Bounding Expected Values on Random Polygons

by

Michael William Berglund

(Under the direction of Jason Cantarella)

Abstract

Random walks of various types have been studied for more than a century. Recently, a new measure on the space of fixed total length random walks in 2 and 3 dimensions was introduced. We will develop de Finetti-style results to help better understand this measure. Along the way, we will demonstrate how to apply these results to better understand these polygons by bounding the expectations of any locally determined quantity, such as curvature or torsion.

Index words: Closed random walk, statistics on Riemannian manifolds, random knot, random polygon

Bounding Expected Values on Random Polygons

by

Michael William Berglund

B.S. The University of West Georgia, 2007

M.A. The University of Georgia, 2011

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2014

Bounding Expected Values on Random Polygons

by

Michael William Beglund

Approved:

Professor: Jason Cantarella

Committee: William H. Kazez
Theodore Shifrin
Michael Usher

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2014

Bounding Expected Values on Random Polygons

Michael William Berglund

April 23, 2014

Contents

Chapter 1

Introduction

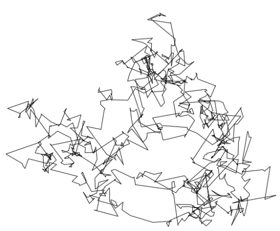

Figure 1.1: Closed 1000 edged Polygon

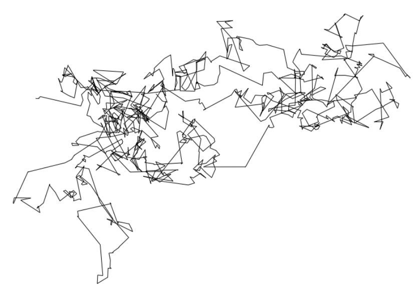

Figure 1.2: Open 1000 edged Polygon

A polymer is a long, flexible, chain-like molecule formed from simpler building block molecules called monomers. When a polymer's ends are joined, it is called a closed polymer, or sometimes a ring polymer. On the other hand, to indicate that the ends are not necessarily joined, we specify that a polymer is open or linear. In a dilute solution with a good solvent, repulsive intermolecular forces between the solvent and monomer subunits dominate over intramolecular interactions, pulling the polymer to occupy relatively large volume. On the other hand, in a solution with a poor solvent, the repulsive intramolecular forces dominate and the chain contracts to occupy relatively small volume. In between these two types of solvents is the theta solvent, where the intermolecular polymer-solvent repulsion balances exactly with the intramolecular monomer-monomer repulsion. Under the theta condition, the polymer behaves like a random walk [36].

In this model, the monomers are rigid rods of a fixed length and their orientations and positions are independent of each other. We will choose to ignore the steric effect, which is the repulsion of atoms whose electron clouds overlap, so that we entertain the possibility that two monomers co-exist at the same place. As such, these polymers are a type of random walk.

Definition 1. Let $\{X_k\}_{k=1}^{\infty}$ be a sequence of independent, identically distributed random variables in $\mathbb{R}^m$. For each positive integer n, we let S_n denote the sum $X_1 + X_2 + \cdots + X_n$. The sequence $\{S_n\}_{n=1}^{\infty}$ is called a random walk in $\mathbb{R}^m$.

Definition 2. Let $\{S_n\}_{n=1}^{\infty}$ be a random walk formed from the random variables $\{X_k\}_{k=1}^{\infty}$. A random open polygon chain with m edges, or open polygon for short, is a polygonal chain whose vertex set is given by $\{\mathbf{0}, S_1, S_2, \ldots, S_m\}$. A random closed polygonal chain, or closed polygon for short, is an open polygon in which $S_m = \mathbf{0}$.

(For more information, Panagiotou-Milllett-Lambropoulou provide a detailed exposition in the introduction of [33]).

In keeping with the connection to polymers, let us for the moment restrict our view to $\mathbb{R}^3$. Defining a probability measure on the space of open polygonal chains is elementary: take a probability measure on $\mathbb{R}^3$ and then sample each edge of the chain independently according to this measure. One specific example arises from require the measure to only be supported on a sphere of appropriate radius. This choice will lead to a probability measure supported on equilateral polygonal arms.

Upon endowing this space with a Riemannian manifold structure, we can then induce the submanifold metric to the subset of closed polygonal chains and obtain a measure from its volume form. It turns out that this closure condition imposes conditions on the distribution of edges that are difficult to understand.

Currently, sampling algorithms used to model random walk representations of polymers, a random walk that requires fixed edge length, are very computationally expensive (such as that shown by Moore-Grosberg in [30]). The Moore-Grosberg algorithm, for instance, is based on computing successive piecewise-polynomial distributions for diagonal lengths and directly sampling from these distributions. These polynomials, however, are of high degree, use large coefficients, and involve many near-cancellations, leading to numerical problems in evaluating them accurately. These methods are discussed by Hughes in [24] in Section 2.5.4. In addition, known theorems show asymptotic behavior without explicit bounds. For example:

Theorem 3. [*42*] *The probability $P(n)$ that an n-gon is knotted is given by* $P(n) = 1 - \exp(-\alpha n + o(n))$ *where α is positive and $o(n)$ denotes an expression $f(n)$ such that for every positive constant ϵ, there exists some $N \in \mathbb{N}$ for which $|f(n)| \leq \epsilon |g(n)|$ for all $n \geq N$.*

Theorem 4. [*10*] *For any given knot type K and any given $p \in (0,1)$, there exists $\alpha > 0$ independent of the number n such that if n is large enough, then with probability at least p, in any random equilateral polygon of length n, there exists a knot of type K with a nice neighborhood of size at least $\alpha\sqrt{n}$ so that the polygon is contained in this neighborhood as a non-trivial loop. Here, a neighborhood of size r of a compact set S is the set of all points whose distance from from S is less than or equal to r, and a neighborhood of K is called nice if it is homotopic to K via a strong deformation retract.*

While these theorems initially look quite good, in practice knots are fairly rare, even for somewhat large values of n. As an explicit example, we sampled 100,000 random 300-gons using the method of [30], but found that only 6,755 of them were knotted, of which only 367 had crossing number larger than 8 and none had crossing number greater than 12.

In this paper, we will consider a different probability measure on polygon spaces—one in which we fix only the total length of the polygon, and allow the individual edge lengths to vary. This probability measure was inspired by the connection between Grassmannians and the set of fixed perimeter polygons established by Hausmann and Knutson in [20] and [21]). These manifolds, which have been studied extensively, have opened the way for many (e.g. Kapovich-Millson [27] and Howard-Manon-Millson [23]) to examine the symplectic and algebraic geometry of polygon spaces. In [5], Cantarella-Deguchi-Shonkwiler develop both a probability measure, known as the symmetric measure, and an algorithm for directly sampling polygons with respect to this measure by utilizing this connection. From this, they calculate a number of exact expectations of physically significant quantities associated with polygons, such as the squared chord length and radius of gyration for both planar and spatial polygons [5], and the total curvature of spatial polygons [6].

