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)

A cube of cheese no larger than a die
May bait the trap to catch a nibbling mie.
— attributed to Chauncey Depew
by Ambrose Bierce, *The Devil's Dictionary* (1911)

4 Regular Homotopy and Hexahedral Meshing

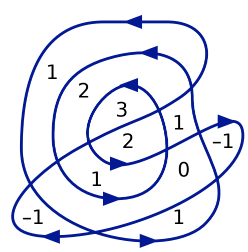
This lecture is concerned with closed curves in the plane that are smooth, but not necessarily. Intuitively, a *regular closed curve* is a closed curve with no sharp corners. Two regular closed curves are *regularly homotopic* if one can be continuously deformed into the other without introducing any sharp corners at any time. The *turning number* of a regular closed curve is the number of times its tangent vector rotates counterclockwise during a single traversal of the curve. I will prove the *Whitney-Graustein theorem*: two regular closed curves in the plane are regularly homotopic if and only if they have the same turning number. Then I'll describe an application of this theorem to hexahedral meshing.

4.1 Winding Numbers

But first, a warmup exercise.

Recall that a *loop* in the plane is a continuous function $\alpha\colon [0,1] \to \mathbb{R}^2$ such that $\alpha(0) = \alpha(1)$. Let p be an arbitrary point that is not in the image of α, and consider an infinite ray r based at p. We say that r is ***generic*** if the set $\{t \mid \alpha(t) \in r\}$ is finite and excludes the values 0 and 1. An intersection point $\alpha(t) \in r$ is called a ***crossing*** if the points $\alpha(t-\varepsilon)$ and $\alpha(t+\varepsilon)$ lie on opposite sides of r, for all sufficiently small $\varepsilon > 0$. The crossing is ***positive*** if the triangle $(0, \alpha(t), \alpha(t+\varepsilon))$ is oriented counterclockwise, and ***negative*** otherwise. The ***winding number*** of α around p, denoted $wind(\alpha, p)$, is the number of positive crossings minus the number of negative crossings, for any (generic) ray.

An argument similar to the proof of Lemma ≥ 2 (the easy half of the Jordan Curve Theorem) implies that this definition is independent of the choice of ray r. If we continuously move the ray, crossings can appear and disappear, but they always do so in matched pairs: one positive and one negative. Moreover, two points in the same connected component of $\mathbb{R}^2 \setminus \operatorname{im}\alpha$ define the same winding number; in particular, if p is in the unbounded component of $\mathbb{R}^2 \setminus \operatorname{im}\alpha$, then $wind(\alpha, p) = 0$. If α is *simple*, the winding number with respect to every interior point is either 1 or -1.



Winding numbers.

Essentially the same argument implies that if we continuously deform a loop without touching a fixed point, say the origin 0, the winding number around that point is constant during the entire deformation, even if we allow the basepoint to move during the deformation. The type of deformation we allow is called a ***free homotopy***, which is just a homotopy through loops. More formally, a ***free homotopy*** between two loops α and β in $\mathbb{R}^2 \setminus 0$ is a function $h\colon [0,1]^2 \to \mathbb{R}^2 \setminus 0$ such that $h(0,t) = \alpha(t)$ and $h(1,t) = \beta(t)$ for all t, and $h(s,0) = h(s,1)$ for all s.

In fact, winding numbers exactly characterize free homotopy classes of loops in the punctured plane.

Theorem 4.1. *Two loops α and β are freely homotopic in $\mathbb{R}^2 \setminus 0$ if and only if $wind(\alpha,0) = wind(\beta,0)$.*

Proof: Let $\zeta\colon [0,1] \to S^1$ denote the loop $\zeta(t) = (\cos 2\pi t, \sin 2\pi t)$. For any integer k, let $\zeta^k(t) = \zeta(kt) = \zeta(kt \bmod 1)$; the loop ζ^k clearly has winding number k. We have already argued that any two freely homotopic loops in $\mathbb{R}^2 \setminus 0$ have the same winding number. To complete the proof of the theorem, we show that any loop α in $\mathbb{R}^2 \setminus 0$ is freely homotopic to $\zeta^{wind(\alpha)}$, which immediately implies that loops with the same winding number are homotopic.

