CHAPTER

winding number rotation number reasonably tame index numerum complicationum Amplitudo I and the Art des Viel care integ der Gran von fingular rder tangent winding number curliness Whitney degree Whitney index total signed curvature index

winding number

Generic and Regular Curves

Turning and turning in the widening gyre The falcon cannot hear the falconer; Things fall apart; the centre cannot hold; Mere anarchy is loosed upon the world

- William Butler Yeats, "The Second Coming" (1921)

In this chapter we study two numerical invariants of closed curves in the plane, called the *winding number* (with respect to a point) and the *rotation number*. Informally, the winding number of a closed curve is the number of times the curve winds around a point, and the rotation number is the number of rotations made by the tangent vector during one traversal of the curve. The formal definitions of these invariants for *arbitrary* curves are subtle, but for reasonably tame curves, they are intuitive and can be computed quickly by elementary algorithms.

Winding and rotation numbers have been known at least informally since 1770, when they were discussed in a seminal paper of Meister [22]. Especially in complex analysis, the winding number is also known as the *index* of a curve with respect a point [28, 33]. The rotation number has been given many other names, including numerum complicationum [22], Amplitudo [11], l'espèce du polygone [31], die Art des Vielecks [49], curvatura integra [6], der Grad von f [16], angular order [26, 27], tangent winding number [8, 14, 42], curliness [17], Whitney degree [18], Whitney index [33], total signed curvature [12], (unfortunately) *index* [3,4], and (even more unfortunately) winding number [21, 33, 37, 46].

11

12

13

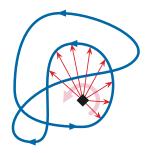
14

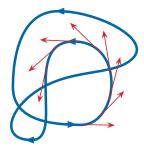
18

19

To avoid difficulties with pathological cases, we explicitly consider only closed curves that have a finite number of of self-intersection points, each of which is a pairwise

generic closed curve normal closed curve stable closed curve vertices edges loop compactness argument winding number wind alpha q tame





8

9

11

12

13

14

16

17

18

19

22

23

A curve with winding number 1 around a point; the same curve has turning number 0.

crossing; we call such curves **generic**. (Whitney called generic curves **normal** [47]; other authors call them **stable** [30].) Any curve can be made generic by an arbitrarily small perturbation. A generic closed curve with at least one self-intersection point can be represented combinatorially by a planar embedding of a planar graph in which every vertex has degree 4. For this reason, we consistently refer to the self-intersection points of a generic curve as its **vertices** and the subpaths between successive vertices as its **edges**. Conversely, any 4-regular plane graph describes a finite set of generic closed curves that intersect transversely at a finite number of points.



Stick to loops, or talk about cycles everywhere instead? We need the basepoint to talk about signed vertices, but not for anything else.

5.1 Winding Numbers

Recall that a *loop* in the plane is a continuous function $\gamma \colon [0,1] \to \mathbb{R}^2$ such that $\gamma(0) = \gamma(1)$. For any loop γ and any point q = (a,b) that is not in the image of γ , there are continuous functions $\rho \colon [0,1] \to \mathbb{R}^+$ and $\theta \colon [0,1] \to \mathbb{R}$ such that

$$\gamma(t) = (a + \rho(t)\cos(2\pi\theta(t)), b + \rho(t)\sin(2\pi\theta(t))).$$

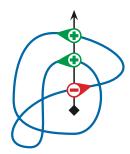
The function ρ is uniquely given by $\rho(t) = \|\gamma(t) - q\|$. Proving that an appropriate angle function θ exists for *arbitrary* loops requires a compactness argument similar to Lemma $\ref{lem:property}$; the function θ is unique up to addition of an arbitrary integer. The **winding number** of γ around q, denoted **wind**(γ , q), is the quantity $wind(\gamma, q) := \theta(1) - \theta(0)$. Intuitively, $wind(\gamma, q)$ is the number of times that γ winds counterclockwise around q. The fact that γ is a loop immediately implies that $wind(\gamma, q)$ is an integer.

Signed Crossings

If the loop γ is reasonably tame, the winding number $wind(\gamma,q)$ can be computed as follows. Let ρ be an arbitrary infinite ray based at q that crosses the loop γ a finite number of times. (The existence of such a ray can be taken as the definition

of "reasonably tame"!) Call a crossing **positive** if γ crosses ρ from right to left, and **negative** otherwise. Results from previous chapters imply that the winding number $wind(\gamma,q)$ is equal to the number of positive crossings minus the number of negative crossings; we leave the details of the proof as an exercise for the reader.

positive crossing negative crossing exercise for the reader Umlaufsatz



Two positive crossings and a negative crossing.

This characterization immediately implies a linear-time algorithm to compute winding numbers of polygonal loops that is only slightly more complex than the PointinPolygon algorithm from Chapter ??. The subroutine CrossSign returns +1 for each positive crossing and -1 if for each negative crossing. (The final return statement correctly handles the case where q lies on the loop, even when the loop is not generic; see below.)

```
\frac{\text{WindingNumber}(P[0..n-1],q):}{wind \leftarrow 0}
P[n] \leftarrow P[0]
for i \leftarrow 0 to n-1
wind \leftarrow wind + \text{CrossSign}(q,P[i],P[i+1])
return wind
```

3

5

6

10

14

16

17

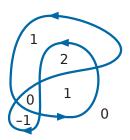
```
\frac{\text{CROSSSIGN}(q,r,s):}{\text{if } r.x < s.x}
sign \leftarrow -1
swap \ r \leftrightarrow s
else
sign \leftarrow +1
\text{if } (q.x < s.x) \text{ or } (q.x \ge r.x)
\text{return } 0
else \text{ if } \Delta(q,r,s) < 0
\text{return } 0
else \text{ if } \Delta(q,r,s) > 0
\text{return } sign
else
\text{return } sign/2
```

Computing the winding number of a polygon around a point.

An argument similar to the proof of Lemma ?? implies that any two points in the same component of $\mathbb{R}^2 \setminus \operatorname{im} \gamma$ define the same winding number—as we continuously translate the ray from one basepoint to the other, crossings may appear and disappear, but always in matched pairs, one positive and one negative. In particular, if p is in the unbounded component, then $\operatorname{wind}(\gamma,p)=0$. If γ is a simple loop, the Jordan curve theorem implies that the winding number with respect to every interior point is either -1 or 1, depending on whether the curve turns clockwise or counterclockwise; Hopf [16] called this observation the $\operatorname{Umlaufsatz}$ ("Circulation Theorem"). More generally, winding

uncrossing moves Seifert decomposition exercise for the reader free homotopy numbers within adjacent components of $\mathbb{R}^2 \setminus \operatorname{im} \gamma$ differ by exactly 1, with the larger winding number on the left side of the curve.