Our main goal will be to examine two main questions about this probability measure. In Chapter 1, we will establish a numerical algorithm to efficiently calculate a natural class of expectations with respect to this measure, given by Theorem 76. In Chapter 2, we will apply current de Finetti-style results to justify the intuition that as the edges of the polygons increase, the closure constraint has a lessened effect on local configurations of sub-arms through Theorem 91. We will then create an analogous set of de Finetti results in Chapter 3 to justify that same intuition with space polygons in Theorem 114.

1.1 Arms

Let us begin by talking about polygonal arms.

Definition 5. Consistent with [5], we will define a polygonal arm with n edges in $\mathbb{R}^d$ to be a curve in $\mathbb{R}^d$ specified by a sequence of vertices $(v_0, v_2, \ldots, v_n, v_{n+1})$ consisting of the n line segments joining consecutive pairs of vertices.

$\mathbb{R}^d$ acts on a polygonal arm by translation. While this changes the vertices, it does not change the edge vectors. We can therefore think of the equivalence class of a polygonal arm with n edges up to translation on $\mathbb{R}^d$ as an ordered set of n vectors in $\mathbb{R}^d$.

Definition 6. Following the exposition given in [6], define $\mathscr{A}_d(n)$ to be the moduli space of open n-edge polygonal arms in $\mathbb{R}^d$ up to translation. Likewise, define $\mathscr{P}_d(n)$ to be the moduli space of closed n-edge polygons in $\mathbb{R}^d$ up to translation.

We can further consider the space of polygonal arms up to the action of the group of dilations. For reasons that will be apparent later, we will choose a canonical representative by selecting the polygonal arm which has a total length of 2.

Definition 7. Taking our notation to match that in [5], let $Arm_d(n)$ and $Pol_d(n)$ denote respectively the moduli space of open polygonal arms with n edges in $\mathbb{R}^d$ of total length 2 up to translation in $\mathbb{R}^d$ and the the moduli space of closed polygons with n edges in $\mathbb{R}^d$ of total length 2 up to translation in $\mathbb{R}^d$.

Finally, we can also consider the space of arms up to the action of translations, dilations, and rotations.

Definition 8. As in [5], for $d > 2$, let $\overline{Arm}_d(n)$ and $\overline{Pol}_d(n)$ denote respectively the moduli space of open polygonal arms with n edges in $\mathbb{R}^d$ of total length 2 up to translations and rotations and the moduli space of closed polygons with n edges in $\mathbb{R}^d$ of total length 2 up to translations and rotations.

Definition 9. As in [5], let $\overline{Arm_2}(n)$ and $\overline{Pol_2}(n)$ denote respectively the moduli space of open polygonal arms with n edges in $\mathbb{R}^2$ of total length 2 up to translations, rotations, and reflections and the moduli space of closed polygons with n edges in $\mathbb{R}^2$ of total length 2 up to translations, rotations, and reflections.

Here, we need the additional identification via reflections so that we can realize $\overline{Arm_2}(n)$ as the subset of $\overline{Arm_3}(n)$ of arms contained in the xz-plane, as two planar arms related by reflection across some axis L in the xz-plane are likewise related by the rotation in $\mathbb{R}^3$ about that axis by an angle of π.

Proposition 10. *From [5], we have a commutative diagram given by:*

$$\begin{tikzcd}
 & O(2) \arrow[rr] \arrow[dl, leftrightarrow, "\cong"'] \arrow[dd, dotted] & & Arm_2(n) \arrow[rr] \arrow[dd, dotted] & & \overline{Arm_2}(n) \arrow[dd] \\
O(2) \arrow[rr] \arrow[dd] & & Pol_2(n) \arrow[ur, hook] \arrow[rr] \arrow[dd] & & \overline{Pol_2}(n) \arrow[ur, hook] \arrow[dd] & \\
 & SO(3) \arrow[rr, dotted] \arrow[dl, leftrightarrow, dotted, "\cong"'] & & Arm_3(n) \arrow[rr, dotted] & & \overline{Arm_3}(n) \\
SO(3) \arrow[rr] & & Pol_3(n) \arrow[ur, hook, dotted] \arrow[rr] & & \overline{Pol_3}(n) \arrow[ur, hook] &
\end{tikzcd}$$

1.2 Quaternions

We will now introduce the associative division algebra known as the quaternions. The quaternions share a special connection with our polygon spaces, but we will need to first derive a number of important properties the quaternions enjoy.

There are a number of ways to represent quaternions, and the proofs for many of the properties of the quaternions can be simplified greatly by the choice of representation. We will begin our discussion of the quaternions by defining them first as vector space over $\mathbb{R}$.

Definition 11. The quaternions, $\mathbb{H}$, are a four dimensional vector space over $\mathbb{R}$. Throughout this document, we will let $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ be a fixed basis for $\mathbb{H}$.

Proposition 12. *There is an isomorphism of vector spaces between the quaternions and* $\left\{ \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} : z, w \in \mathbb{C} \right\}$

Proof. It is simple to verify that sending $a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ to the matrix $\begin{bmatrix} a + b\mathbf{i} & c + d\mathbf{i} \\ -c + d\mathbf{i} & a - b\mathbf{i} \end{bmatrix}$ provides such an isomorphism.

□

Proposition 13. *Using this isomorphism, we define a binary operation on $\mathbb{H}$ as multiplication of the matrix representatives.*

Proof. Given $\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$ and $\begin{bmatrix} x & y \\ -\overline{y} & \overline{x} \end{bmatrix}$, consider the product:

$$\begin{aligned}
\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \begin{bmatrix} x & y \\ -\overline{y} & \overline{x} \end{bmatrix} &= \begin{bmatrix} zx - w\overline{y} & zy + w\overline{x} \\ -x\overline{w} - \overline{zy} & -\overline{w}y + \overline{zx} \end{bmatrix} \\
&= \begin{bmatrix} zx - w\overline{y} & zy + w\overline{x} \\ -\overline{\overline{x}w + zy} & \overline{-w\overline{y} + zx} \end{bmatrix} \\
&= \begin{bmatrix} zx - w\overline{y} & zy + w\overline{x} \\ -\overline{zy + w\overline{x}} & \overline{zx - w\overline{y}} \end{bmatrix}
\end{aligned}$$

which has the form $\begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}$ for $a = zx - w\overline{y}$ and $b = zy + w\overline{x}$. As such, the matrix product of the matrix representation of two quaternions is the matrix representation of a new quaternion. □

Proposition 14. *Quaternionic multiplication enjoys the following four properties: Let $x, y, z \in \mathbb{H}$ and $a, b \in \mathbb{R}$ be arbitrary.*

1. *Left distributivity:* $(x + y)z = xz + yz$

2. *Right distributivity:* $x(y + z) = xy + xz$

3. *Compatibility with scalars:* $(ax)(by) = (ab)(xy)$

4. *Associativity:* $x(yz) = (xy)z$

Proof. These properties are all inherited from matrix multiplication. □

Proposition 15. *There is a multiplicative identity.*

Proof. While this is also inherited from matrix multiplication, we must also point out that the multiplicative identity of matrix multiplication is in fact a quaternion. Specifically, that $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ has the form $\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$. □

Additionally, we note here that in the basis representation, the multiplicative identity is give by the basis element **1**.