Now let α be an arbitrary loop in $\mathbb{R}^2 \setminus 0$. Without loss of generality, we can assume that $\alpha(0)$ lies on the positive x-axis; otherwise, rotate α using a free homotopy. We can also safely assume that $\alpha(t)$ lies on the positive x-axis for only a finite set of real values t.

Let $\alpha^*\colon [0,1] \to S^1$ denote the function $\alpha^*(t) = \alpha(t)/\|\alpha(t)\|$. The loops α and α^* are freely homotopic; in particular, they have the same winding number. There is a unique function $\bar{\alpha}\colon [0,1] \to \mathbb{R}$ such that $\alpha^* = \zeta \circ \bar{\alpha}$. Because $\alpha^*(0) = \alpha^*(1)$, the function value $\bar{\alpha}(1)$ must be an integer.

For all $0 < t < 1$, the point $\alpha(t)$ lies on the positive x-axis if and only if $\bar{\alpha}(t)$ is an integer. If moreover $\bar{\alpha}(t-\varepsilon) < \bar{\alpha}(t) < \bar{\alpha}(t+\varepsilon)$ for all sufficiently small $\varepsilon > 0$, then $\alpha(t)$ is a positive crossing. Similarly, if $\bar{\alpha}(t-\varepsilon) > \bar{\alpha}(t) > \bar{\alpha}(t+\varepsilon)$ for all sufficiently small $\varepsilon > 0$, then $\alpha(t)$ is a negative crossing. It follows that $\bar{\alpha}(1) = wind(\alpha)$.[1]

Finally, define the function $\bar{h}\colon [0,1]^2 \to \mathbb{R}$ by setting $\bar{h}(s,t) = (1-s)\bar{\alpha}(t) + st \cdot wind(\alpha)$, and let $h = \zeta \circ \bar{h}$. Tedious definition-crunching implies that $h\colon [0,1]^2 \to S^1$ is a homotopy from α^* to $\zeta^{wind(\alpha)}$. Thus, α is freely homotopic to $\zeta^{wind(\alpha)}$, which completes the proof. □

4.2 Regular Closed Curves and Regular Homotopy

More formally, a ***parameterized regular closed curve*** is a path $\gamma\colon [0,1] \to \mathbb{R}^2$ satisfying the following conditions:

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

More succinctly, a differentiable function $\gamma\colon [0,1] \to \mathbb{R}^2$ is a parameterized regular closed curve if and only if both γ and its derivative γ' are loops and γ' avoids the origin.

Two regular closed curves γ and δ are ***equivalent***, denoted $\boldsymbol{\gamma \sim \delta}$, if they differ only by reparameterization, that is, if there is a continuous function $\eta\colon \mathbb{R} \to \mathbb{R}$ such that $\eta(t+1) = \eta(t) + 1$ and $\gamma(t) = \delta(\eta(t) \bmod 1)$ for all t. If such a function η exists, its derivative must be positive everywhere. It is easy to check that $\sim$ is an equivalence relation; the equivalence classes are called ***regular closed curves***, and its elements are called ***parameterizations***.

A ***regular homotopy*** is a homotopy through parameterized regular closed curves, that is, a function $h\colon [0,1]^2 \to \mathbb{R}^2$ such that for all s, the function $t \mapsto h(s,t)$ is a parameterized regular closed curve.

[1]In fact, $\bar{\alpha}(1)$ is a much less fragile *definition* of the winding number of α, since it does not assume that α is in any way well-behaved.

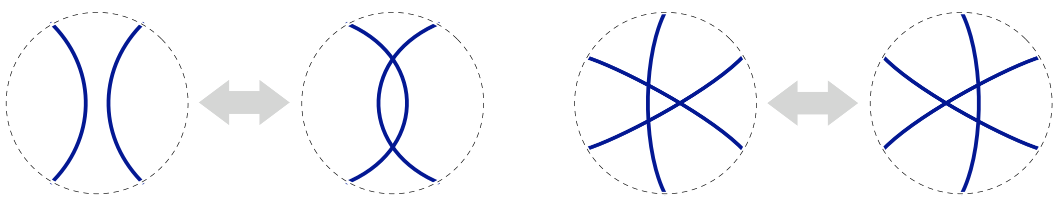
Two parameterized regular closed curves γ and δ are ***regularly homotopic***, denoted $\gamma \simeq \delta$ if there is a regular homotopy h such that $h(0,t) = \gamma(t)$ and $h(1,t) = \delta(t)$ for all t.