Alexander [1] used this characterization of the winding number as its definition; in fact, it is already implicit in the work of Meister [22] and Möbius [24].



Alexander numbering: Winding numbers of points in each region.

The definition of winding number can be extended to points that lie on the curve. For any point $\gamma(t)$ that is not a vertex, we define $wind(\gamma, \gamma(t))$ to be the average of the winding numbers on either side of $\gamma(t)$. For any vertex x, we define $wind(\gamma, x)$ to be the average of the four regions incident to x; if two of these incident regions are identical, that region is counted twice in the average. (These cases are properly handled by the last return statement in the subroutine WhichSide.)

9

11

12

13

15

16

17

18

19

23

24

25

26

27

28

29

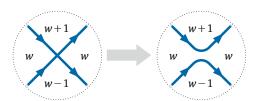
Cycle Decomposition

We can also characterize winding numbers by decomposing the loop into *simple* loops using *uncrossing moves*. An uncrossing move modifies the curve only in a small neighborhood of an intersection point, by replacing two intersecting subpaths with two new non-crossing paths. Because the curves are directed, the choice of new paths is unique, as shown below. Uncrossing a self-intersection point of a loop splits it into two smaller loops; uncrossing an intersection point of two loops merges them into one larger loop. Uncrossing every self-intersection decomposes γ into a collection of *non-crossing* simple loops, called the *Seifert decomposition* of γ [38].

It is not hard to show that starting from any generic set of closed curves, uncrossing moves do not change the sum of the winding numbers; we leave the proof as an exercise for the reader. Thus, if we can decompose any loop γ into *simple* loops $\gamma_1, \ldots, \gamma_k$ by uncrossing moves, then $wind(\gamma, q)$ is the number of counterclockwise loops γ_i that contain q minus the number of clockwise loops γ_i that contain q.

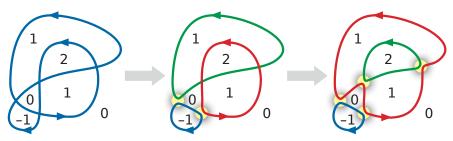
Free Homotopy

If we continuously deform γ without touching the point q, the winding number $wind(\gamma,q)$ remains constant. Recall that a *free homotopy* between two loops γ and δ in $\mathbb{R}^2 \setminus \{q\}$ is a function $h \colon [0,1]^2 \to \mathbb{R}^2 \setminus \{p\}$ such that $h(0,t) = \gamma(t)$ and $h(1,t) = \delta(t)$ for all t, and h(s,0) = h(s,1) for all s. As long as γ remains reasonably tame during a



compactness argument
One can mechanically
verify
regular curve
regular closed curve

An uncrossing move. The curve does not change outside the dotted circle.



Winding numbers are preserved by uncrossing.

free homotopy, crossings with the fixed ray r appear and disappear in positive-negative pairs. Proving that *arbitrary* free homotopies preserve the winding number requires another compactness argument, first provided by Hopf [16].

Hopf also proved that any two loops with the same winding number around a fixed point are freely homotopic, using the following direct construction. Let $\zeta: [0,1] \to S^1$ denote the loop $\zeta(t) = (\cos 2\pi t, \sin 2\pi t)$ that winds once counterclockwise around the unit circle at constant speed. For any integer k, let $\zeta^k(t) = \zeta(kt) = \zeta(kt \mod 1)$; the loop ζ^k clearly has winding number k around the origin. Fix a loop $\gamma: [0,1] \to \mathbb{R}^2 \setminus 0$. Following the definition of the winding number, choose continuous functions $r: [0,1] \to \mathbb{R}^+$ and $\theta: [0,1] \to \mathbb{R}$ such that $\gamma(t) = r(t) \cdot \zeta(\theta(t))$. One can mechanically verify that the function $h: [0,1]^2 \to \mathbb{R}^2 \setminus 0$, defined by setting

$$h(s,t) := (s + (1-s)r(t)) \cdot \zeta(s \cdot t \cdot wind(\gamma,0) + (1-s)\theta(t)),$$

is a free homotopy from γ to $\zeta^{wind(\gamma,0)}$. We conclude that γ is freely homotopic to any loop δ such that $wind(\gamma,0) = wind(\delta,0)$.

Theorem 5.1 (Hopf [16]). Two loops γ and δ are freely homotopic in $\mathbb{R}^2 \setminus p$ if and only if $wind(\gamma, p) = wind(\delta, p)$.

5.2 Rotation Numbers

Regular Curves

2

4

6

8

9

13

14

15

16

17

18

19

21

Following Whitney [47], we call a loop γ in the plane **regular** if it has no sharp corners or cusps. More formally, a **regular closed curve** is a function $\gamma: [0,1] \to \mathbb{R}^2$ satisfying the following conditions:

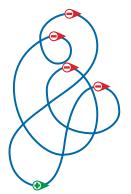
rotation number extreme point happy point sad point

- $\gamma(0) = \gamma(1)$;
- γ has a well-defined, continuous derivative $\gamma': [0,1] \to \mathbb{R}^2$;
- $\gamma'(0) = \gamma'(1)$; and
- $\gamma'(t) \neq (0,0)$ for all *t*.

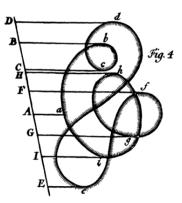
More succinctly, a regular closed curve is a differentiable loop $\gamma \colon [0,1] \to \mathbb{R}^2$ whose derivative γ' is a loop that avoids the origin.

The **rotation number** of a regular closed curve γ is the winding number of its derivative γ' around the origin, or more intuitively, the total number of times that a person walking once around the curve turns counterclockwise.

For reasonably tame regular curves, the rotation number can be more easily described as follows. Call a point $\gamma(t)$ *extreme* if the derivative vector $\gamma'(t)$ points in some fixed direction. An extreme point $\gamma(t)$ is *happy* if γ lies locally to the left of the tangent ray, and *sad* if γ is locally to the right of the tangent ray; for a generic direction, every extreme point is either happy or sad. If the fixed direction is parallel to the positive x-axis, happy points have neighborhoods that curve up \sim and sad points have neighborhoods that curve down \sim . Our earlier characterization of winding numbers by positive and negative crossings implies that $rot(\gamma)$ is the number of happy points minus the number of sad points. This characterization of the rotation number was certainly known to Gauss [11] but may have been known even earlier [22].



Four sad points and one happy point.