Proposition 16. *Quaternionic multiplication admits a unique inverse for each non-zero quaternion.*

Proof. For a non-zero quaternion $\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$, we see that the determinant of this matrix is $z\overline{z} + w\overline{w} = \|z\|^2 + \|w\|^2$ is a non-zero real number. As such, it has a matrix inverse given by $\dfrac{1}{\|z\|^2 + \|w\|^2} \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix}$. This matrix has the appropriate form, and so the inverse is indeed a quaternion. □

Proposition 17. $\mathbb{H}$ *is an associative division algebra over the base field* $\mathbb{R}$.

Proof. We have shown in Propositions 13, 14, and 15 that the vector space $\mathbb{H}$ over $\mathbb{R}$ has a bilinear product which is associative and possesses an identity element. As such it is an associative algebra. We also showed in Proposition 16 that every non-zero quaternion has a multiplicative inverse. As such, we have shown that $\mathbb{H}$ is a division algebra. □

Definition 18. We define the vector space isomorphism between $\mathbb{H}$ and $\mathbb{R}^4$ to be $G : \mathbb{H} \to \mathbb{R}^4$ so that $G(q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}) = (q_0, q_1, q_2, q_3)$.

Proposition 19. *In the basis representation, quaternionic multiplication of* $\mathbf{x} = x_1\mathbf{1} + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}$ *and* $\mathbf{y} = y_1\mathbf{1} + y_2\mathbf{i} + y_3\mathbf{j} + y_4\mathbf{k}$ *is given by:*

$$\begin{aligned}\mathbf{xy} &= (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4)\mathbf{1}\\ &+ (x_1y_2 + x_2y_2 + x_3y_4 - x_4y_3)\mathbf{i}\\ &+ (x_1y_3 - x_2y_4 + x_3y_2 + x_4y_2)\mathbf{j}\\ &+ (x_1y_4 + x_2y_3 - x_3y_2 + x_4y_2)\mathbf{k}\end{aligned}$$

Proof. Writing these as matrices, we have

$$\begin{aligned}\mathbf{xy} &= \begin{bmatrix} x_1 + x_2\mathbf{i} & x_3 + x_4\mathbf{i} \\ -x_3 + x_4\mathbf{i} & x_1 - x_2\mathbf{i} \end{bmatrix} \begin{bmatrix} y_1 + y_2\mathbf{i} & y_3 + y_4\mathbf{i} \\ -y_3 + y_4\mathbf{i} & y_1 - y_2\mathbf{i} \end{bmatrix}\\ &= \begin{bmatrix} (x_1 + x_2\mathbf{i})(y_1 + y_2\mathbf{i}) & (x_1 + x_2\mathbf{i})(y_3 + y_4\mathbf{i}) \\ (-x_3 + x_4\mathbf{i})(y_1 + y_2\mathbf{i}) & (-x_3 + x_4\mathbf{i})(y_3 + y_4\mathbf{i}) \end{bmatrix}\\ &+ \begin{bmatrix} (x_3 + x_4\mathbf{i})(-y_3 + y_4\mathbf{i}) & (x_3 + x_4\mathbf{i})(y_1 - y_2\mathbf{i}) \\ (x_1 - x_2\mathbf{i})(-y_3 + y_4\mathbf{i}) & (x_1 - x_2\mathbf{i})(y_1 - y_2\mathbf{i}) \end{bmatrix}\end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4) & (x_1y_3 - x_2y_4 + x_3y_2 + x_4y_2) \\ -(x_1y_3 - x_2y_4 + x_3y_2 + x_4y_2) & (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4) \end{bmatrix} \\
&+ \begin{bmatrix} (x_1y_2 + x_2y_2 + x_3y_4 - x_4y_3) & (x_1y_4 + x_2y_3 - x_3y_2 + x_4y_2) \\ (x_1y_4 + x_2y_3 - x_3y_2 + x_4y_2) & -(x_1y_2 + x_2y_2 + x_3y_4 - x_4y_3) \end{bmatrix} \mathbf{i} \\
&= (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4)\mathbf{1} \\
&+ (x_1y_2 + x_2y_2 + x_3y_4 - x_4y_3)\mathbf{i} \\
&+ (x_1y_3 - x_2y_4 + x_3y_2 + x_4y_2)\mathbf{j} \\
&+ (x_1y_4 + x_2y_3 - x_3y_2 + x_4y_2)\mathbf{k}
\end{aligned}
$$

□

Definition 20. The real part of a quaternion $\mathbf{q} = q_0\mathbf{1}+q_1\mathbf{i}+q_2\mathbf{j}+q_3\mathbf{k}$ is q_0 and the imaginary part of a quaternion is $q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$. We say that a quaternion is real if $q_1 = q_2 = q_3 = 0$. Likewise, we say that a quaternion is purely imaginary if $q_0 = 0$.

Definition 21. The norm of a quaternion is norm of the corresponding vector in $\mathbb{R}^4$.

Proposition 22. *The norm of a quaternion is equal to the square root of the determinant of the matrix representation of the quaternion.*

Proof. We saw in the proof of Proposition 16 that the determinant of the matrix representation of the quaternion $\mathbf{q} = q_0\mathbf{1} + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ is given by $q_0^2 + q_1^2 + q_2^2 + q_3^2$. Which we can easily verify to be the square of the norm of the corresponding vector. □

Now that we have established that the quaternions are a normed associative algebra with a natural global isometry with $\mathbb{R}^4$, we will introduce two new endomorphisms. In discussing them, it will prove useful to have a quick reference of the products of basis vectors which we will provide now.

Proposition 23. *The products of the basis vectors* $\mathbf{q_1 q_2}$ *can be summarized in the following table:*

		$\mathbf{q_2}$			
		1	**i**	**j**	**k**
$\mathbf{q_1}$	**1**	**1**	**i**	**j**	**k**
	i	**i**	**-1**	**k**	**-j**
	j	**j**	**-k**	**-1**	**i**
	k	**k**	**j**	**-i**	**-1**

Definition 24. The conjugate of a quaternion $\mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ is the quaternion $\overline{\mathbf{q}} = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k}$.

