

# Stiefel Manifolds and Polygons

Clayton Shonkwiler

Department of Mathematics, Colorado State University; clay@shonkwiler.org

## Abstract

Polygons are compound geometric objects, but when trying to understand the expected behavior of a large collection of random polygons – or even to formalize what a random polygon is – it is convenient to interpret each polygon as a point in some parameter space, essentially trading the complexity of the object for the complexity of the space. In this paper I describe such an interpretation where the parameter space is an abstract but very nice space called a *Stiefel manifold* and show how to exploit the geometry of the Stiefel manifold both to generate random polygons and to morph one polygon into another.

## Introduction

The goal is to describe an identification of polygons with points in certain nice parameter spaces. For example, every triangle in the plane can be identified with a pair of orthonormal vectors in 3-dimensional space  $\mathbb{R}^3$ , and hence planar triangle shapes correspond to points in the space  $\text{St}_2(\mathbb{R}^3)$  of all such pairs, which is an example of a Stiefel manifold. While this space is defined somewhat abstractly and is hard to visualize – for example, it cannot be embedded in Euclidean space of fewer than 5 dimensions [9] – it is easy to visualize and manipulate points in it: they're just pairs of perpendicular unit vectors. Moreover, this is a familiar and well-understood space in the setting of differential geometry. For example, the group  $SO(3)$  of rotations of 3-space acts transitively on it, meaning that there is a rotation of space which transforms any triangle into any other triangle.

This identification between triangles and pairs of orthonormal vectors generalizes: we will represent an  $n$ -gon in the plane by a pair of orthonormal vectors in  $\mathbb{R}^n$ . The space  $\text{St}_2(\mathbb{R}^n)$  of such pairs is again a Stiefel manifold – in general, the *Stiefel manifold*  $\text{St}_k(\mathbb{R}^n)$  is the collection of all  $k$ -tuples of mutually orthonormal vectors in  $\mathbb{R}^n$  – and again there is a natural transitive action, now by the group  $SO(n)$  of all rotations of  $\mathbb{R}^n$ . The existence of transitive group actions is special in both probability theory and differential geometry, and in particular means that the Stiefel manifold (and, hence, the collection of planar  $n$ -gons) has a natural probability measure (*Haar measure*) which is invariant under the action.

There are two pleasing consequences of this identification of polygons with points in a Stiefel manifold:

1. It is easy to sample (pseudo-)random points from Haar measure on the Stiefel manifold  $\text{St}_2(\mathbb{R}^n)$ : apply Gram–Schmidt to a pair of  $n$ -dimensional vectors with entries chosen from a standard Gaussian distribution. This produces a simple algorithm for stochastically generating large ensembles of random polygons. The resulting polygons have edges whose lengths are on the order of  $\frac{1}{\sqrt{n}}$  times the diameter of the polygon, which is qualitatively different from polygons produced from, e.g., random Gaussian vertices, which have edge lengths of the same order as the diameter.
2. It is straightforward to find explicit parametrized paths between points in the Stiefel manifold, yielding a procedure for generating efficient interpolations between any two polygons with the same number of edges.

Therefore, the Stiefel manifold is potentially useful both as a generative tool and for finding aesthetically-pleasing morphs between pre-determined shapes.

The *Mathematica* notebook which generated all of the figures in this paper includes reference implementations of all of the algorithms described below and is available at <https://shonkwiler.org/bridges2019>.<sup>1</sup> It may be helpful to follow along in that notebook to see the implementation details.

## From Polygons to Stiefel Manifolds

The basic construction is surprisingly straightforward; see [3] for a more detailed but still elementary introduction to these ideas. The goal is a translation-invariant and scale-invariant representation of polygons, so represent an  $n$ -gon as a sequence of edge vectors  $e_1, \dots, e_n$  which, up to scale, we normalize to have  $|e_1| + \dots + |e_n| = 2$ . This representation is really of an *ordered* polygon, since it implies a choice of which edge is first, second, and so on. Notice that a cyclic permutation of the edge labels does not change the visual appearance of the polygon at all.