9

11

12

13

14

15

16

17

18

19

21

22

23

24

25

28

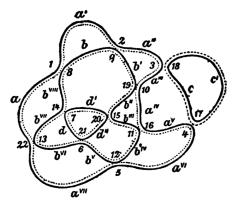
A figure from Meister [22]

Gauss [11] also observed that the rotation number of any loop γ is equal to the sum of the rotation numbers of the simple loops in the Seifert decomposition of γ ; see also Kauffman [17]. Any simple loop has rotation number -1 or +1, depending on whether it is oriented clockwise or counterclockwise; the rotation number is equal to the winding number around any interior point.

Formally, to maintain a set of regular curves, we must smooth the curves slightly near the former intersection point after the uncrossing move. The precise details of the smoothing are surprisingly unimportant. Suppose we smooth by replacing the sharp corners with tiny circular arcs. These smoothing arcs subtend *approximately* the same

angle, so uncrossing and smoothing changes the total rotation number by an arbitrarily small amount. But the rotation number is an integer, so in fact there is no change at all!

positive vertex negative vertex writhe



A Seifert decomposition into one negative and three positive cycles, from Gauss [11]

Signed Vertices

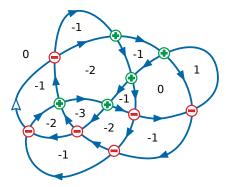
5

8

9

12

Alternatively, we can compute the rotation number of a curve directly from its graph representation, without decomposing it into cycles. Without loss of generality, we assume that the basepoint $\gamma(0) = \gamma(1)$ is not a vertex of γ . Let $x = \gamma(u) = \gamma(v)$ be a vertex of γ , for some 0 < u < v < 1. We call x a **positive vertex** if the tangent vector $\gamma'(u)$ is counterclockwise from $\gamma'(v)$, and a **negative vertex** otherwise. Equivalently, a vertex x is positive if the winding number of γ with respect to a point moving just next to the curve increases *the first time* it passes by x, and negative otherwise.



Five positive and six negative vertices; the white arrowhead on the far left is the basepoint.

For any vertex x, define sgn(x) = 1 if x is a positive vertex, and sgn(x) = -1 if x is a negative vertex. Following standard practice in knot theory, we define the **writhe** of γ

to be the number of positive vertices minus the number of negative vertices; that is,

$$writhe(\gamma) := \sum_{\text{vertices } x \text{ of } \gamma} \operatorname{sgn}(x).$$

The writhe depends on the basepoint used to express it as a loop; sliding the basepoint over a vertex changes the sign of that vertex and thus changes the writhe by 2. To simplify notation, let $wind_0(\gamma) = wind(\gamma, \gamma(0))$ denote the average of the winding numbers just to the left and right of $\gamma(0)$.

The following theorem was first proved by Titus [41, Theorem 2], generalizing a direct analytical argument of Whitney [47, Theorem 2] for the special case where $\gamma(0)$ is incident to the outer face, so $wind_0(\gamma) = \pm \frac{1}{2}$. However, the result was actually known to Gauss [11] about a century before Whitney. Quine [34] describes a generalization of this theorem to non-generic curves.

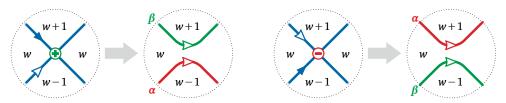
Theorem 5.2. For any generic regular curve γ , we have $rot(\gamma) = 2 wind_0(\gamma) + writhe(\gamma)$.

Proof: We prove the theorem by induction on the number of vertices. As a base case, first suppose γ is a simple loop. Trivially $writhe(\gamma) = 0$. If the interior of γ lies to the left of γ , then $rot(\gamma) = 1$ and $wind_0(\gamma) = \frac{1}{2}$; Otherwise, $rot(\gamma) = -1$ and $wind_0(\gamma) = -\frac{1}{2}$.

Now suppose γ is not simple. Let x be the closest vertex of γ to the basepoint; that is, let u be the smallest value such that $\gamma(u) = \gamma(\nu)$ for some $0 < u < \nu < 1$, and let $x = \gamma(u) = \gamma(\nu)$. An uncrossing move at x breaks γ into two closed curves α and β such that $rot(\gamma) = rot(\alpha) + rot(\beta)$; without loss of generality, let α contain the old basepoint $\gamma(0)$. The inductive hypothesis immediately implies that

$$rot(\gamma) = 2wind_0(\alpha) + 2wind_0(\beta) + writhe(\alpha) + writhe(\beta).$$

Parametrize α and β as loops with basepoints as close as possible to the former vertex x.



Uncrossing positive and negative vertices; white arrowheads indicate basepoints.

Call a vertex of γ bichromatic if it is an intersection point of α and β . As we increase the parameter t from 0 to 1, the winding number $wind(\alpha, \beta(t))$ increases by 1 at each positive bichromatic vertex and decreases by 1 at each negative bichromatic vertex. But $wind(\alpha, \beta(0)) = wind(\alpha, \beta(1))$, so there must be exactly the same number of positive and negative bichromatic vertices. Every vertex of α of β is a vertex of γ with the same sign. The only vertex of γ that is neither bichromatic nor a vertex of α or β is the removed vertex x. We conclude that $writhe(\gamma) = writhe(\alpha) + writhe(\beta) + sgn(x)$.

Now consider the initial winding number terms. If x was a positive vertex and $wind_0(\gamma) = w - \frac{1}{2}$, as illustrated above, then $wind_0(\alpha) + wind_0(\beta) = w$. Similarly, if x was a positive vertex and $wind_0(\gamma) = w + \frac{1}{2}$, then $wind_0(\alpha) + wind_0(\beta) = w$. In either case, we have $2wind_0(\gamma) = 2wind_0(\alpha) + 2wind_0(\beta) - sgn(x)$, and the theorem follows immediately.

regular homotopy regularly homotopic gamma homotopic delta

We can now take Theorem 5.2 to be the *definition* of the rotation number for arbitrary generic curves, including pathological curves for which no definition based on curvature or angles is possible. The theorem has two immediate useful corollaries.

Corollary 5.3 (Gauss [11]). Every generic closed curve with rotation number 0 has at least one vertex. For any integer $r \neq 0$, every generic closed curve with rotation number r has at least |r| - 1 vertices.

Corollary 5.4. A generic curve has an even number of vertices if and only if its rotation number is odd.

5.3 Regular Homotopy

1

2

3

4

5

6

8

9

11

13

14

18

19

25

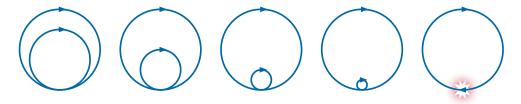
26

28

29

A **regular homotopy** is a function $h: [0,1]^2 \to \mathbb{R}^2$ such that for all s, the function $t \mapsto h(s,t)$ is a regular closed curve, and the partial derivative $\partial h/\partial t$ is a free homotopy between loops in $\mathbb{R}^2 \setminus 0$. Two regular closed curves γ and δ are **regularly homotopic**, denoted $\gamma \simeq \delta$, if there is a regular homotopy h such that $h(0,\cdot) = \gamma$ and $h(1,\cdot) = \delta$.

