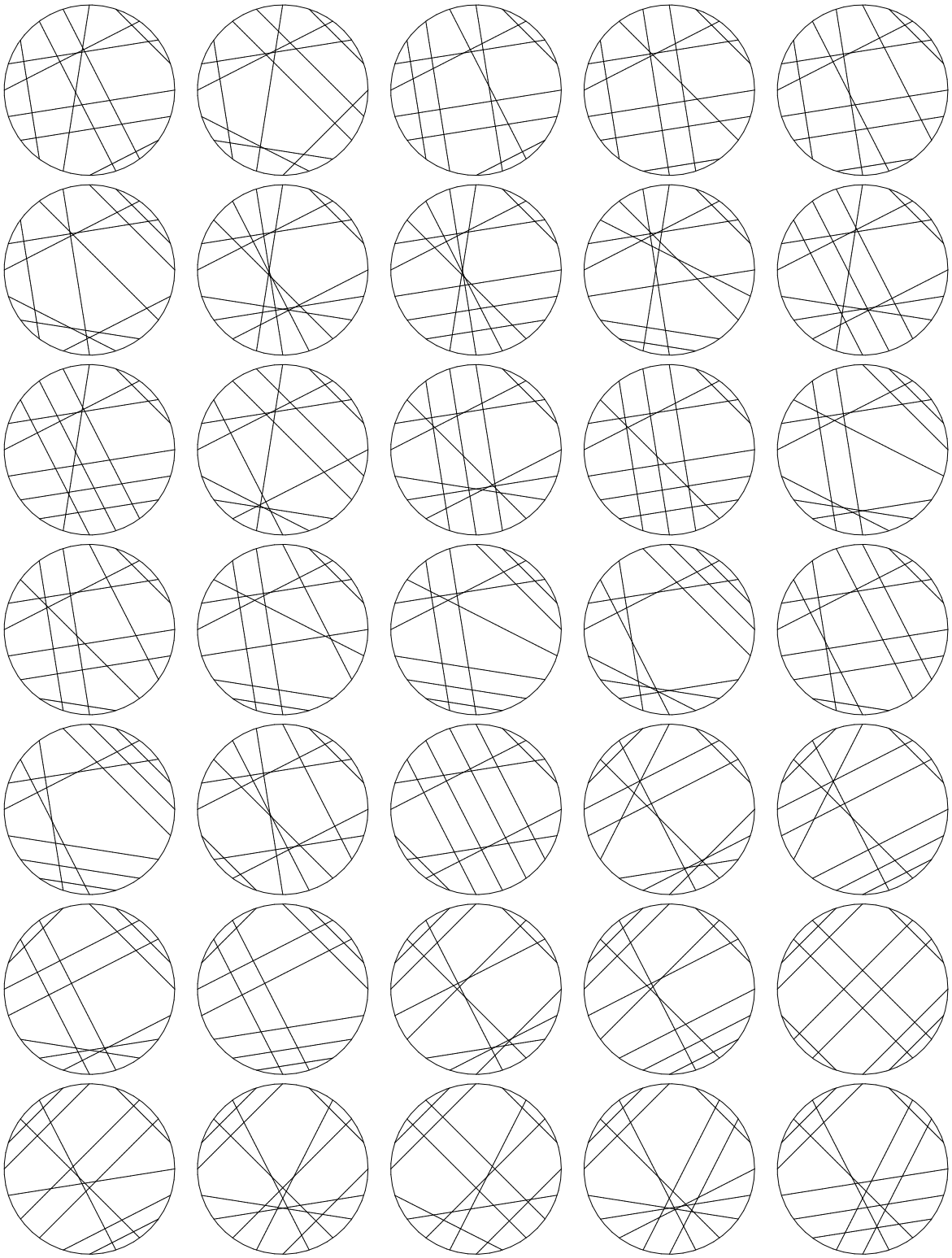
Prime Gauss diagrams with 10 **crossings (364)**