In this vector representation, the condition that the polygon closes is equivalent to the vector sum  $e_1 + \dots + e_n$  being the zero vector. Generating random tuples of vectors satisfying  $\sum e_k = 0$  and  $\sum |e_k| = 2$  is not easy, and finding nice paths in the collection of such tuples is even harder, but a clever trick originally due to Hausmann and Knutson [6] (cf. [2]) comes to the rescue.

Interpret the edge vectors  $e_k$  as complex numbers, and then define the complex numbers  $z_k$  by  $z_k^2 = e_k$ . There is ambiguity in the choice of each  $z_k$ , but the payoff is worth it: letting each  $z_k = x_k + \mathbf{i}y_k$  and thinking of  $\vec{x} = (x_1, \dots, x_n)$  and  $\vec{y} = (y_1, \dots, y_n)$  as vectors in  $\mathbb{R}^n$ , the closure condition becomes

$$0 = \sum e_k = \sum z_k^2 = \sum (x_k^2 - y_k^2 + 2\mathbf{i}x_k y_k) = \|\vec{x}\|^2 - \|\vec{y}\|^2 + 2\mathbf{i}\langle \vec{x}, \vec{y} \rangle,$$

where  $\langle \vec{x}, \vec{y} \rangle$  is the standard dot product on  $\mathbb{R}^n$ , and the normalization condition becomes

$$2 = \sum |e_k| = \sum |z_k|^2 = \sum (x_k^2 + y_k^2) = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$

In other words, a pair  $(\vec{x}, \vec{y})$  of  $n$ -dimensional vectors corresponds to a closed polygon of perimeter 2 if and only if

$$\|\vec{x}\| = \|\vec{y}\| = 1 \quad \text{and} \quad \langle \vec{x}, \vec{y} \rangle = 0.$$

Pairs of vectors satisfying the above conditions are called *orthonormal 2-frames* for  $\mathbb{R}^n$ , and the space of all such pairs is the *Stiefel manifold*  $\text{St}_2(\mathbb{R}^n)$ .

Sampling a random point on the Stiefel manifold  $\text{St}_2(\mathbb{R}^n)$  is easy: generate two random vectors  $\vec{u}$  and  $\vec{v}$  from the standard  $n$ -dimensional Gaussian distribution and then apply Gram–Schmidt (or another orthogonalization algorithm which preserves spherical symmetry) to get an orthonormal pair  $(\vec{x}, \vec{y})$ . In *Mathematica*, this can be done with the following function:

```
StiefelSample[n_] := Orthogonalize[RandomVariate[NormalDistribution[], {2, n}]];
```

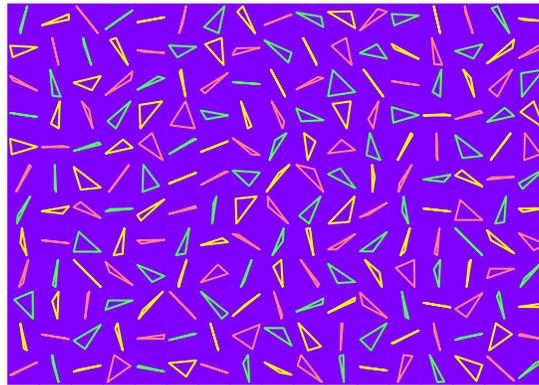
Alternatively, the same distribution is generated by sampling an  $n \times 2$  standard Gaussian matrix  $A$  and letting  $\vec{x}$  and  $\vec{y}$  be the columns of  $Q$  in the QR decomposition  $A = QR$ .<sup>2</sup>

<sup>1</sup>There is also a PDF version available at this location for readers who may not have access to *Mathematica*.

<sup>2</sup>This requires some care since the QR decomposition is not unique. Due to how NAG and LAPACK implement the QR decomposition [8], *Mathematica*, *Matlab*, *NumPy*, *SageMath*, *Julia*, and many other software packages always return a  $Q$  with nonpositive (1, 1) entry, producing a non-uniform distribution of polygons unless this is corrected for.