Lemma 4.2. *Let γ and δ be arbitrary parameterized regular closed curves. If γ and δ are equivalent, then γ and δ are regularly homotopic.*

Proof: Suppose $\gamma(t) = \delta(\eta(t) \bmod 1)$. Let $H\colon [0,1] \times \mathbb{R} \to \mathbb{R}$ be defined by $H(s,t) := (1-s)\eta(t) + st$, and let $h\colon [0,1]^2 \to \mathbb{R}^2$ be defined by $h(s,t) := \delta(H(s,t) \bmod 1)$. We easily verify that $h(0,t) = \gamma(t)$ and $h(1,t) = \delta(t)$ for all t, and that $H(s,t+1) = H(s,t) + 1$ for all s and t. Moreover, $\frac{\partial}{\partial t}H(s,t) = (1-s)\eta'(t) + s > 0$ for all s and t. We conclude that h is a regular homotopy from γ to δ. $\square$

In light of this lemma, we say that two regular closed curves are regularly homotopic if any parameterizations of those curves are regularly homotopic.

A regular closed curve α is ***normal*** if it has a finite number of self-intersections, and each self-intersection is a pairwise crossing. Any normal curve can be described by a planar embedding of a planar (multi-)graph where every vertex has degree 4. (However, not all connected 4-regular plane graphs describe normal curves.) A regular homotopy between two normal curves can be described combinatorially as a sequence of so-called ***Whitney moves***, of which there are two types: creating or destroying bigons, and flipping triangles.

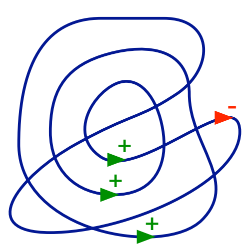


Whitney moves; the curve does not change outside the circles

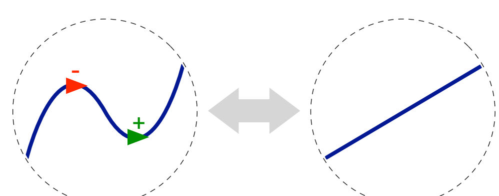
4.3 Turning Numbers and the Whitney-Graustein Theorem

In 1937, Whitney characterized the regular homotopy classes of regular curves in the plane in terms of ***turning numbers***. The turning number $turn(\gamma)$ of a regular closed curve γ is the winding number of its derivative γ' around the origin.

Equivalently, call a point $\gamma(t)$ *extreme* if the derivative vector $\gamma'(t)$ points in some fixed direction (say, to the right). For a generic direction, there are a finite number of extreme points of exactly two types: $\gamma(t)$ is ***happy*** if γ is locally to the left of the tangent ray, and ***sad*** if γ is locally to the right of the tangent ray. If the fixed direction is to the right, happy points have neighborhoods that curve up $\smile$ and sad points have neighborhoods that curve down $\frown$. The turning number of γ is the number of happy points minus the number of sad points.



A regular closed curve with turning number 2.



Creating or destroying a happy-sad pair.

It is not hard to see that regularly homotopic curves have equal turning numbers; indeed, there is a one-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 and σ are regularly homotopic, their derivatives are homotopic in $\mathbb{R}^2 \setminus 0$, so γ and δ have the same turning number. Alternately, one can argue that a regular homotopy can only create or destroy happy and sad points in matching pairs.

Surprisingly, the converse is true as well! The following result appears in a seminal 1937 paper of Hassler Whitney [7], who attributes both the theorem and its proof to William Graustein.