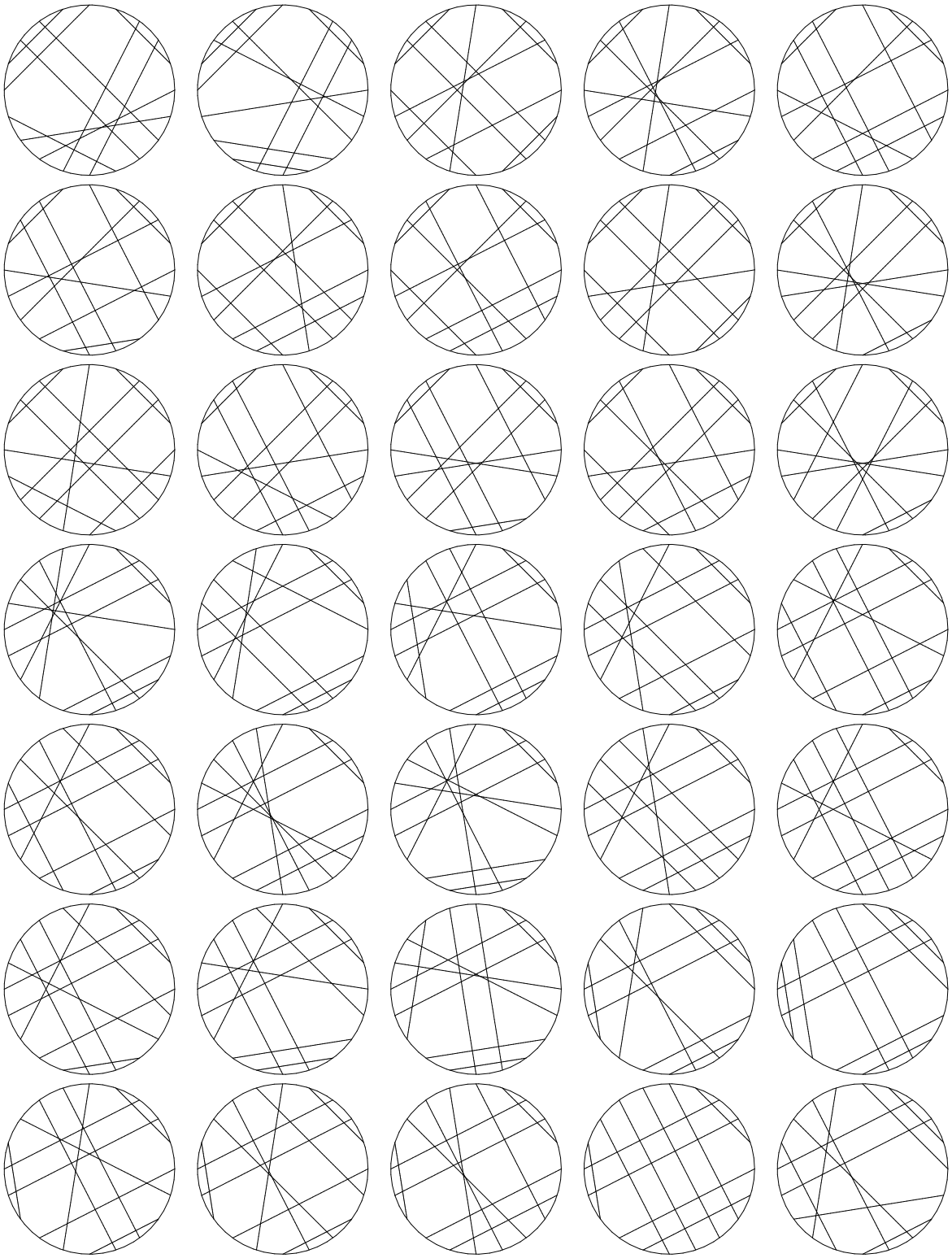


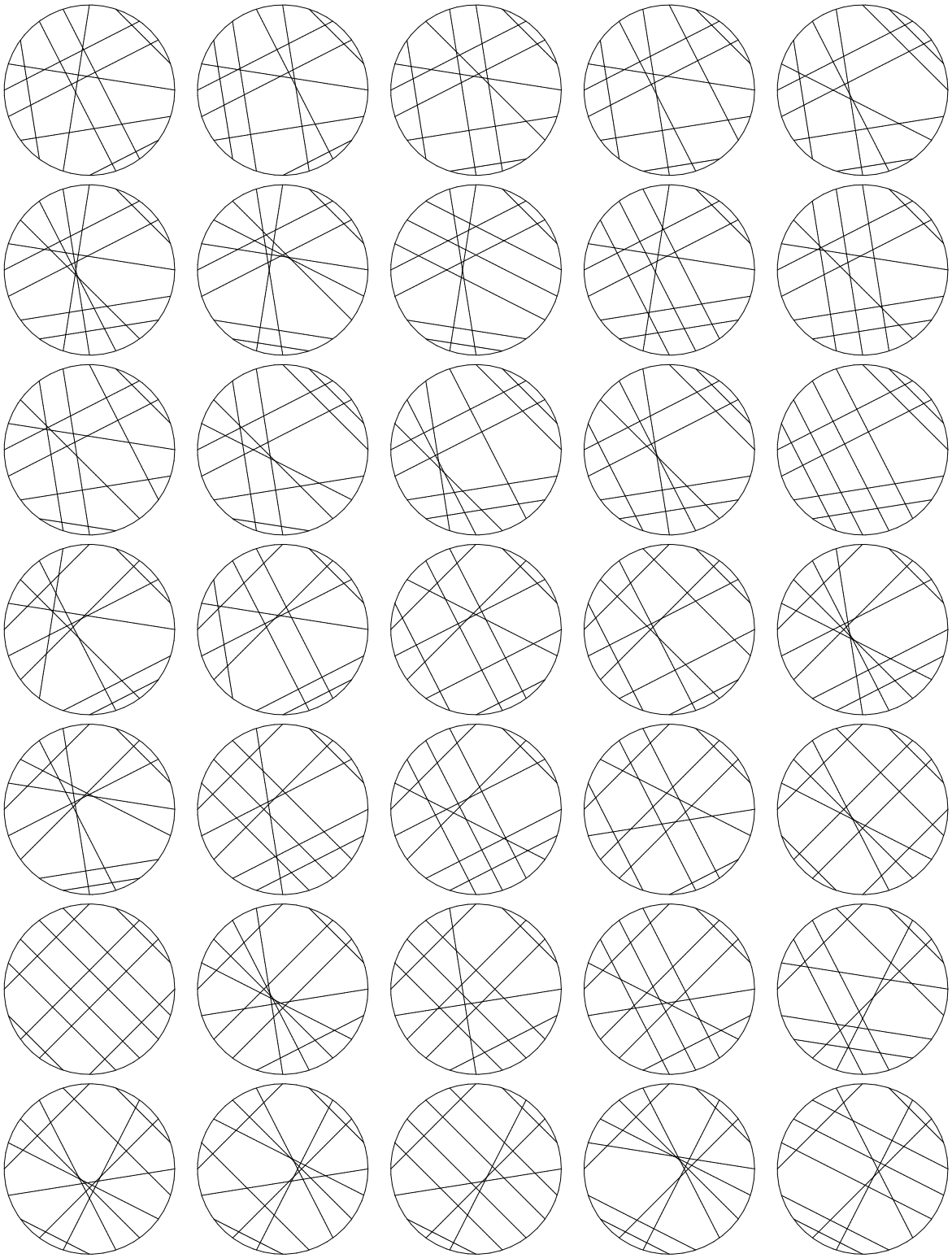