## Random $n$ -Gons

Given a point  $(\vec{x}, \vec{y})$  on the Stiefel manifold, turning it into a polygon is also simple: the  $k$ th edge of the polygon is simply  $(x_k + \mathbf{i}y_k)^2$ . Using this technique, it is easy to generate large ensembles of random polygons as in Figure 1, which illustrates the slightly remarkable fact – suggested by Wesley Woolhouse [13] and Lewis Carroll [4, Pillow Problem #58] in the 19th century – that the majority of random triangles are obtuse [3].



**Figure 1:** Wallpaper (joint with Anne Harding), showing 204 random triangles, of which 177 (or about 86.8%) are obtuse, compared to a true probability of  $\frac{3}{2} - \frac{3\ln 2}{\pi} \approx 0.838$ .

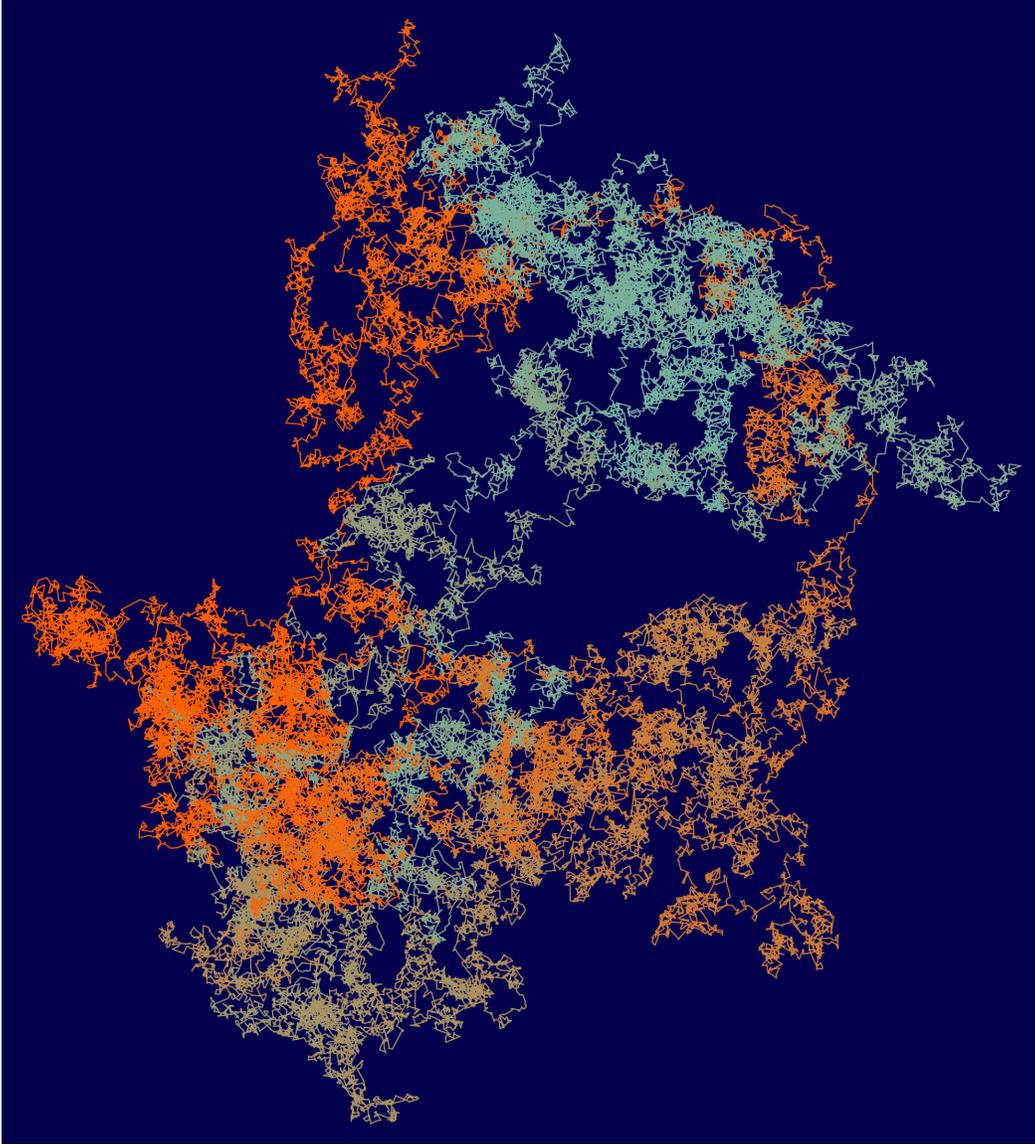
The 100,000-gon in Figure 2 illustrates one curious feature of this parametrization of polygon space: the polygons need not be convex or even embedded. However, producing convex  $n$ -gons is simple: generate a random  $n$ -gon, then permute the edges until the edge directions are sorted by increasing angle with the  $x$ -axis. Here is the convexification of a random 200-gon:  $\bigcirc$ . As discussed in [3], permuting the edges is an isometry of the Stiefel manifold, so this procedure corresponds to uniformly sampling the convex  $n$ -gons inside the space of all  $n$ -gons; see Figure 3 for an example of a random convex pentagon. Unlike taking the convex hull of a collection of random points, this gives precise control on the number of sides of the resulting convex polygon.