Theorem 4.3 (Whitney-Graustein). *Two regular closed curves in $\mathbb{R}^2$ are regularly homotopic if and only if their turning numbers are equal.*

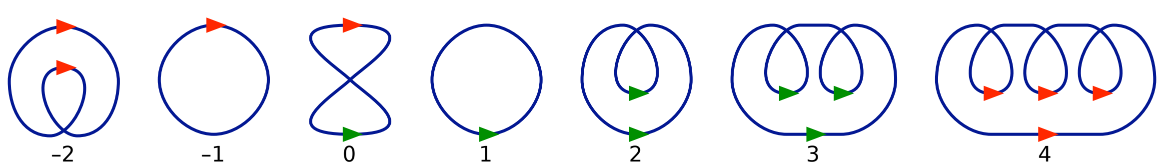
Proof: Let γ and δ be parameterized regular closed curves with the same turning number. Without loss of generality, we assume that both γ and δ have arc-length 1. (We can scale each curve using a regular homotopy if necessary.)

We also assume without loss of generality that γ and δ are parameterized by arc length, that is, $\|\gamma'(t)\| = \|\delta'(t)\| = 1$ for all t. Let $\ell_\gamma(t) = \int_0^t \|\gamma'(u)\|\,du$ denote the length of the prefix $\gamma([0,t])$. The function $\ell_\gamma\colon [0,1] \to [0,1]$ is a continuous and strictly increasing bijection, and therefore has a continuous and increasing inverse ℓ_γ^{-1}. Let $\bar{\gamma}(t) := \gamma(\ell_\gamma^{-1}(t))$; by construction, $\|\bar{\gamma}'(t)\| = 1$ for all t. Because ℓ^{-1} is continuously increasing, $\bar{\gamma}$ is equivalent to (and therefore regularly homotopic to) γ.

Because δ and γ have the same turning number, their derivatives γ' and δ' have the same winding number and are therefore homotopic in 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 $\mathit{turn}(\gamma) = 0$.

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: $\int_0^1 \alpha(t)\,dt = 0$. Let $c\colon [0,1] \to \mathbb{R}^2$ be defined by $c(s) := \int_0^1 h'(s,t)\,dt$; this 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 $h^*\colon [0,1]^2 \to \mathbb{R}^2$ be 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. $\square$

For *normal* curves, one can also prove this theorem combinatorially, by describing a sequence of Whitney moves transforming any normal curve to a canonical normal curve with the same winding number. One possible choice of canonical curves is shown below, but there are other possibilities. I'll leave the details of this approach as an exercise. Very recently, Nowik [5] proved that $\Theta(n^2)$ Whitney moves are always necessary and sometimes sufficient to transform one normal curve into another, where n is the total number of self-intersection points in both curves.



Canonical regular curves for each turning number.

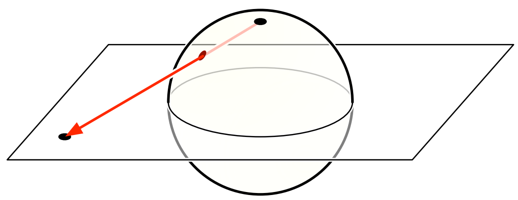
4.4 Regular Curves on the Sphere

The Whitney-Graustein theorem almost immediately implies a similar classification of regular closed curves on the sphere $S^2 = \{(x,y,z) \mid x^2 + y^2 + z^2 = 1\} \subset R^3$, which 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 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 things are not so simple. Consider the ***stereographic projection*** $\phi : S^2 \setminus (0,0,1) \to \mathbb{R}^2$, defined by setting

$$\phi(x,y,z) := \left(\frac{x}{1-z}, \frac{y}{1-z}\right).$$

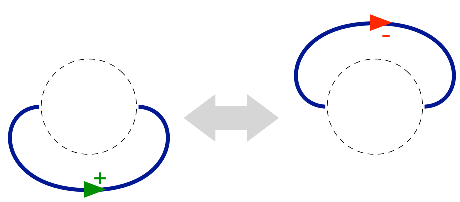
The projection of any point $p \in S^2$ can be determined geometrically by extending a line through p and the north pole $(0,0,1)$; the intersection of this line with the xy-plane is $\phi(p)$. A closed curve γ on the sphere is regular if and only if, after rotating the sphere so that γ avoid the north pole, the projection $\phi(\gamma)$ is a regular closed curve in the plane.



Stereographic projection.

Say that a regular curve on the sphere is ***even*** if its projection to the plane has even turning number, and ***odd*** otherwise. We easily observe that the parity of a curve is invariant under rotations of the sphere.