Proposition 25. *For any* $\mathbf{q}, \mathbf{r} \in \mathbb{H}$, $\overline{(\mathbf{qr})} = (\overline{\mathbf{r}})(\overline{\mathbf{q}})$.

Proof. Consider the map $f_{\mathbf{q}} : \mathbb{H} \to \mathbb{H}$ given by $f_{\mathbf{q}}(\mathbf{r}) = \overline{(\mathbf{qr})} - (\overline{\mathbf{r}})(\overline{\mathbf{q}})$. We can see that for any $\mathbf{q} \in \mathbb{H}$, the map $f_{\mathbf{q}}$ is linear. We can then specify what happens to any $\mathbf{r}$ by looking at what happens to the basis vectors.

Write $\mathbf{q} = q_0\mathbf{1} + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$. It is then clear that $f_{\mathbf{q}}(\mathbf{1}) = 0$, as $\mathbf{1} = \overline{\mathbf{1}}$ is the multiplicative identity.

Likewise, $f_{\mathbf{q}}(\mathbf{i}) = \overline{(\mathbf{qi})} + (\mathbf{i}\overline{\mathbf{q}}) = (-q_1\mathbf{1} - q_0\mathbf{i} - q_3\mathbf{j} + q_2\mathbf{k}) + (q_1\mathbf{1} + q_0\mathbf{i} + q_3\mathbf{j} - q_2\mathbf{k}) = 0$,
$f_{\mathbf{q}}(\mathbf{j}) = \overline{(\mathbf{qj})} + (\mathbf{j}\overline{\mathbf{q}}) = (-q_2\mathbf{1} + q_3\mathbf{i} - q_0\mathbf{j} - q_1\mathbf{k}) + (q_2\mathbf{1} - q_3\mathbf{i} + q_0\mathbf{j} + q_1\mathbf{k}) = 0$, and
$f_{\mathbf{q}}(\mathbf{k}) = \overline{(\mathbf{qk})} + (\mathbf{k}\overline{\mathbf{q}}) = (-q_3\mathbf{1} - q_2\mathbf{i} + q_1\mathbf{j} - q_0\mathbf{k}) + (q_3\mathbf{1} + q_2\mathbf{i} - q_1\mathbf{j} + q_0\mathbf{k}) = 0$

Since this is a linear map that evaluates to zero on each basis vector, we see that $f_{\mathbf{q}}$ is the zero map for any $\mathbf{q}$. Hence, we may conclude that for any $\mathbf{q}, \mathbf{r} \in \mathbb{H}$ we have the equality $\overline{(\mathbf{qr})} = (\overline{\mathbf{r}})(\overline{\mathbf{q}})$ □

Proposition 26. *The square of the norm of a quaternion* $\mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ *is given by* $\|\mathbf{q}\|^2 = \mathbf{q}\overline{\mathbf{q}}$

Proof. We have defined already the norm of the quaternion to be $\|\mathbf{q}\| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}$, so we need only check that $\mathbf{q}\overline{\mathbf{q}} = q_0^2 + q_1^2 + q_2^2 + q_3^2$:

$$\begin{aligned}
\mathbf{q}\overline{\mathbf{q}} &= \begin{bmatrix} q_0 + q_1\mathbf{i} & q_2 + q_3\mathbf{i} \\ -q_2 + q_3\mathbf{i} & q_0 - q_1\mathbf{i} \end{bmatrix} \begin{bmatrix} q_0 - q_1\mathbf{i} & -q_2 - q_3\mathbf{i} \\ q_2 - q_3\mathbf{i} & q_0 + q_1\mathbf{i} \end{bmatrix} \\
&= \begin{bmatrix} q_0^2 + q_1^2 + q_2^2 + q_3^2 & (q_0 + q_1\mathbf{i})(-q_2 - q_3\mathbf{i} + q_2 + q_3\mathbf{i}) \\ (q_0 - q_1\mathbf{i})(-q_2 + q_3\mathbf{i} + q_2 - q_3\mathbf{i}) & q_0^2 + q_1^2 + q_2^2 + q_3^2 \end{bmatrix} \\
&= (q_0^2 + q_1^2 + q_2^2 + q_3^2)\mathbf{1}
\end{aligned}$$

□

Proposition 27. *The norm of any quaternion is equal to the norm of its conjugate.*

Proof. Given $\mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, we have that $\overline{\mathbf{q}} = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k}$. Working out the norm then gives us:

$$\begin{aligned}
\|\overline{\mathbf{q}}\| &= (q_0^2 + (-q_1)^2 + (-q_2)^2 + (-q_3)^2)^{1/2} \\
&= (q_0^2 + q_1^2 + q_2^2 + q_3^2)^{1/2} \\
&= \|\mathbf{q}\|
\end{aligned}$$

□

Proposition 28. *The norm of a product of quaternions is the product of their norms.*

Proof. From Proposition 22, we see that the norm of a quaternion is the square root of the determinant of the matrix representation of the quaternion. Recall that the determinant of a product of matrices is the product of the determinants of the matrices. Likewise, the square root of a product of positive numbers is the product of the square roots of the numbers. The result follows immediately from these two facts. □

Definition 29. The Hopf map on $\mathbb{H}$ is the mapping given by $\mathbf{q} \mapsto \overline{\mathbf{q}}\mathbf{i}\mathbf{q}$

Proposition 30. *The* Hopf *map takes any quaternion to a purely imaginary quaternion. Specifically, for* $\mathbf{q} = q_0\mathbf{1} + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, *we have*

$$\mathrm{Hopf}(\mathbf{q}) = (q_0^2 + q_1^2 - q_2^2 - q_3^2)\mathbf{i} + 2(q_1q_2 - q_0q_3)\mathbf{j} + 2(q_0q_2 + q_1q_3)\mathbf{k}$$