Since every quadrilateral tiles the plane, the Stiefel manifold approach gives a way to generate random planar tilings: generate a random point in  $\text{St}_2(\mathbb{R}^4)$ , turn it into a quadrilateral, and then tile the plane with that quadrilateral. Figure 4 shows tilings produced in this way using representatives from each of the three basic classes of quadrilaterals: convex, reflex, and crossed.<sup>3</sup> A beautiful fact proved in [3] (though the result was in some sense anticipated by the science educator, astronomer, future priest, and past Senior Wrangler James Maurice Wilson in 1866 [12]) is that, with respect to this measure on quadrilateral space, exactly  $1/3$  of quadrilaterals are convex,  $1/3$  are reflex, and  $1/3$  are crossed.

## Paths in Polygon Space

In addition to ease of sampling, one of the main mathematical motivations for using the Stiefel manifold model for polygons is that it is straightforward to find paths between pairs of points on the Stiefel manifold, and hence between any two given  $n$ -gons. An obvious choice for such a path would be a geodesic, but unfortunately it is not generally possible to explicitly parametrize geodesics between arbitrary pairs of points in Stiefel manifolds. Instead, the following procedure, used to produce Figure 5, is modeled off the construction of

<sup>3</sup>The crossed quadrilaterals do not tile the plane in the traditional sense, as there are usually overlaps. However, they still tile in a generalized sense: it is possible to choose a fixed orientation on the boundary of the quadrilateral and to cover the plane with copies of the quadrilateral-with-oriented-boundary so that at every point in the plane that does not lie on an edge, the sum of the winding numbers of all the quadrilateral boundaries around that point is  $+1$ .



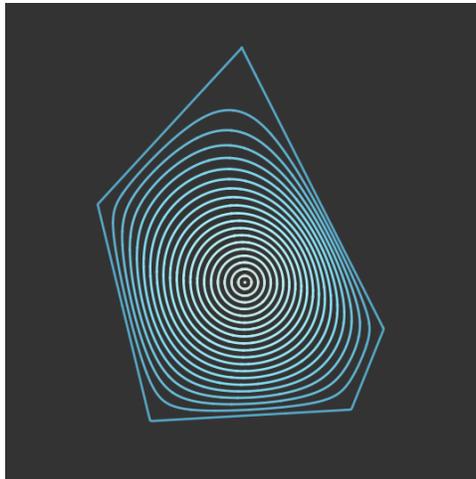
**Figure 2:** A random 100,000-gon, colored so that the first edge is a bright orange (●), the 50,000th is a light green (●), and the edge colors between are interpolated sinusoidally.