A regular homotopy between normal curves on the sphere can again be modeled by a sequence of Whitney moves on the sphere. 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 north pole. This event can be modeled in the plane by a ***Whitney flip***, which takes the topmost arc of the curve and moves it 'through infinity' to an arc below the curve, or vice versa.



A Whitney flip. Everything *inside* the circle is unchanged.

A Whitney flip changes the turning number of the planar curve by 2, by replacing a happy point with a sad point or vice versa. Any regular homotopy can be modeled by a sequence of Whitney moves and Whitney flips, two regularly homotopic curves on the sphere have the same parity. Conversely, we can change the turning number of any regular curve in the plane by any even number by a sequence of Whitney moves and Whitney flips, so any two regular curves on the sphere with the same parity are regularly homotopic.

Corollary 4.4. *Two regular closed curves in S^2 are regularly homotopic if and only if they are both even or both odd.*

We can also extend the notion of regular homotopy to *sets* of regular closed curves.

Corollary 4.5. *Let C be a finite set of n regular closed curves on the sphere, of which k are even. Then C is regularly homotopic to a set of n disjoint non-nested curves, of which k are small figure-8s and the remaining $n - k$ are small circles.*

Nowik's $\Theta(n^2)$ bound on the worst-case number of Whitney moves also applies to sets of normal curves on the sphere [5].

4.5 The Mitchell-Thurston Hex Mesh Theorem

Many applications in scientific computing call for three-dimensional to be decomposed into a mesh of geometrically simpler pieces. One of the most sought-after types of decomposition is a ***hexahedral mesh***, or simply *hex mesh*. A hex(ahedron) is a polyhedron combinatorially equivalent to a cube: six facets, each with four edges, meeting at eight vertices of degree 3. A hex mesh is a set of hexahedra in which the intersection of any two hexes is either the empty set, a vertex of both, an edge of both, or a facet of both.

Now suppose we are given a quadrilateral mesh of the *surface* of some three-dimensional object. Not surprisingly, a quad mesh is a set of quadrilaterals, any two of which share a common edge, a common vertex, or nothing at all; the union of the quads is the surface in question. When can we extend this surface quad mesh to a hex mesh of the interior body? Specifically, we want a hex mesh whose boundary facets are precisely the quadrilaterals in the input surface mesh.

As a geometry problem, this is still open, even for some very simple special cases. (See in particular Bern and Eppstein's *bicuboid* [1].) As a topological question, however, there is a very simple solution, at least for surfaces homeomorphic to the sphere, independently discovered by Bill Thurston [6] and Scott Mitchell [4]. In this setting, we don't require the quads and hexes to have any particular geometry—they're just balls and disks with certain constraints on their intersection pattern.

Let X be a topological space. A ***cube*** in X is a continuous injective map $Q\colon [0,1]^d \to X$. A ***facet*** of Q is another cube $Q'\colon [0,1]^{d-1} \to X$ that is equal to the restriction of Q to one of the $2d$ facets of the reference hypercube $[0,1]^d$ (ignoring the fixed coordinate). A ***face*** of Q is either Q itself or a face of a facet of Q. A ***(geometric) cube complex*** in X is a set $\mathcal{Q}$ of cubes in X that satisfy the following properties:

- If $Q \in \mathcal{Q}$, then every facet of Q is also in $\mathcal{Q}$.
- For any pair of cubes $Q, Q' \in \mathcal{Q}$ whose images in X intersect, there is another cube Q'' that is a face of both Q and Q, such that $\operatorname{im} Q'' = \operatorname{im} Q \cap \operatorname{im} Q'$.

The ***underlying space*** of a cube complex is the union of the images of its cubes.

When we speak of 'a quad mesh of the sphere', we really mean a finite cube complex in $\mathbb{R}^3$ whose underlying space is the standard unit sphere S^2; Similarly, 'a hex mesh of the ball' means a finite cube complex in $\mathbb{R}^3$ whose underlying space is the standard unit ball B^3. (Obviously, we can replace B^3 and S^2 with any subset X of $\mathbb{R}^3$ homeomorphic to B^3 and its boundary, but imposing our own geometry makes the following argument simpler.)