Proof. The proof of this claim is a straightforward calculation:

$$\begin{aligned}
\overline{\mathbf{q}}\mathbf{i}\mathbf{q} &= \begin{bmatrix} q_0 - q_1\mathbf{i} & -q_2 - q_3\mathbf{i} \\ q_2 - q_3\mathbf{i} & q_0 + q_1\mathbf{i} \end{bmatrix} \begin{bmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{bmatrix} \begin{bmatrix} q_0 + q_1\mathbf{i} & q_2 + q_3\mathbf{i} \\ -q_2 + q_3\mathbf{i} & q_0 - q_1\mathbf{i} \end{bmatrix} \\
&= \mathbf{i} \begin{bmatrix} q_0 - q_1\mathbf{i} & -q_2 - q_3\mathbf{i} \\ q_2 - q_3\mathbf{i} & q_0 + q_1\mathbf{i} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} q_0 + q_1\mathbf{i} & q_2 + q_3\mathbf{i} \\ -q_2 + q_3\mathbf{i} & q_0 - q_1\mathbf{i} \end{bmatrix} \\
&= \mathbf{i} \begin{bmatrix} q_0 - q_1\mathbf{i} & q_2 + q_3\mathbf{i} \\ q_2 - q_3\mathbf{i} & -q_0 - q_1\mathbf{i} \end{bmatrix} \begin{bmatrix} q_0 + q_1\mathbf{i} & q_2 + q_3\mathbf{i} \\ -q_2 + q_3\mathbf{i} & q_0 - q_1\mathbf{i} \end{bmatrix} \\
&= \mathbf{i} \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & (q_0 - q_1\mathbf{i})(2q_2 + 2q_3\mathbf{i}) \\ (q_0 + q_1\mathbf{i})(2q_2 - 2q_3\mathbf{i}) & -q_0^2 - q_1^2 + q_2^2 + q_3^2 \end{bmatrix} \\
&= \mathbf{i} \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_0q_2 + q_1q_3) + 2(q_0q_3 - q_1q_2)\mathbf{i} \\ 2(q_0q_2 + q_1q_3) + 2(q_1q_2 - q_0q_3)\mathbf{i} & -q_0^2 - q_1^2 + q_2^2 + q_3^2 \end{bmatrix} \\
&= \begin{bmatrix} (q_0^2 + q_1^2 - q_2^2 - q_3^2)\mathbf{i} & 2(q_1q_2 - q_0q_3) + 2(q_0q_2 + q_1q_3)\mathbf{i} \\ 2(q_0q_3 - q_1q_2) + 2(q_0q_2 + q_1q_3)\mathbf{i} & (-q_0^2 - q_1^2 + q_2^2 + q_3^2)\mathbf{i} \end{bmatrix} \\
&= (q_0^2 + q_1^2 - q_2^2 - q_3^2)\mathbf{i} + 2(q_1q_2 - q_0q_3)\mathbf{j} + 2(q_0q_2 + q_1q_3)\mathbf{k}
\end{aligned}$$

□

Proposition 31. *For all* $\mathbf{q} \in \mathbb{H}$, $\|\mathrm{Hopf}(\mathbf{q})\| = \|\mathbf{q}\|^2$

Proof. $\|\mathrm{Hopf}(\mathbf{q})\| = \|\overline{\mathbf{q}}\mathbf{i}\mathbf{q}\| = \|\overline{\mathbf{q}}\| * \|\mathbf{i}\| * \|\mathbf{q}\| = \|\mathbf{q}\| * 1 * \|\mathbf{q}\| = \|\mathbf{q}\|^2$

□

Definition 32. Let $\mathbb{IH}$ be the additive subgroup of the ring $\mathbb{H}$ formed from the purely imaginary quaternions.

Definition 33. Define the map $I : \mathbb{IH} \to \mathbb{R}^3$ to be the linear isomorphism which sends $(0\mathbf{1} + q_1\mathbf{i} + q_2\mathbf{j} + q_1\mathbf{k}) \mapsto (q_1, q_2, q_3)$.

Proposition 34. *Where $\pi : \mathbb{R}^4 \to \mathbb{R}^3$ is projection to the last three coordinates, we have that $I = \pi \circ G|_{\mathbb{IH}}$*

Proof. To see this, we need only recall that G takes a quaternion to the corresponding vector in $\mathbb{R}^4$.

$\square$

Definition 35. Define the map $\mathrm{Hopf}_{\mathbb{R}} : \mathbb{H} \to \mathbb{R}^3$ to be $\mathrm{Hopf}_{\mathbb{R}} = I \circ \mathrm{Hopf}$.

Proposition 36. $\mathrm{Hopf}_{\mathbb{R}}$ *is a surjection. Moreover,* $\mathrm{Hopf}_{\mathbb{R}}|_{\mathbb{IH}}$ *is a surjection.*

Proof. In coordinates,
$\mathrm{Hopf}_{\mathbb{R}}(q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}) = (q_0^2 + q_1^2 - q_2^2 - q_3^2, 2(q_1q_2 - q_0q_3), 2(q_0q_2 + q_1q_3))$. In particular, notice that if $\mathbf{q} \in \mathbb{IH}$, then $\mathrm{Hopf}_{\mathbb{R}}(\mathbf{q})$ can be simplified to $(q_1^2 - q_2^2 - q_3^2, 2q_1q_2, 2q_1q_3)$. For a given element $(x, y, z) \in \mathbb{R}^3$, if $\mathbf{q} \in \left(\mathbb{IH} \bigcap \mathrm{Hopf}_{\mathbb{R}}^{-1}(\{(x, y, z)\})\right)$, then we know that:

1. $q_1^2 - q_2^2 - q_3^2 = x$
2. $2q_1q_2 = y$
3. $2q_1q_3 = z$

Let us treat this in two cases:

Case 1: $y^2 + z^2 \neq 0$.

In this case, we see that q_1 is a solution to the equation $q_1^2 - \left(\dfrac{y}{2q_1}\right)^2 - \left(\dfrac{z}{2q_1}\right)^2 = x$, which we can write as $q_1^4 - xq_1^2 - \dfrac{y^2 + z^2}{4} = 0$. Letting $w = q_1^2$, this becomes the quadratic equation $w^2 - xw - \dfrac{(y^2+z^2)}{4} = 0$. Given that this quadratic has a negative constant term, we know that it has one positive and one negative zero. Let w_0 denote the positive zero. Then we have shown that the quaternion $0\mathbf{1} + \sqrt{w_0}\mathbf{i} + \dfrac{y}{2\sqrt{w_0}}\mathbf{j} + \dfrac{z}{2\sqrt{w_0}}\mathbf{k}$ maps to (x,y,z).

Case 2: $y^2 + z^2 = 0$.

In this case, we have that $y = 0$ and $z = 0$. If $x \leq 0$, then setting $q_1 = 0$ and $q_2 = q_3 = \sqrt{\dfrac{-x}{2}}$ produces a preimage of (x,y,z). If $x > 0$, then setting $q_1 = \sqrt{x}$, and $q_2 = q_3 = 0$ will produce a preimage of (x,y,z). Hence we have shown that $\mathrm{Hopf}_{\mathbb{R}}$ is surjective, and that for all $(x,y,z) \in \mathbb{R}^3$, $\left(\mathbb{IH} \bigcap \mathrm{Hopf}_{\mathbb{R}}^{-1}(\{(x,y,z)\})\right) \neq \emptyset$. $\square$

Now that we know that $\mathrm{Hopf}_{\mathbb{R}}$ is surjective onto $\mathbb{R}^3$, we may wonder what the preimage of any given point is. From Proposition 31, we can see that any point $\vec{p} \in \mathbb{R}^3$ has $\mathrm{Hopf}_{\mathbb{R}}^{-1}(\{\vec{p}\}) \subset S^3(\sqrt{\|\vec{p}\|}) \subset \mathbb{H}$. The specific subset turns out to be a great circle of S^3, as we will show now.