geodesics in the Grassmannian discussed below, and hence is a reasonably efficient strategy for traveling between two points, though it almost never produces a geodesic. Given the starting and ending points  $(\vec{x}_0, \vec{y}_0), (\vec{x}_1, \vec{y}_1) \in \text{St}_2(\mathbb{R}^n)$ , let  $\vec{x}(t)$  be the direct rotation of  $\vec{x}_0$  towards  $\vec{x}_1$  and let  $\vec{y}(t)$  be the normalized projection of the direct rotation of  $\vec{y}_0$  towards  $\vec{y}_1$  onto the perpendicular space  $\vec{x}(t)^\perp$ . Then  $(\vec{x}(t), \vec{y}(t))$  is a path in  $\text{St}_2(\mathbb{R}^n)$  from  $(\vec{x}_0, \vec{y}_0)$  to  $(\vec{x}_1, \vec{y}_1)$ .<sup>4</sup>

A more mathematically satisfying approach is to pass from  $\text{St}_2(\mathbb{R}^n)$  to the corresponding *Grassmann manifold*  $\text{Gr}_2(\mathbb{R}^n)$  of 2-dimensional linear subspaces of  $\mathbb{R}^n$ , sending a pair of orthonormal vectors to the subspace they span. This gives an orientation-independent representation of polygons [2], meaning that it does not distinguish between a polygon and a rotated version of the same polygon.

Geodesics in the Grassmannian are well-understood [1, 5]: if  $P_0, P_1 \in \text{Gr}_2(\mathbb{R}^n)$ ,  $M_i$  is an  $n \times 2$  matrix

<sup>4</sup>Reversing the roles of the  $\vec{x}_i$  and the  $\vec{y}_i$  in this procedure generally produces a very similar path.



**Figure 3:** A frame from the animation Nucleation [11]. The Riemann mapping theorem guarantees there is a conformal transformation from the disk to the randomly-generated convex pentagon; the curves show the images of concentric circles under this map.