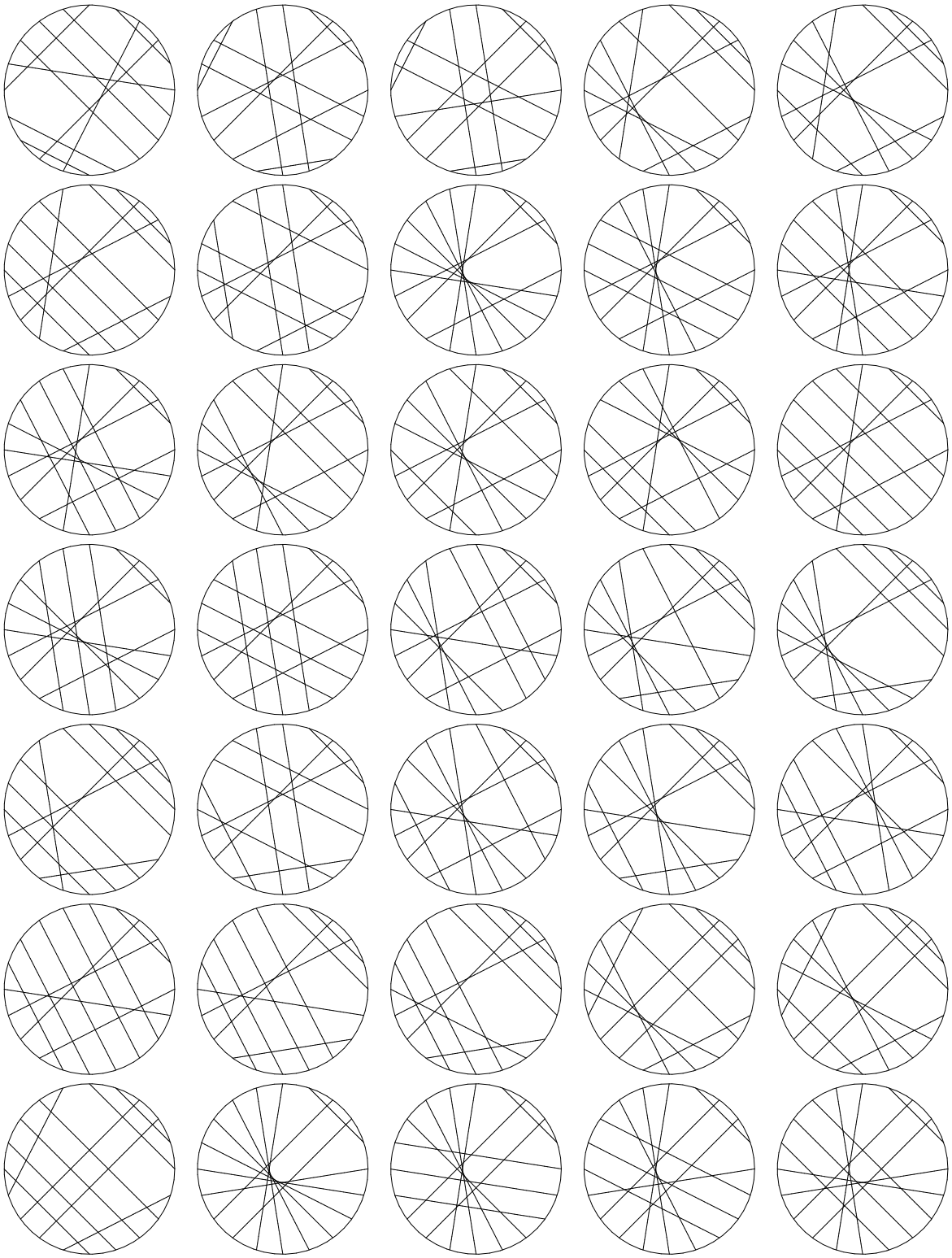


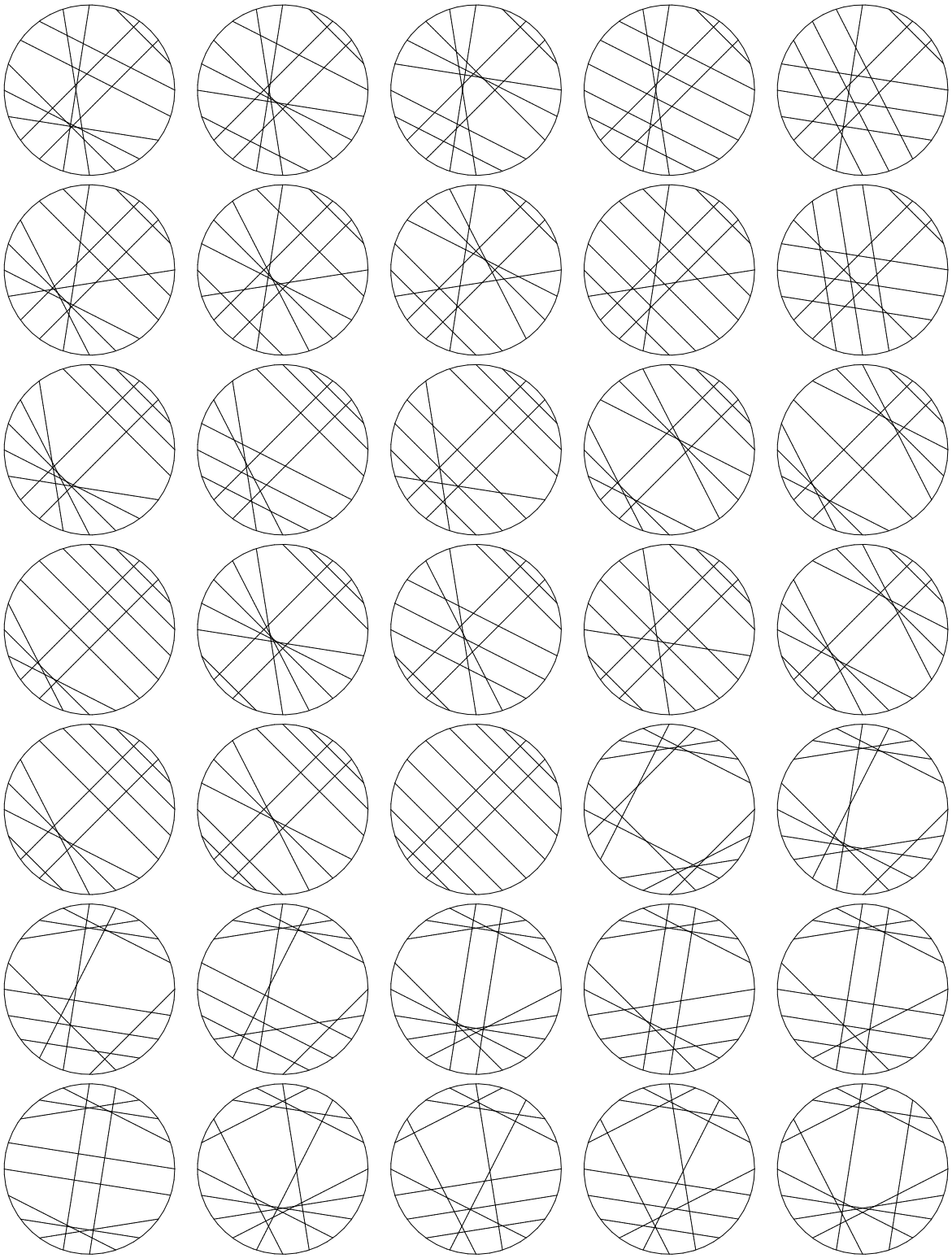