Proposition 37. *The quaternion* $\mathbf{q}$ *satisfies* $\overline{\mathbf{q}}\mathbf{i}\mathbf{q} = \mathbf{i}$ *if and only if we can express* $\mathbf{q} = a\mathbf{1} + b\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}$ *with* $a^2 + b^2 = 1$.

Proof. Write the quaternion as $q = z\mathbf{i} + w\mathbf{j}$ for $z = a + b\mathbf{i}$ and $w = c + d\mathbf{i}$. Writing the product in terms of matrices, we have:

$$\begin{aligned}
\overline{\mathbf{q}}\mathbf{i}\mathbf{q} &= \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix} \begin{bmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{bmatrix} \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \\
&= \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix} \begin{bmatrix} z\mathbf{i} & w\mathbf{i} \\ \overline{w}\mathbf{i} & -\overline{z}\mathbf{i} \end{bmatrix} \\
&= \begin{bmatrix} (|z|^2 - |w|^2)\mathbf{i} & 2\overline{z}w\mathbf{i} \\ 2\overline{w}z\mathbf{i} & (|w|^2 - |z|^2)\mathbf{i} \end{bmatrix}.
\end{aligned}$$

This is equal to $\mathbf{i}$ if and only if $(|z|^2 - |w|^2) = 1$ and $2\overline{z}w = 0$. This pair of equation is satisfied precisely when $w = 0$ and $|z| = 1$. □

Proposition 38. *The two quaternions* $\mathbf{q}$ *and* $\mathbf{r}$ *have the same image under* $\mathrm{Hopf}_{\mathbb{R}}$ *if and only if there exists some complex number* $z \in \mathbb{C}$ *of unit length so that* $\mathbf{r} = z\mathbf{q}$.

Proof. Suppose that $\mathrm{Hopf}_{\mathbb{R}}(\mathbf{q}) = \mathrm{Hopf}_{\mathbb{R}}(\mathbf{r})$. Then we know that $\overline{\mathbf{q}}\mathbf{i}\mathbf{q} = \overline{\mathbf{r}}\mathbf{i}\mathbf{r}$. We may express this as $\mathbf{i} = (\overline{\mathbf{q}})^{-1}(\overline{\mathbf{r}}\mathbf{i}\mathbf{r})\mathbf{q}^{-1} = (\overline{\mathbf{q}^{-1}}\,\overline{\mathbf{r}})\mathbf{i}(\mathbf{r}\mathbf{q}^{-1}) = (\overline{\mathbf{r}\mathbf{q}^{-1}})\mathbf{i}(\mathbf{r}\mathbf{q}^{-1})$. By Proposition 37, we know that this is only possible if $\mathbf{r}\mathbf{q}^{-1}$ is a unit length complex number, z. Finally, we may write $\mathbf{r}\mathbf{q}^{-1} = z$ as $\mathbf{r} = z\mathbf{q}$ as desired. □

Proposition 39. *There is a surjective map* $H : \mathbb{H}^n \to \mathscr{A}_3(n)$ *that restricts to a surjective map* $h : S^{4n-1}(\sqrt{2}) \subset \mathbb{H}^n \to Arm_3(n)$.

Proof. First, note that an element of $\mathscr{A}_3(n)$ is an equivalence class of polygonal arms up to translation. As such, it is representable uniquely as an ordered set of n edges. Define $H_2 : \mathbb{H}^n \to (\mathbb{R}^3)^n$ to be the map that applies $\mathrm{Hopf}_{\mathbb{R}}$ coordinate-wise. Define the map H_3 to be the map that sends an element $(v_1, v_2, \ldots, v_n) \in (\mathbb{R}^3)^n$ to the equivalence class of polygonal arms represented by the edge set $\{v_1, v_2, \ldots, v_n\}$. Finally, define $H = H_3 \circ H_2$. Surjectivity of H follows from that of $\mathrm{Hopf}_{\mathbb{R}}$.

Next, we claim that the restriction $h = H|_{S^{4n-1}(\sqrt{2})}$ gives a surjection to $Arm_3(n)$. In fact more is true, we shall show that $S^{4n-1}(\sqrt{2}) = H^{-1}(Arm_3(n))$, and surjection onto this subset will follow from H being a surjection to $\mathscr{A}_3(n)$. The proof of this claim follows from the following lemma, which we will prove following this proof.

Lemma 40. *For any* $\vec{q} \in \mathbb{H}^n$, *the length of the polygonal arm* $H(\vec{\mathbf{q}})$ *is equal to* $\|\tilde{\mathbf{q}}\|^2$.

By the lemma, we see that the set $S^{4n-1}(r) \subset \mathbb{H}^n$ is precisely the preimage of the set of equivalence classes of polygonal arms of total length r^2. As such, we have that h is is a surjection from $S^{4n-1}(\sqrt{2}) = H^{-1}(Arm_3(n))$ to $Arm_3(n)$. □

Proof of Lemma 40. For a given $\tilde{\mathbf{q}} = (\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}) \in \mathbb{H}^n$, note that the total length of $H(\tilde{\mathbf{q}})$ is equal to the sum of the lengths of its edges. Each edge, e_i, is obtained as $\operatorname{Hopf}(\mathbf{q_i})$, and so its length is equal to $\|\operatorname{Hopf}(\mathbf{q_i})\|$. Putting this together, we see that the length of $H(\vec{\mathbf{q}})$ is given by:

$$L = \sum_{i=1}^{n} \|e_i\| = \sum_{i=1}^{n} \|\operatorname{Hopf}(\mathbf{q_i})\| = \sum_{i=1}^{n} \|\mathbf{q_i}\|^2 = \|\vec{\mathbf{q}}\|^2.$$

□

Proposition 41. *Given scalars $a, b \in \mathbb{R}$, we have that*
$\operatorname{Hopf}(a\mathbf{1} + b\mathbf{j}) = \operatorname{Hopf}(a\mathbf{i} + b\mathbf{k}) = (a^2 - b^2)\mathbf{i} + (2ab)\mathbf{k}$

Proof. Applying the coordinate form of the Hopf map to these quaternions produces the desired result. □

Since we have that $\mathscr{A}_2(n) \subsetneq \mathscr{A}_3(n)$, it is natural to wonder what the preimage of $\mathscr{A}_2(n)$ is under the map H. We will now show that it is in fact a $2n$-dimensional subspace of $\mathbb{H}^n$. We will further show that there is a natural embedding of $\mathbb{C}^n$ into $\mathbb{H}^n$ for which the composition with H is simply the coordinate-wise squaring map (after identifying the polygon in $\mathscr{A}_2(n)$ with the vector of edges in $\mathbb{C}^n$).