Some early sources [21], including Whitney's original paper [47], omit the partial derivative condition form the definition of regular homotopy, but the following construction shows that it is necessary. Consider a closed curve that winds once around the unit circle and once around a smaller circle tangent at its lowest point. There is a free homotopy from this curve to the unit circle that shrinks the inner circle to its point of tangency. The deforming curve is always regular; however, at the instant when the inner circle vanishes, the tangent vector at the top point of the inner circle must either vanish or change discontinuously.



A homotopy through regular closed curves, but not a regular homotopy.

It is not hard to see that regularly homotopic curves have equal rotation numbers; indeed, there is a two-line proof: For any regular homotopy h from γ to δ , the partial derivative $\partial h/\partial t$ is a (free) homotopy from γ' to δ' that avoids the origin. Thus, if γ

Whitney-Graustein theorem

and σ are regularly homotopic, their derivatives are homotopic in $\mathbb{R}^2 \setminus 0$, so γ and δ have the same rotation number.

Surprisingly, the converse is true as well. The following result is commonly called the Whitney-Graustein theorem, because it appears in a seminal paper of Whitney [47], who attributes both the theorem and its proof to the geometer William C. Graustein. However, the theorem was first proved more than 30 years earlier by Boy, in the same PhD thesis where he describes the immersion of the projective plane now known as Boy's surface [5,6]. Moreover, the result may have been known, at least informally, as early as Meister [22].²

The Whitney-Graustein Theorem. Two regular closed curves in \mathbb{R}^2 are regularly homotopic if and only if their rotation numbers are equal.

Proof (Graustein): Let γ and δ be regular closed curves with the same rotation number. Without loss of generality, we assume that both γ and δ have arc-length 1, and that are parametrized by arc length, that is, $\|\gamma'(t)\| = \|\delta'(t)\| = 1$ for all t. (We can scale and reparametrize each curve via regular homotopy if necessary.)

Because γ and δ have the same rotation number, their derivatives γ' and δ' have the same winding number around the origin, and are therefore homotopic in the unit circle S^1 (not just in $\mathbb{R}^2 \setminus 0$). Let $h' \colon [0,1]^2 \to S^1$ be a homotopy from γ' to δ' . If necessary, perturb h' so that every loop $h'(s,\cdot)$ is non-constant; this perturbation is only necessary if $rot(\gamma) = 0$.

A loop $\alpha \colon [0,1] \to \mathbb{R}^2 \setminus 0$ is the derivative of a regular closed curve if and only if its center of mass is the origin: $\int_0^1 \alpha(t) dt = 0$. The "center of mass" function $c \colon [0,1] \to \mathbb{R}^2$, defined by setting $c(s) := \int_0^1 h'(s,t) dt$, is a loop whose basepoint is the origin. For all s, the loop $h'(s,\cdot)$ lies on S^1 and is non-constant, so its center of mass c(s) lies in the open interior of S^1 . In particular, $h'(s,t) \neq c(s)$ for all s and t. Thus, the function $h^* \colon [0,1]^2 \to \mathbb{R}^2 \setminus 0$ defined by $h^*(s,t) := h'(s,t) - c(s)$ is a homotopy from γ to δ through derivatives of regular closed curves. We conclude that the function $h \colon [0,1]^2 \to \mathbb{R}^2$ defined by $h(s,t) := \int_0^t h^*(s,u) du$ is a regular homotopy from γ to δ .

5.4 Combinatorial Homotopy

When reasoning about equivalence classes of generic curves in the plane, it is often more convenient to use the following combinatorial framework, first employed (at least implicitly) by Boy [5,6] and later developed by Titus and Francis [9,10,43]. A similar framework developed independently by Alexander and Briggs [2] and Reidemeister [35,36] is now the standard definition of knot equivalence. More recent work of Arnold [3,4] significantly extended this framework and revived interest in the study of combinatorial invariants for immersed curves in the plane.

Titus Moves

6

8

9

10

11

12

14

15

16

18

19

20

21

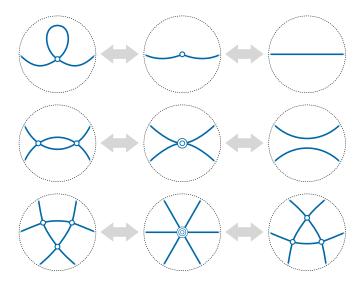
22

Any sufficiently tame homotopy can be described combinatorially by a sequence of elementary local transformations of the 4-regular plane graph representing the evolving curve. In the absence of a standard name, I will call these local transformations *Titus moves*, following Francis [9, 10].³ There are three different Titus moves:

Titus moves monogon bigon triangle compactness argument canonical regular curve

- Type I: create or destroy a **monogon** (a face with degree 1);
- Type II: create or destroy a **bigon** (a face with degree 2);
- Type III: invert a **triangle** (a face with degree 3).

During each of these moves, the curve is unchanged outside a small neighborhood of the affected face. Continuous deformations of the curve that do not change its graph representation are simply ignored.



Type I, type II, and type III Titus moves. The curve does not change outside the dotted circles.

It is not hard to see that any Titus move can be executed by a free homotopy. Moreover, any move of type II or III can be executed by a *regular* homotopy, but a type-I move cannot, because it changes the parity of the turning number. In fact, *any* free homotopy between generic curves is equivalent to a sequence of Titus moves, and *any* regular homotopy between generic curves is equivalent to a sequence of type-II and type-III Titus moves. It is possible to prove these equivalence claims using compactness arguments and careful case analysis [2,9,10,35,36].⁴ However, in light of Theorems 5.1 and 5.3, we instead give a relatively simple algorithmic proof below.

Canonical Curves

To prove that any regular homotopy can be decomposed into Titus moves, we describe an algorithm that transforms any generic curve γ into a *canonical* regular curve with

canonical regular curve!inner canonical regular curve!inner!outer

the same rotation number using Titus moves. To transform γ into a different curve δ with the same rotation number, we can run the algorithm forward to transform γ into the corresponding canonical curve, and then run the algorithm backward to transform the canonical curve into δ .

We actually define two sequences of canonical curves, the **inner** canonical curves I_r and the **outer** canonical curves O_r . When r=0, these curves coincide; $I_0=O_0$ is a figure-8 consisting of two loops, one clockwise and one counterclockwise. Otherwise, each inner curve I_r consists of an oriented cycle with |r|-1 disjoint, similarly oriented loops inside, and each outer curve O_r consists of an oriented cycle with |r|+1 disjoint oppositely oriented loops outside. Corollary 5.3 implies that each inner curve I_r has the minimum number of vertices for its rotation number.