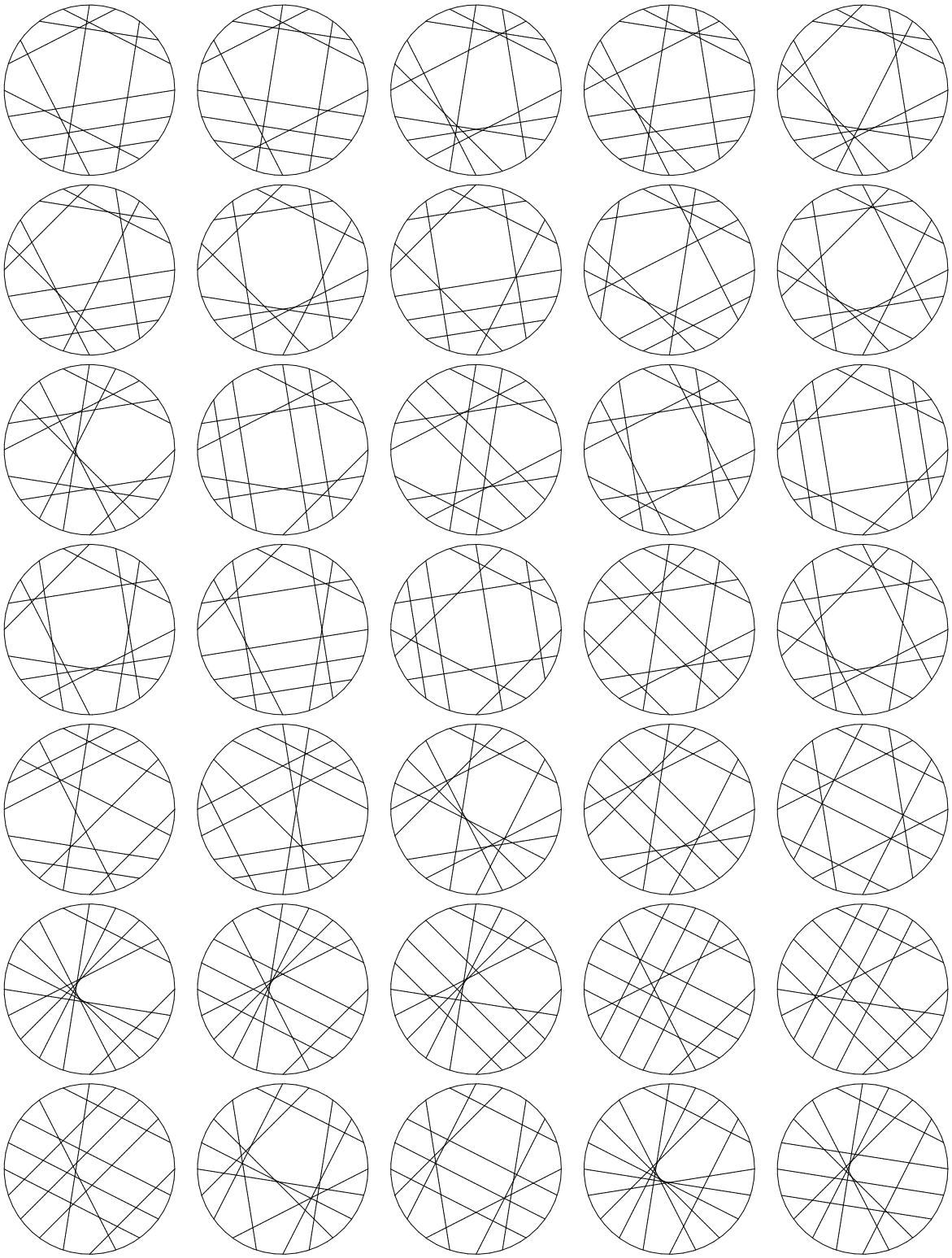


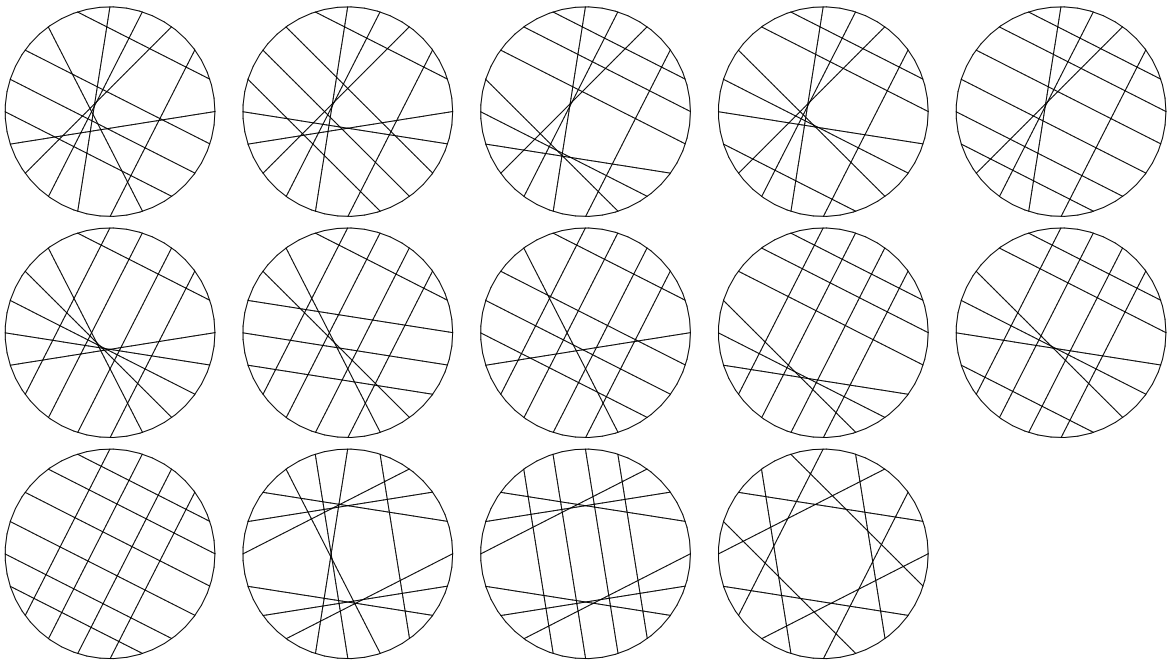












[Note: Not all subjects in the References section were, in fact referenced. We felt that a long bibliography would be a useful source list for future REUs and other individuals studying this subject.]