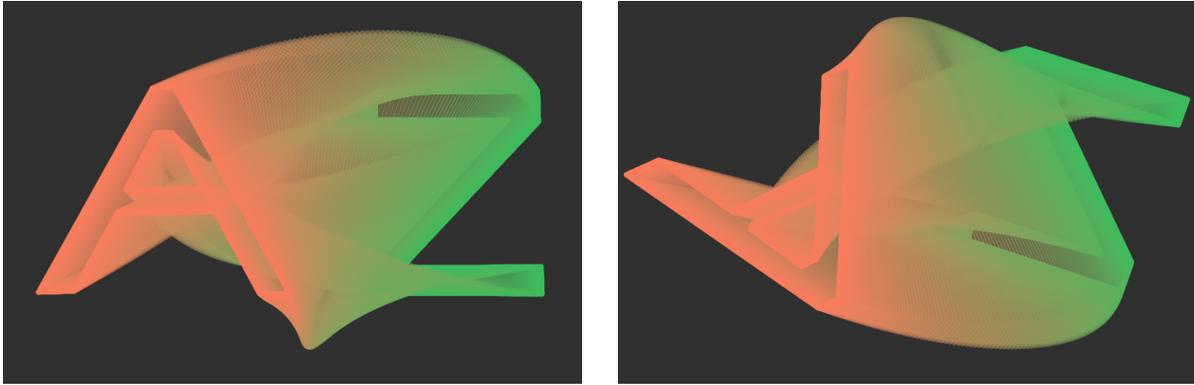


**Figure 4:** Tilings of the plane by convex, reflex, and crossed quadrilaterals.

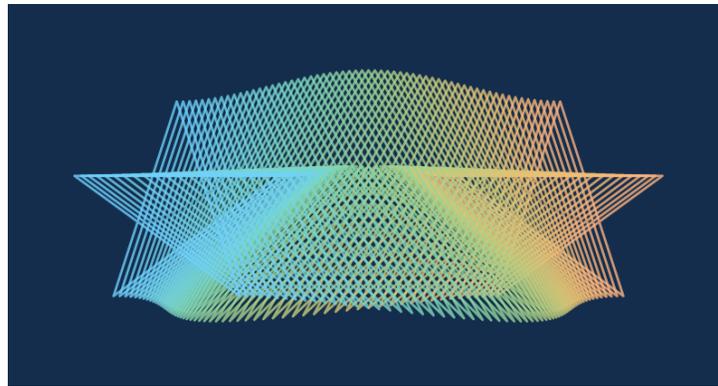
with orthonormal columns spanning  $P_i$ , and  $U\Sigma V^T$  is the singular value decomposition of  $M_0^T M_1$ , let  $(\vec{x}_0, \vec{y}_0)$  be the columns of  $M_0 U$  and  $(\vec{x}_1, \vec{y}_1)$  be the columns of  $M_1 V$ . Then  $(\vec{x}_i, \vec{y}_i)$  is an orthonormal basis for  $P_i$ , and the geodesic in  $\text{Gr}_2(\mathbb{R}^n)$  from  $P_0$  to  $P_1$  corresponds to the path  $(\vec{x}(t), \vec{y}(t))$  in  $\text{St}_2(\mathbb{R}^n)$  where  $\vec{x}(t)$  is the direct rotation of  $\vec{x}_0$  towards  $\vec{x}_1$  and  $\vec{y}(t)$  is the direct rotation of  $\vec{y}_0$  towards  $\vec{y}_1$ . The effect of choosing a different orthonormal basis for an element of  $G_2(\mathbb{R}^n)$  is to rotate the corresponding polygon; the bases for the starting and ending points that arise as above optimally register the polygons to each other. This loss of control of orientation may not be desirable if the original orientations were important.

One curious feature of the Stiefel manifold parametrization of polygon space – which is either a detriment or an opportunity, depending on circumstance – is that there are many choices in how to lift a polygon to a point in the Stiefel manifold. Since this is really a parametrization of *ordered* polygons, it depends on an ordering of the edges. This means that cyclically permuting the edge labels produces  $n$  different paths between visually identical starting and ending polygons.

In fact, there are even more choices: lifting a polygon involved interpreting each edge vector  $e_k$  as a complex number and finding  $z_k$  so that  $z_k^2 = e_k$ . There are generally two possible choices for the square root, so this produces  $2^n$  possible lifts of a given ordered  $n$ -gon. One consequence is that there are many non-trivial paths from a polygon to itself coming from paths between different lifts. Figure 6 shows such a path between different lifts of a pentagonal star.



**Figure 5:** Left: A to Z, showing a path in  $\text{St}_2(\mathbb{R}^{12})$ . Right: the geodesic in  $\text{Gr}_2(\mathbb{R}^{12})$  with the same starting and ending polygons, now registered to each other.



**Figure 6:** The geodesic in  $\text{Gr}_2(\mathbb{R}^5)$  between two different lifts of the same pentagonal star.

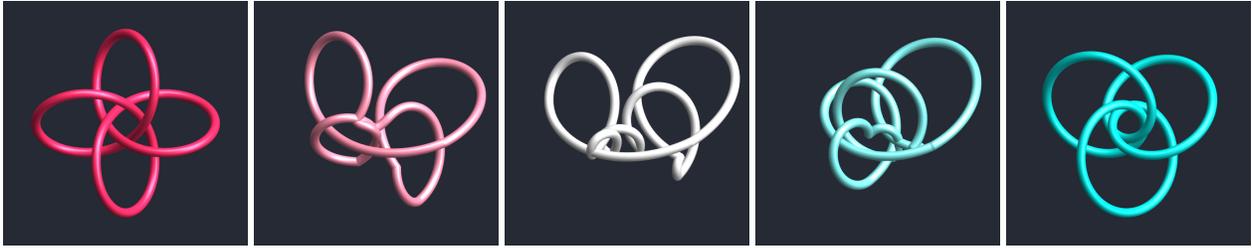
### A Generalization to $n$ -Gons in Space