Definition 42. Define the maps $C_1, C_2 : \mathbb{C} \to \mathbb{H}$ by $C_1(a + b\mathbf{i}) = a\mathbf{1} + b\mathbf{j}$ and
$C_2(a + b\mathbf{i}) = a\mathbf{i} + b\mathbf{k}$.

Proposition 43. *C_1 and C_2 are linear $\mathbb{R}$-module homomorphisms.*

Proof. Since C_1 and C_2 are defined by sending the basis vectors of the $\mathbb{R}$-module $\mathbb{C}$ to distinct basis vectors of the $\mathbb{R}$-module $\mathbb{H}$ and extending linearly, we can see the result immediately. □

Definition 44. Define the maps $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ to be projection onto the first and third coordinates.

Definition 45. Define the maps $i : \mathbb{R}^2 \to \mathbb{C}$ be the linear isomorphism that sends the vector (a, b) to the number $a + b\mathbf{i}$.

Proposition 46. *Both maps* $i \circ \pi \circ \mathrm{Hopf}_{\mathbb{R}} \circ C_j : \mathbb{C} \to \mathbb{C}$ *for* $j \in \{1, 2\}$ *are equal to the map* $z \mapsto z^2$.

Proof. First, note that for $z = a + b\mathbf{i}$, we have from Proposition 30 that $(\mathrm{Hopf}_{\mathbb{R}} \circ C_j)(z) = I(\mathrm{Hopf}(C_j(a + b\mathbf{i})))$ will be equal to the vector $(a^2 - b^2, 0, 2ab)$. Next, the projection π sends this to the vector $(a^2 - b^2, 2ab)$. Finally, i sends this to the number $(a^2 - b^2) + 2ab\mathbf{i}$. It is then a simple computation to verify that when written as $z = a + b\mathbf{i}$, we have that $z^2 = (a^2 - b^2) + 2ab\mathbf{i}$. □

Proposition 47. *There is a surjective map* $P : \mathbb{C}^n \to \mathscr{A}_2(n)$ *that restricts to a map* $p : S^{2n-1}(\sqrt{2}) \subset \mathbb{C}^n \to Arm_2(n)$.

Proof. We first recall that an element of $\mathscr{A}_2(n)$ is representable uniquely as an ordered set of n edges. We define $H_4 : \mathbb{C}^n \to \mathbb{C}^n$ to be the coordinate-wise squaring map, $(z_1, z_2, \ldots, z_n) \mapsto (z_1^2, z_2^2, \ldots, z_n^2)$. We then define H_5 to be the map that sends an element $(v_1, v_2, \ldots, v_n) \in (\mathbb{R}^2)^n$ to the equivalence class of polygonal arms represented by the edge set $\{v_1, v_2, \ldots, v_n\}$. Where $i : \mathbb{R}^2 \to \mathbb{C}$, is the isomorphism of vector spaces given in the previous proposition, define $P = H_5 \circ i^{-1} \circ H_4$. Surjectivity follows from the surjectivity of each of the three individual maps.

For a given vector $\vec{z} = \vec{a} + \vec{b}\mathbf{i}$, we then have that the length of the polygonal arm $P(\vec{z})$ is given by:

$$L = \sum_{j=1}^{n} |e_j| = \sum_{j=1}^{n} \|(a_j^2 - b_j^2, 2a_j b_j)\| = \sum_{j=1}^{n} |(a_j^2 - b_j^2) + 2a_j b_j \mathbf{i}| = \sum_{j=1}^{n} |z_j^2|$$
$$= \sum_{j=1}^{n} |z_j||z_j| = \sum_{j=1}^{n} |z_j||\overline{z_j}| = \sum_{j=1}^{n} |z_j \overline{z_j}| = \sum_{j=1}^{n} z_j \overline{z_j} = \|\vec{z}\|^2.$$

Hence, the image of a complex vector $\vec{z}$ has total length $\|z\|^2$. This, along with surjectivity, shows that $Arm_2(n) = P^{-1}(S^{2n-1}(\sqrt{2})$. So $p = P|_{S^{2n-1}(\sqrt{2})}$ is the desired surjection. $\square$

1.3 Closed Polygons

We now have a way to model polygonal arms as points on a sphere. This leads us to wonder what we can say about the subset that corresponds to closed polygons. In this section we will show that this subset is a well known manifold and use the Hopf map to establish a probability measure on our polygon spaces. First, recall that we have defined the following spaces:

- $\mathscr{A}_d(n)$, the moduli space of open n-edged polygonal arms in $\mathbb{R}^d$ up to translation.
- $\mathscr{P}_d(n)$, the subset of $\mathscr{A}_d(n)$ consisting of the closed polygons.
- $Arm_d(n)$, the subset of $\mathscr{A}_d$ consisting of all the polygons with total length 2.
- $Pol_d(n)$, the subset of $Arm_d(n)$ consisting of the closed polygons.

Lemma 48. *For $\mathbf{q} \in \mathbb{H}$ expressed as $\mathbf{q} = a + b\mathbf{j}$ where $a, b \in \mathbb{C}$, we have that* $\mathrm{Hopf}_{\mathbb{R}}(\mathbf{q}) = (|a|^2 - |b|^2, 2Im(a\bar{b}), 2Re(a\bar{b}))$

Proof. Writing $a = x + y\mathbf{i}$ and $b = z + w\mathbf{i}$, we have that $\mathbf{q} = a + b\mathbf{j} = x + y\mathbf{i} + z\mathbf{j} + w\mathbf{k}$. From Proposition 30, we see that $\mathrm{Hopf}_{\mathbb{R}}(\mathbf{q}) = (x^2 + y^2 - z^2 - w^2, 2(yz - xw), 2(xz + yw))$. We can easily see that $|a|^2 - |b|^2 = x^2 + y^2 - z^2 - w^2$. Meanwhile, $a\bar{b} = (x + y\mathbf{i})(z - w\mathbf{i}) = (xz + yw) + (yz - xw)\mathbf{i})$, so we see that we indeed have $\mathrm{Hopf}_{\mathbb{R}}(\mathbf{q}) = (|a|^2 - |b|^2, 2Im(a\bar{b}), 2Re(a\bar{b}))$. $\square$