11

12

13

14

15

16

17

18

19

20

23

24

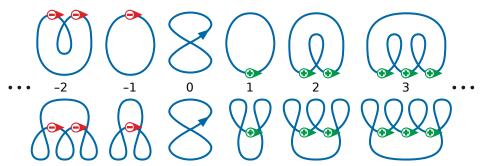
25

26

27

29

30



Inner and outer canonical curves for each rotation number.

The following algorithm appears to be folklore. Our presentation most closely follows Nowik [29], but similar algorithms were described earlier by Francis [8], Mehlhorn and Yap [20,21], and Vegter [46].

Theorem 5.5. Any generic closed curve γ with n vertices can be transformed into the outer canonical curve $O_{rot(\gamma)}$ by a sequence of $O(n^2)$ type-II and type-III Titus moves.

Proof: Let α be a simple subpath of γ that lies on the outer face and contains no crossings, and let β be the other subpath of γ between the endpoints of α . In the main part of the algorithm, we repeat the following steps until β is a simple path. For purposes of analysis, suppose the path β has m vertices when the current iteration begins.

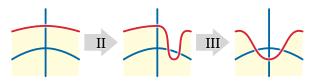
First, we find a simple subloop of β as follows. Parametrize β as a function $\beta \colon [0,1] \to \mathbb{R}^2$, and let t be the smallest number in [0,1] such that $\beta(s) = \beta(t)$ for some $0 \le s < t$. The restriction of β to the interval [s,t] is a simple loop; call it ℓ . Suppose there are i vertices in the interior of ℓ and 2b vertices on ℓ itself, *not* including the basepoint $\beta(s) = \beta(t)$. Note that $i + 2b \le m - 1$.

Now we shrink the loop ℓ so that there are no vertices in its interior. We can move any interior vertex that is adjacent to a vertex of ℓ outside ℓ using two Titus moves, as shown below. Removing all i interior vertices requires a total of 2i moves and adds 2i vertices to ℓ .

Whitney trick

One iteration of the main algorithm:

Find a simple loop, make it empty, slide it into position, and cancel opposing loops.



Moving an internal vertex outside ℓ .

At this point, the interior of ℓ intersects β in exactly 2i+2b simple, disjoint subpaths, each represented by a single edge between two vertices of ℓ . We move the interior subpaths outside ℓ one at a time, each with a type II move; removing all subpaths requires a total of i+b moves.

2

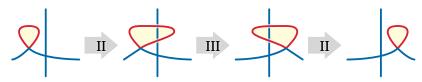
6

8

9

13

Next we slide ℓ along the curve into the subpath α . Consider the subpath π of β starting at the basepoint of ℓ and ending at some point of α . We translate ℓ past each vertex on π using three Titus moves, as shown below. Sliding ℓ up to α requires at most 3(m-1) Titus moves, one for each vertex along π .



Sliding an empty loop across a vertex.

Finally, when the loop ℓ reaches α , if there is already another loop in α on the opposite side from ℓ , we cancel these two loops using a sequence of four Titus moves often called the *Whitney trick*.⁵

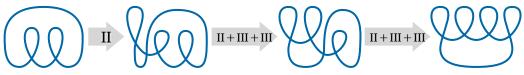


Altogether, each iteration requires at most $3i + b + 3(m - 1) + 4 \le 6m - 2$ moves. We have m = n at the beginning of the algorithm, and each iteration decreases m by at

strangeness

least 1, so the total number of moves used so far is at most $\sum_{m=1}^{n} (6m-2) = 3n^2 + n = O(n^2)$.

When the main algorithm ends, the original curve γ has been transformed into one of the two canonical curves $I_{rot(\gamma)}$ or $O_{rot(\gamma)}$. If the current curve is $O_{rot(\gamma)}$ (in particular, if $rot(\gamma) = 0$), we are done. Otherwise, we transform $I_{rot(\gamma)}$ into $O_{rot(\gamma)}$ using $3|rot(\gamma)| - 2 \le 3n + 1$ moves as follows: First we perform a type-II move to create a single outside loop; then we slide each of the $|rot(\gamma)| - 1$ inner loops to the outside using the the same sequence of three moves described earlier.



Outing an inner canonical curve.

Corollary 5.6. Two generic closed curves are regularly homotopic if and only if they are connected by a sequence of type-II and type-III Titus moves.

Strangeness

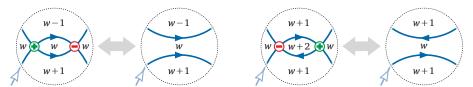
More recently, Nowik [29] proved that the $O(n^2)$ upper bound from Theorem 5.5 is tight in the worst case, using an invariant of generic curves called **strangeness**, first introduced by Arnold [3,4]. Shumakovich [39] gave the following explicit formula for strangeness, which we will take as its definition:

$$strange(\gamma) = wind^{+}(\gamma) \cdot wind^{-}(\gamma) + \sum_{\text{vertices } x \text{ of } \gamma} \operatorname{sgn}(x) \cdot wind(\gamma, x)$$

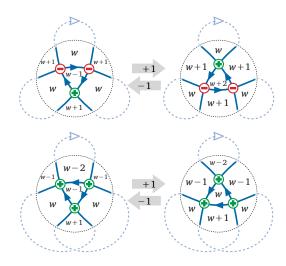
Here, sgn(x) denotes the sign of vertex x: either +1 or -1, depending on whether x is a positive or negative vertex, respectively. Recall that the winding number around any vertex x is the average of the winding numbers in the four regions incident to x.

Although the definition of strangeness depends on the choice of basepoint, its value does not. If we slide the basepoint over a positive vertex x with winding number w, then x changes to a negative vertex, so $\operatorname{sgn}(x) \cdot \operatorname{wind}(\gamma, x)$ decreases by 2w, but $\operatorname{wind}^+(\gamma)$ increases from w to w+1, so $\operatorname{wind}^+(\gamma) \cdot \operatorname{wind}^-(\gamma)$ increases by 2w.

Any Titus-II move either adds or deletes two vertices with the same winding number, one positive and one negative, and therefore leaves strangeness unchanged. Similar case analysis, illustrated below, implies that any Titus-III move either increases or decreases strangeness by exactly 1. The dashed lines in the second figure indicate only the order in which the three subpaths are traversed, not the actual topology of the overall curve. Both figures omit cases that can be obtained by reflecting or reversing the curve or moving the basepoint.



Titus-II moves do not change strangeness; white arrowheads indicate basepoints.