A generalization of this construction yields a correspondence between polygons in  $\mathbb{R}^3$  and points in the Stiefel manifold  $\text{St}_2(\mathbb{C}^n)$ . The idea is essentially the same as for planar polygons, except now instead of thinking of the edge vectors as complex numbers, we think of them as purely imaginary quaternions, and instead of taking a square root we take a section of the Hopf map  $q \mapsto \bar{q} \mathbf{i} q$ . In this way a closed polygon in space can be turned into a pair  $(\vec{x}, \vec{y})$  of vectors in  $\mathbb{C}^n$  which are orthonormal with respect to the standard Hermitian inner product on  $\mathbb{C}^n$ . Specifically,  $(\vec{x}, \vec{y}) \in \text{St}_2(\mathbb{C}^n)$  is mapped to the polygon in  $\mathbb{R}^3$  whose  $\ell$ th edge is  $(\bar{x}_\ell - y_\ell \mathbf{j}) \mathbf{i} (x_\ell + y_\ell \mathbf{j})$ , which is a purely imaginary quaternion, meaning it can be written as  $a \mathbf{i} + b \mathbf{j} + c \mathbf{k}$  for some  $a$ ,  $b$ , and  $c$ , and hence interpreted as a three-dimensional vector.<sup>5</sup> See [2] for details.

Again, we can both sample random  $n$ -gons and, since it is straightforward to find paths connecting pairs of points in  $\text{St}_2(\mathbb{C}^n)$  or geodesics in the corresponding Grassmannian  $\text{Gr}_2(\mathbb{C}^n)$ , morph any polygon in  $\mathbb{R}^3$  to any other in a natural way, as in Figure 7, which shows a path in  $\text{St}_2(\mathbb{C}^{1000})$  between 1000-gons which form (3, 4) and (4, 3) torus knots.

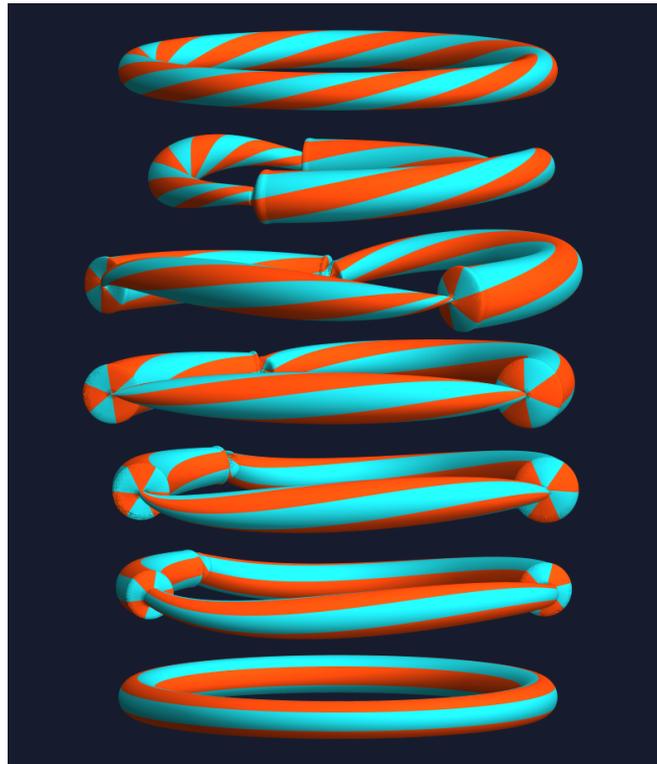
Just as the lift of a planar polygon depended on choices of square roots, lifting a space polygon to  $\text{St}_2(\mathbb{C}^n)$  depends on choices of sections of the Hopf map, meaning there is an entire circle's worth of possible preimages for each edge. This circle corresponds to possible framings on each edge, and lifts of polygons can be interpreted as *inflatable elastic knots*, which are bendable, twistable closed curves with variable thickness.

<sup>5</sup>Here  $\mathbf{i}$ ,  $\mathbf{j}$ , and  $\mathbf{k} = \mathbf{ij}$  are the standard quaternion units.



**Figure 7:** A path in  $\text{St}_2(\mathbb{C}^{1000})$  from a (3, 4)-torus knot to a (4, 3)-torus knot.

Geodesics in the Grassmannian minimize an appropriate generalized Kirchhoff energy [10] which penalizes bending, twisting, and varying the thickness. Figure 8 shows a geodesic between two different framings of the regular 200-gon: the apparent singularities arise because it is energetically advantageous to concentrate most of the bending and twisting in a few locations and moreover because the energy being minimized penalizes bending and twisting more where the knot is thicker, so these locations tend to have (nearly) zero cross-sectional radius. Notice, in particular, that there are three singularities, corresponding to the fact that the ending frame has three twists in it.