REFERENCES

[Ai] F. AICARDI. *Tree-like Curves,* Advances in Soviet Mathematics. **21** (1994), 1-31.

[Ad] C.C. ADAMS. *The Knot Book,* W.H. Freeman and Company (New York) 1994.

[Ar1] V. I. ARNOLD. *Plane Curves, Their Invariants, Perestroikas and Classifications,* Advances in Soviet Mathematics. **21** (1994), 33-91.

[Ar2] V. I. ARNOLD. *Remarks on the Enumeration of Plane Curves,* Amer. Math. Soc. Transl. **173** (1996), 17-32.

[Ar3] V. I. ARNOLD. *Topological invariants of plane curves and caustics,* University Lecture Series **5**. American Math Society, Providence, 1994.

[BB] A. BARKER, I. BIRINGER. *On defect of plane curves,* Summer Oregon State REU Program, 2003.

[BR1] B. BOLLOBÁS, O. RIORDAN. *A polynomial of graphs on orientable surfaces,* Proc. London Math. Soc. **83** (2001), 513-531.

[BR2] B. BOLLOBÁS, O. RIORDAN. *A polynomial of graphs on surfaces,* Math. Ann. **323** (2002), 81-96

[CE] CAIRNS, G AND ELTON D. *The planarity problem for signed Gauss words,* J. Knot Theory Ramifications, **2**(1993), 359-367.

[CHB] M. CARDWELL, R. HAUT, K. BARRESE. *Knot invariants and their implications for closed plane curves,* Summer Oregon State REU Program, 2005.

[Ca] J.S. CARTER. *Classifying immersed curves,* Proceedings of the American Mathematical Society, **111**(1991), 281-287.

[CP] S. CHMUTOV, I. PAK. *The Kauffman bracket of virtual links and the Bolloba's-Riordan polynomial*, 2006 version of the preprint arXiv:math.GT/0404475.

[DGZ] F. DUZHIN, S. GUSSEIN-ZADE. *On the number of topological types of plane curves.* Uspekhi Math. Nauk. **53**, 197-198 (1998) (english translation in *Russian Math. Surveys* **53**, 626627 (1998)).

[JZJ] J. JACOBSEN, P. ZINN-JUSTIN. *A Transfer Matrix approach to the Enumeration of Knots.* J. Knot Theor. Ramif. **11** (2002), 739-758.

[LW] X.-S. LIN, Z.-H. WANG. *Integral geometry of plane curves and knot invariants.* J. of Differential Geom. **44** (1996), 74-95.

[Luo] C. LUO. *Proof of Arnold's Conjectures About Plane Curves,* PhD thesis, Brown University, 1997.

[Po] M. POLYAK. *Invariants of curves and fronts via Gauss diagrams,* Topology, **37**(1998), 989-1009.

[SZJ] G. SCHAEFFER, P. ZINN-JUSTIN. *On the Asymptotic Number of Plane Curves and Alternating Knots.* Experiment. Math. **13** (2004), 483-493.

[Wh] H. WHITNEY. *On regular closed curves in the plane,* Compositio Mathematica **4**(1937), 276-284.

[Knot] KNOTILUS A useful knot drawing program used as a template for many of our figures. *http://srankin.math.uwo.ca/cgi-bin/retrieve.cgi/html/start.html*

OHIO STATE UNIVERSITY
E-mail address: chmutov@mps.ohio-state.edu

COLBY COLLEGE
E-mail address: tahulse@colby.edu

WHITMAN COLLEGE
E-mail address: lumaa@whitman.edu

OREGON STATE UNIVERSITY
E-mail address: rowellp@onid.orst.edu