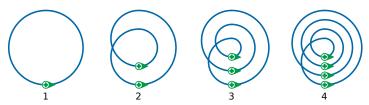


Titus-III moves either increment or decrement strangeness; white arrowheads indicate basepoints; dashed arcs indicate traversal order.

We now easily observe that each outer canonical curve O_r has strangeness 0. On the other hand, for any $r \ge 1$, let S_r denote the "spiral" curve consisting of r nested counterclockwise loops; this curve has r-1 positive vertices, no negative vertices, rotation number r, and strangeness r(r+1)/2. It follows that any sequence of Titus moves that transforms S_r into O_r (or the reversal of S_r into O_{-r}) must include at least $r(r+1)/2 = (n+1)(n+2)/2 = \Omega(n^2)$ Titus-III moves. In fact, S_{n-1} is the curve of maximum strangeness with n vertices [3,32,39,50].

4

5



Maximum-strangeness curves for each rotation number.

exercise for the reader generic arrangement of curves exercise for the reader stereographic projection even curve on the sphere odd curve on the sphere Generalizations

A similar algorithmic argument implies that for any finite set P of points, two generic curves, each with at most n crossings, that are freely homotopic in $\mathbb{R}^2 \setminus P$ are connected by a sequence of $O(n^2)$ Titus moves, each of which takes place in an open set disjoint from P. In particular, any two curves with the same winding number around a point P are connected by $O(n^2)$ Titus moves that avoid P; moreover, the $O(n^2)$ bound is tight in the worst case. We leave the details as an exercise for the reader.

We can also easily extend the definitions of free and regular homotopy to multiple curves. A *generic arrangement of curves* is a finite set of closed curves that intersect (each other or themselves) in a finite number of points, always transversely. Any 4-regular plane graph represents a generic arrangement of curves; conversely, any generic arrangement where each connected component has at least one intersection point can be represented by a 4-regular plane graph. We leave the following straightforward extension of our earlier algorithm as an exercise for the reader.

9

11

12

13

14

15

16

17

18

19

20

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

Theorem 5.7. Let $\Gamma = \{\gamma_1, \gamma_2, ..., \gamma_m\}$ and $\Delta = \{\delta_1, \delta_2, ..., \delta_m\}$ be two be two generic arrangements of curves in the plane with a total of n intersection points, where for each i, the curves γ_i and δ_i are regularly homotopic. There is a sequence of $O(n^2)$ type-II and type-III Titus moves that transforms Γ into Δ .

5.5 Regular Curves on the Sphere

Regular closed curves on the sphere $S^2 = \{(x,y,z) \mid x^2 + y^2 + z^2 = 1\} \subset R^3$ are defined just as they are in the plane: A differentiable function $\gamma \colon [0,1] \to S^2$ is a regular closed curve if and only if both γ and its derivative γ' are loops, and γ' avoids the origin. In this case, however, γ' is a loop in $\mathbb{R}^3 \setminus 0$, and it is not hard to show that any two loops in $\mathbb{R}^3 \setminus 0$ are freely homotopic. Perhaps any two regular closed curves on the sphere are regularly homotopic?

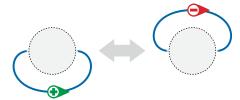
A different perspective should immediately convince you that the situation is not so straightforward. Recall the stereographic projection map $\phi: S^2 \setminus (0,0,1) \to \mathbb{R}^2$, defined as $\phi(x,y,z) = (x/(1-z),y/(1-z))$. A closed curve γ on the sphere is regular if and only if, after arbitrarily rotating the sphere so that γ avoids the north pole, the projection $\phi(\gamma)$ is a regular closed curve in the plane.

Call a regular curve on the sphere **even** if its projection to the plane has even rotation number and **odd** otherwise. We easily observe that the parity of a curve is invariant under rotations of the sphere. Equivalently, by Corollary 5.4 a generic curve on the sphere is even if and only if it has an odd number of vertices.

A regular homotopy between generic regular curves on the sphere can again be described by a sequence of Titus moves on the sphere. Stereographically projecting the resulting evolving curve onto the plane *almost* gives us a regular homotopy in the plane, except at moments where the spherical curve passes over the point of projection. Each

Whitney flip

such event can be modeled in the plane by a *flip*, which takes the topmost arc of the curve and moves it 'through infinity' to an arc below the curve, or vice versa.



A flip. Everything inside the circle is unchanged.

Each flip changes the rotation number of the planar curve by 2, by replacing a happy point with a sad point or vice versa. Any regular homotopy on the sphere can be modeled by a sequence of Titus moves and flips in the plane, and therefore two regularly homotopic curves on the sphere have the same parity. Conversely, we can increase or decrease the rotation number of any regular curve in the plane by any even number by a sequence of Titus moves and flips; thus, any two regular curves on the sphere with the same parity are regularly homotopic.

Corollary 5.8. Two regular closed curves in S^2 are regularly homotopic if and only if they are both even or both odd.

Corollary 5.9. Let $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_m\}$ be a generic arrangement of k even and m - k odd closed curves on the sphere, with a total of n intersection points. There is a sequence of $O(n^2)$ type-II and type-III Titus moves that transforms Γ into a collection of m disjoint non-nested curves, of which k are circles and m - k are figure-8s. The $O(n^2)$ bound is tight in the worst case.

Notes

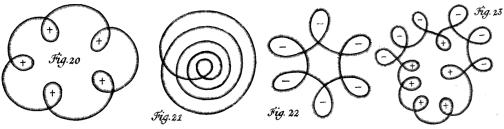
1. (page 1) Although many earlier authors studied both irregular convex polygons and regular star polygons, starting perhaps with Thomas of Bradwardine in the early 1300s [7], Meister was the first author to define polygons as sequences of *arbitrary* points connected by line segments. Meister's contributions were almost completely forgotten; many modern sources give credit for the formalization of arbitrary polygons to Poinsot [31], who was apparently unaware of Meister's then forty-year-old work (and whose definition of polygon strangely forbade repeated vertices). As Grünbaum [13] eloquently puts it, "Although Meister's paper is mentioned quite often in very complimentary ways, it seems that few—if any—of the writers even just looked at the paper." Remarkably, Grünbaum focuses entirely on Meister's definition of regular polygons, completely missing his discussion of "complications" and sums of angles, as well as the point of Poinsot's corresponding discussion of "species" of polygons.

Muphry's Law

A modern translation of Meister's work on polygons [22] and polyhedra [23] is sorely overdue. (In accordance with Muphry's Law, I fully expect some future author to notice yet another gem hidden in Meister's paper and complain that I didn't notice it.)