Theorem 4.6 (Mitchell-Thurston). *A quad mesh Q of the sphere can be extended to a hex mesh of the ball if and only if the number of quads in Q is even.*

Proof: One direction follows from an easy inductive argument: The empty complex obviously has zero boundary quads. Removing any cube from any other cube complex destroys some number $k < 6$ of boundary quads, but creates $6 - k$ more, so the total change in the number of boundary quads is $6 - 2k$, which is even. Thus, every (pure three-dimensional) cube complex has an even number of boundary quads.

On the other hand, let Q be an even quad mesh of the sphere. The combinatorial dual Q^* of Q is a connected 4-regular plane graph, which we can interpret as a set of regular closed curves on the sphere. By Corollary 4.5, there is a regular homotopy of those curves to a collection of disjoint, non-nested circles and figure-8s. Because Whitney moves preserve the parity of the number of intersection points, there are an even number of figure-8s.

Now think of the regular homotopy as the intersection of a shrinking sphere with a collection of *surfaces*. Edges sweep out sections of surfaces, and vertices sweep out curves where two such surfaces cross. A Whitney-II move occurs at every local minimum or maximum of the radius function along some intersection curve. A Whitney-III move occurs whenever three surfaces meet at a common point. Within the innermost sphere, cap off the small circles and join the figure-8s in pairs with self-intersecting tubes.

We now have a collection of surfaces inside the unit ball that *locally* resembles the dual of a hex mesh. These surfaces subdivide the ball into a cell complex, whose vertices are intersection points of three surfaces (one of which may be the bounding sphere), whose edges are segments of intersection curves bounded by vertices, and whose sheets are maximal subsets of surfaces bounded by edges. Locally, the cell complex resembles the dual of a hex mesh; the neighborhood of each interior vertex is homeomorphic to the intersection of three planes.

However, in general, the cell complex contains structures inconsistent with a dual hex mesh: edges incident to zero or one vertices, edges incident only to boundary vertices, sheets incident to the same edge more than once, and sheets incident only to boundary edges. To eliminate these features, we add several spheres to the surface arrangement:

- one just inside the bounding sphere;
- one around each interior vertex;
- one around the interior of each edge, intersecting the sphere(s) around its endpoint(s), but not containing the endpoint(s) themselves;
- two surrounding each edge without vertices;
- one around each sheet, intersecting the spheres around its edges and vertices, but not containing the edges and vertices themselves;

Adding these spheres increases the complexity of the surface arrangement by a (not particularly small) constant factor. The resulting cell complex is dual to a hex mesh of the ball, whose boundary facets are the original surface quad mesh. □

The hex mesh that results from this procedure is considerably more complex than the given quad mesh. If the input quad mesh has complexity n, then Nowik's analysis [5] implies that we may need $\Theta(n^2)$ Whitney moves to reach a collection of circles and eights. Thus, this algorithm results in a hex mesh of complexity $\Theta(n^2)$ in the worst case. A more recent algorithm of Eppstein [3] computes a topological hex mesh with complexity $O(n)$.

References

[1] M. Bern and D. Eppstein. Flipping cubical meshes. *Proc. 10th Int. Meshing Roundtable*, 19–29, 2001. ⟨http://www.andrew.cmu.edu/user/sowen/abstracts/Be812.html⟩. Preliminary version of [2].

[2] M. Bern, D. Eppstein, and J. Erickson. Flipping cubical meshes. *Eng. Comput.* 18:173–187, 2002. Full version of [1].

[3] D. Eppstein. Linear-complexity hexahedral mesh generation. *Comput. Geom. Theory Appl.* 12:3–16, 1999.

[4] S. A. Mitchell. A characterization of the quadrilateral meshes of a surface which admit a compatible hexahedral mesh of the enclosed volume. *Proc. 13th Ann. Symp. Theoret. Aspects Comput. Sci.*, 456–476, 1996. Lecture Notes Comput. Sci. 1046, Springer-Verlag.

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

[6] W. Thurston. Hexahedral decomposition of polyhedra. Posting to sci.math, 25 October 1993. ⟨http://www.ics.uci.edu/~eppstein/gina/Thurston-hexahedra.html⟩. Absent from the Google Groups sci.math archive.

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