**Figure 8:** The geodesic in the Grassmannian  $\text{Gr}_2(\mathbb{C}^{200})$  between the regular 200-gon with the standard blackboard framing (bottom) and the regular 200-gon with a thrice-twisted frame (top).

Natural choices of framing for a polygonal curve include the discrete Frenet–Serret frame (which often approximates a material frame, for example in proteins [7]) and the discrete Bishop frame (which minimizes twist), but the framing ambiguity provides an opportunity for experimentation.

## Summary and Conclusions

The Stiefel manifold parametrizations of polygons both in the plane and in space provide new tools for generative production of polygons and for morphing one polygon into another. While the primary motivation for developing this machinery was to find an exactly solvable model for ring polymers (see [2] for much more detail and background), I believe that this approach should also be useful for both generative and deterministic art, especially since it requires at most a bit of linear algebra to implement.

## Acknowledgements

I am very grateful to the three anonymous referees whose thoughtful comments have substantially improved this paper. This work was supported by a grant from the Simons Foundation (#354225).

## References

- [1] Å. Björck and G. H. Golub. “Numerical methods for computing angles between linear subspaces.” *Mathematics of Computation*, vol. 27, no. 123, 1973, 579–594.
- [2] J. Cantarella, T. Deguchi, and C. Shonkwiler. “Probability theory of random polygons from the quaternionic viewpoint.” *Communications on Pure and Applied Mathematics*, vol. 67, no. 10, 2014, pp. 1658–1699.
- [3] J. Cantarella, T. Needham, C. Shonkwiler, and G. Stewart. “Random triangles and polygons in the plane.” *The American Mathematical Monthly*, vol. 126, no. 2, 2019, pp. 113–134. [arXiv:1702.01027](https://arxiv.org/abs/1702.01027).
- [4] L. Carroll. *Curiosa Mathematica, Part II: Pillow-problems thought out during wakeful hours*, 3rd ed. MacMillan, London, 1894.
- [5] A. Edelman, T. A. Arias, and S. T. Smith. “The geometry of algorithms with orthogonality constraints.” *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, 1999, pp. 303–353.
- [6] J.-C. Hausmann and A. Knutson. “Polygon spaces and Grassmannians.” *L’Enseignement Mathématique*, vol. 43, no. 1–2, 1997, pp. 173–198.
- [7] S. Hu, M. Lundren, and A. J. Niemi. “Discrete Frenet frame, inflection point solitons, and curve visualization with applications to folded proteins.” *Physical Review E*, vol. 83, no. 6, 2011, article id. 061908.
- [8] R. B. Lehoucq. “The computation of elementary unitary matrices.” *ACM Transactions on Mathematical Software*, vol. 22, no. 4, 1996, pp. 393–400.
- [9] M. Mahowald. “On the embeddability of the real projective spaces.” *Proceedings of the American Mathematical Society*, vol. 13, no. 5, 1962, pp. 763–764.
- [10] T. Needham. “Knot types of generalized Kirchhoff rods.” Preprint, 2017. [arXiv:1708.09124](https://arxiv.org/abs/1708.09124).
- [11] C. Shonkwiler. *Nucleation*. 2015. <https://shonkwiler.org/nucleation>.
- [12] J. M. Wilson. “On the Four-point problem and similar Geometrical Chance Problems.” *Mathematical Questions with Their Solutions from the Educational Times*, vol. V, 1866, p. 81. <https://hdl.handle.net/2027/hvd.32044102906633?urlappend=%3Bseq=85>.
- [13] W. S. B. Woolhouse. “Solution to Mathematical Question 1835.” *Mathematical Questions with Their Solutions from the Educational Times*, vol. V, 1866, p. 70. <https://hdl.handle.net/2027/hvd.32044102906633?urlappend=%3Bseq=74>.