2. (page 10) Section V of Meister's seminal paper "De angulis figurarum" [22] discusses how the sum of the internal angles of a (not necessarily simple) polygon changes as its vertices move, where each internal angle is defined as the counterclockwise angle (in modern terms, always between 0 and 2π) between two successive edges. In particular, he observes that when an angle collapses to zero, that angle should be immediately replaced by a full circle, which implies that the sum of the angles depends on more than just the number of edges. The last paragraph of the section includes a weak, informal version of Whitney-Graustein theorem:

For brevity's sake, we distinguish between these later positive complications of the perimeter and the former, which we call negative; we do not consider the figures themselves, but their circumscribing curves. And first it is evident that a positively complicated perimeter, for any number of complications, can be reduced to the general forms in Figures 20 and 21; negatively complicated perimeters to the form in Figure 22 with the same number of complications, which differs from Figure 20 only as respective angles are external or internal. Then it is clear that positive complications remove an equal number of negative ones; that if the number of both in the figure are equal, it will return to a simple figure, where the sum of the angles is determined by the number of edges in the usual manner. I also add this: If the sine of the sum is sought from the sines of the individual angles of a figure, it would be the same for all figures with an equal number of sides, whether the perimeter is made more or less complicated.



Figures 20-23 of Meister [22].

When Meister describes how positive complications "remove" negative complications, he is referring to the effect of positive and negative loops on the sum of internal angles, not an actual cancellation like the Whitney trick. I believe the last sentence merely asserts that the sum of internal angles is always a multiple of π (so its sine is always 0).

3. (page 11) Francis [9,10] named a homotopy composed of these elementary moves a *Titus homotopy* in honor of his PhD advisor; Titus appears to have used the moves himself only later [43]. Titus moves have also been called *shadow moves* [44, 45],

perestroikas [3,4], *Reidemeister-type moves* [33], and *basic moves* [29]; however, most authors do not give bother to give them a collective name at all.

compactness argument triangle moves

- **4.** (page 11) Specifically, a compactness argument similar to our proof of Lemma ?? implies that any generic curve can be approximated by a homotopic generic polygon, and that any two homotopic generic polygons are connected by a piecewise-linear homotopy. Alexander and Briggs [2] and Reidemeister [35,36] independently proved that any piecewise-linear homotopy can be decomposed into a sequence of $triangle\ moves$, which replace a single edge pr with a pair of edges pq,qr or vice versa. Any triangle move can be decomposed into smaller triangle moves where the triangle pqr either contains one vertex, contains one self-intersection point, intersects the interior of one edge, or is disjoint from the rest of the polygon. Straightforward case analysis implies that each of these primitive triangle moves is equivalent to zero or one Titus moves. A similar proof was later given by Francis [9, 10]. Mehlhorn and Yap [20, 21] proved that any two generic polygons with the same rotation number are connected by a sequence of $O(n^2)$ triangle moves; Vegter [46] later improved this bound to O(n) using a different normal form and considerably more detailed case analysis.
- **5.** (page 13) The name "Whitney trick" seems to originate with Kauffman [17, 18, 19], who cites Whitney's seminal paper on regular curves [47]; however, this paper does not actually describe the Whitney trick! The trick was actually first used by Boy [5,6] in the construction of his minimal smooth immersion of the projective plane. Hass and Hughes [15] later used the same trick to construct an immersion of the disk with a single triple point, which they called the "kinky disk"; in hindsight, the kinky disk is just Boy's surface with a disk removed. Confusingly, the phrase "Whitney trick" more commonly refers to a generalization of a type-II Titus move to smooth immersions of higher-dimensional manifolds, proposed by Whitney [48].

Bibliography

- James W. Alexander. Topological invariants of knots and links. *Trans. Amer. Math. Soc.* 30(2):275–306, 1928. (4)
 - [2] James W. Alexander and G. B. Briggs. On types of knotted curves. *Ann. Math.* 28(1/4):562–586, 1926–1927. (10, 11, 19)
- Vladimir I. Arnold. Plane curves, their invariants, perestroikas and classifications. Singularities and Bifurcations, 33–91, 1994. Adv. Soviet Math. 21, Amer. Math. Soc. (1, 10, 14, 15, 19)
 - [4] Vladimir I. Arnold. *Topological Invariants of Plane Curves and Caustics*. University Lecture Series 5. Amer. Math. Soc., 1994. (1, 10, 14, 19)

35	[5]	Werner Boy. <i>Über die Curvatura integra und die Topologie geschlossener Flächen</i> . Ph.D. thesis, Univ. Göttingen, 1901. Cited in [6]. (10, 19)	1
	[6]	Werner Boy. Über die Curvatura integra und die Topologie geschlossener Flächen. <i>Math. Ann.</i> 57(2):151–184, 1903. (1, 10, 19, 20)	2
	[7]	Thomas Bradwardine. <i>Geometria speculativa</i> . c. 1325. English translation and commentary in [25]. (17)	4
	[8]	George K. Francis. The folded ribbon theorem: A contribution to the study of immersed circles. <i>Trans. Amer. Math. Soc.</i> 141:271–303, 1969. (1, 12)	6
	[9]	George K. Francis. Titus' homotopies of normal curves. <i>Proc. Amer. Math. Soc.</i> 30:511–518, 1971. (10, 11, 18, 19)	8
	[10]	George K. Francis. Generic homotopies of immersions. <i>Indiana Univ Math. J.</i> 21:1101–1112, 1971/72. (11, 18, 19)	10 11
	[11]	Carl Friedrich Gauß. Nachlass. I. Zur Geometria situs. <i>Werke</i> , vol. 8, 271–281, 1900. Teubner. Originally written between 1823 and 1840. (1, 6, 7, 8, 9)	12 13
	[12]	Eitan Grinspun and Adrian Secord. Introduction to discrete differential geometry: The geometry of plane curves. <i>Discrete Differential Geometry: An Applied Introduction</i> , 3rd edition, 1–4, 2008. SIGGRAPH Asia 2008 Course Notes. (1)	14 15 16
	[13]	Branko Grünbaum. Polygons: Meister was right and Poinsot was wrong but prevailed. <i>Beitr. Algebra Geom.</i> 53(1):57–71, 2012. (17)	17 18
	[14]	Branko Grünbaum and Geoffrey C. Shephard. Rotation and winding numbers for planar polygons and curves. <i>Trans. Amer. Math. Soc.</i> 322(1):169–187, 1990. (1)	19 20
	[15]	Joel Hass and John Hughes. Immersions of surfaces in 3-manifolds. <i>Topology</i> 24(1):97–112, 1985. (19)	21 22
	[16]	Heinz Hopf. Über die Drehung der Tangenten und Sehnen ebener Kurven. <i>Compositio Math.</i> 2:50–62, 1935. (1, 3, 5)	23 24
	[17]	Louis H. Kauffman. <i>Formal Knot Theory</i> . Princeton Math. Notes. Princeton Univ. Press, 1983. (1, 6, 19)	25 26
	[18]	Louis H. Kauffman. <i>On Knots</i> . Princeton Univ. Press, 1987. (1, 19)	27
	[19]	Louis H. Kauffman. Knots and Physics. World Scientific, 2001. (19)	28
	[20]	Kurt Mehlhorn and Chee-Keng Yap. Constructive Hopf's theorem: Or how to untangle closed planar curves. <i>Proc. Int. Conf. Automata Lang. Progr.</i> , 410–423, 1988. Lecture Notes Comput. Sci. 317, Springer-Verlag. Preliminary version of [21]. (12, 19, 21)	29 30 31 32

- [21] Kurt Mehlhorn and Chee-Keng Yap. Constructive Whitney-Graustein theorem: or how to untangle closed planar curves. *SIAM J. Comput.* 20(4):603–621, 1991. Full version of [20]. (1, 9, 12, 19, 20)
- Albrecht Ludwig Friedrich Meister. Generalia de genesi figurarum planarum, et inde pendentibus earum affectionibus. *Novi Commentarii Soc. Reg. Scient. Gott.* 1:144–180 + 9 plates, 1769/1770. Presented January 6, 1770. (1, 4, 6, 10, 18)

1

2

6

14

16

18

- [23] Albrecht Ludwig Friedrich Meister. Commentatio de solidis geometricis pro cognoscenda eorum indole in certos ordines et versus disponendis. *Commentationes Mathematicae (Soc. Reg. Scient. Gott.)* 7:3–74 + 2 plates, 1784/1785. Presented October 1, 1875. (18)
- [24] August F. Möbius. Theorie der elementaren Verwandtschaften. Ber. Sächs. Akad. Wiss. Leipzig, Math.-Phys. Kl. 17:18–57, 1863. Gesammelte Werke 2:433–471, Liepzig, 1886. (4)
- [25] George Molland. Thomas Bradwardine, Geometria Speculativa: Latin Text and English Translation with an Introduction and Commentary. Boethius 18. Steiner Verlag, Stuttgart, 1989. (20)
- [26] Marston Morse. Topological Methods in the Theory of Functions of a Complex Variable. Princeton Univ. Press, 1947. (1)
- [27] Marston Morse and Maurice Heins. Deformation classes of meromorphic functions and their extensions to interior transformations. *Acta Mathematica* 79(1):53–103, 1947. (1)
 - [28] Tristan Needham. Visual Complex Analysis. Oxford University Press, 1999. (1)
- [29] Tahl Nowik. Complexity of planar and spherical curves. *Duke J. Math.* 148(1):107–118, 2009. (12, 14, 19)
- [30] Tahl Nowik. Order one invariants of planar curves. *Adv. Math.* 220:427–440, 2009. ArXiv:0712.4071. (2)
- [31] Louis Poinsot. Mémoire sur les polygones et les polyèdres. *J. École Polytechnique* 10:16–48, 1810. (1, 17)
- [32] Michael Polyak. Invariants of curves and fronts via Gauss diagrams. *Topology* 37(5):989–1009, 1998. (15)
- [33] Michael Polyak. New Whitney-type formulae for plane curves. *Differential and symplectic topology of knots and curves*, 103–111, 1999. Amer. Math. Soc. Translations 190, Amer. Math. Soc. (1, 19)
- John R. Quine. Homotopies and intersection sequences. *Pacific J. Math.* 64(1):233–240, 1976. (8)

35	[35]	Kurt Reidemeister. Elementare Begründung der Knotentheorie. <i>Abh. Math. Sem. Hamburg</i> 5:24–32, 1927. (10, 11, 19)	1
	[36]	Kurt Reidemeister. <i>Knotentheorie</i> . Ergebnisse der Mathematik und ihrer Grenzgebiete 1. Springer, 1932. (10, 11, 19)	2
	[37]	Bruce L. Reinhart. The winding number on two manifolds. <i>Annales de l'institute Fourier</i> 10:271–283, 1960. (1)	4 5
	[38]	Herbert Seifert. Konstruktion dreidimensionaler geschlossener Raume. <i>Berichte Sächs. Akad. Leipzig</i> 83:26–66, 1931. (4)	6 7
	[39]	Aleksandr Shumakovich. Explicit formulas for the strangeness of plane curves. <i>Algebra i Analiz</i> 7(3):165–199, 1995. Corrections in <i>Algebra i Analiz</i> 7(5):252–254, 1995. In Russian. (14, 15, 22)	8 9 10
	[40]	Aleksandr Shumakovich. Explicit formulas for the strangeness of plane curves. <i>St. Petersburg Math. J.</i> 7(3):445–472, 1996. English translation of [39]. (22)	11 12
	[41]	Charles J. Titus. A theory of normal curves and some applications. <i>Pacific J. Math.</i> 10:1083–1096, 1960. (8)	13 14
	[42]	Charles J. Titus. The combinatorial topology of analytic functions on the boundary of a disk. <i>Acta Mathematica</i> 106:45–64, 1961. (1)	15 16
	[43]	Charles J. Titus. Extensions through codimension one to sense preserving mappings. <i>Ann. Inst. Fourier, Grenoble</i> 23(2):215–227, 1973. (10, 18)	17 18
	[44]	Vladimir G. Turaev. Quantum invariants of 3-manifolds and a glimpse of shadow topology. <i>C. R. Acad. Sci. Paris I</i> 313:395–398, 1991. (18)	19 20
	[45]	Vladimir G. Turaev. Shadow links and face models of statistical mechanics. <i>J. Diff. Geom.</i> 36:35–74, 1992. (18)	21 22
	[46]	Gert Vegter. Kink-free deformation of polygons. <i>Proc. 5th Ann. Symp. Comput. Geom.</i> , 61–68, 1989. (1, 12, 19)	23 24
	[47]	Hassler Whitney. On regular closed curves in the plane. <i>Compositio Math.</i> 4:276–284, 1937. (2, 5, 8, 9, 10, 19)	25 26
	[48]	Hassler Whitney. The self-intersections of a smooth n -manifold in $2n$ -space. Ann . $Math. 45(2):220-246, 1944. (19)$	27 28
	[49]	Christian Wiener. Über Vielecke und Vielflache. Teubner, 1864. (1)	29
	[50]	Jianming Yu, Jianyi Zhou, and Jianzhong Pan. Extremal values of the basic invariants of plane curves. <i>J. Knot Theory Ramif.</i> 9(8):1085–1126, 2000. (15)	30 31