Theorem 49 (From Jean-Claude Hausmann and Allen Knutson [20])**.** *Express $\vec{\mathbf{q}} \in \mathbb{H}^n$ as $\vec{\mathbf{q}} = \vec{a} + \vec{b}\mathbf{j}$ for $\vec{a}, \vec{b} \in \mathbb{C}^n$. Then $H(\vec{\mathbf{q}}) \in \mathscr{P}_3(n)$ if and only if $\|\vec{a}\| = \|\vec{b}\|$ and $\langle \vec{a}, \vec{b} \rangle = 0$.*

Proof. To be in $\mathscr{P}_3(n) \subset \mathscr{A}_3(n)$, we require that the edges sum to 0. Since the edges of $H(\vec{\mathbf{q}})$ are the vectors given by applying $\mathrm{Hopf}_{\mathbb{R}}$ to each coordinate, we have that the sum of the edges is given by:

$$\sum_{j=1}^{n}(|a_j|^2 - |b_j|^2, 2Im(a_j\overline{b_j}), 2Re(a_j\overline{b_j})).$$

The sum of the first component simplifies to $\|\vec{a}\|^2 - \|\vec{b}\|^2$, so we see that this is zero if and only if $\|\vec{a}\| = \|\vec{b}\|$. Next, recall that $\langle \vec{a}, \vec{b} \rangle = \sum_{j=1}^{n} a_j\overline{b_j}$. As such, the sum of the second component simplifies to $2Im(\langle \vec{a}, \vec{b} \rangle)$ and the sum of the third component simplifies to $2Re(\langle \vec{a}, \vec{b} \rangle)$. These are both zero if and only if $\langle \vec{a}, \vec{b} \rangle = 0$. □

Proposition 50. *Express $\vec{z} \in \mathbb{C}^n$ as $\vec{z} = \vec{a} + \vec{b}\mathbf{i}$ for $\vec{a}, \vec{b} \in \mathbb{R}^n$. Then $P(\vec{\mathbf{q}}) \in \mathscr{P}_2(n)$ if and only if $\|\vec{a}\| = \|\vec{b}\|$ and $\langle \vec{a}, \vec{b} \rangle = 0$.*

Proof. From Proposition 46, we see that P is given by coordinate-wise squaring followed by the association between a complex number and an edge vector. Looking at Proposition 43, we see that we can do this through a bit of a longer sequence by:

1. first mapping to the $1 \oplus \mathbf{j}$ plane in $\mathbb{H}$ by sending $1 \mapsto \mathbf{1}$ and $\mathbf{i} \mapsto \mathbf{j}$ and extending linearly.
2. Apply the Hopf map and associate the resulting purely imaginary quaternion with a vector in $\mathbb{R}^3$
3. We observed in Proposition 41, every edge of this polygon will lie in the xz-plane, so we lose nothing by focusing on the associated 2-dimensional vector.

From Theorem 49, we see that the polygon will be closed if and only if the quaternionic vector $\vec{\mathbf{q}} = \vec{a} + \vec{b}\mathbf{j}$ satisfies: (1) $\|\vec{a}\| = \|\vec{b}\|$ and $\langle \vec{a}, \vec{b} \rangle = 0$ (as complex vectors). Since the both the norms and the orthogonality of the vectors $\vec{a}$ and $\vec{b}$ are the same whether viewed as a real or complex vector, the result follows directly from the previous theorem. □

Definition 51. Fix $\mathbb{F}$ to be either the field $\mathbb{R}$ or $\mathbb{C}$. We define a k-frame on $\mathbb{F}^n$ to be an ordered k-tuple of linerally independent vectors in $\mathbb{F}^n$. Likewise, we define an orthonormal k-frame to be a k-frame whose vectors have unit length and are pair-wise orthogonal.

Definition 52. Fix $\mathbb{F}$ to be the field $\mathbb{R}$ or $\mathbb{C}$. The Stiefel Manifold $V_k(\mathbb{F}^n)$ is the set of all orthonormal k-frames on $\mathbb{F}^n$.

Example.

- $V_{n+1}(\mathbb{F}^n) \cong \varnothing$
- $V_1(\mathbb{R}^n) \cong S^{n-1}$
- $V_1(\mathbb{C}^n) \cong S^{2n-1}$
- $V_n(\mathbb{R}^n) \cong O(n)$, the set of orthogonal matrices of size $n \times n$
- $V_n(\mathbb{C}^n) \cong U(n)$, the set of unitary matrices of size $n \times n$

Theorem 53.

1. *The preimage of $Pol_3(n)$ under the map $H : \mathbb{H}^n \to \mathscr{A}_3(n)$ is precisely the subset of $S^{4n-1}(\sqrt{2}) \subset \mathbb{H}^n$ that corresponds to the embedding of $V_2(\mathbb{C}^n)$ which sends the orthonormal 2-frame $\{\vec{a}, \vec{b}\}$ to the quaternionic vector $\vec{a}\mathbf{1} + \vec{b}\mathbf{j}$.*

2. *The preimage of $\mathscr{P}_3(n)$ is precisely the subset of $\mathbb{H}^n$ that corresponds to the infinite cone $CV_2(\mathbb{C}^n) = V_2(\mathbb{C}^n) \times [0, \infty)/\{V_2(\mathbb{C}^n) \times \{\vec{0}\}\}$, where the interval coordinate parametrizes the square root of the total length of the polygon.*

3. *The preimage of $Pol_2(n)$ under the map $P : \mathbb{C}^n \to \mathscr{A}_2(n)$ is precisely the subset of $S^{2n-1}(\sqrt{2}) \subset \mathbb{C}^n$ that corresponds to the embedding of $V_2(\mathbb{R}^n)$ which sends the orthonormal 2-frame $\{\vec{a}, \vec{b}\}$ to the complex vector $\vec{a} + \vec{b}\mathbf{i}$.*

4. *The preimage of $\mathscr{P}_2(n)$ is precisely the subset of $\mathbb{H}^n$ that corresponds to the infinite cone $CV_2(\mathbb{R}^n) = V_2(\mathbb{R}^n) \times [0, \infty)/\{V_2(\mathbb{R}^n) \times \{\vec{0}\}\}$, where the interval coordinate parametrizes the square root of the total length of the polygon.*

Moreover, for any polygon $P \in \mathscr{P}_2(n)$, the inverse image of P, $F_2^{-1}(\{p\})$ consists of 2^{n-m} points where m is the number of length 0 edges of p. In the spatial case, the inverse image of a polygon $P \in \mathscr{P}_3(n)$, $F_3^{-1}(\{P\})$ is an embedded copy of an $(n-m)$-dimensional torus, where again we have that m is the number of edges of p that have length 0.

Proof.

1. First, Proposition 39 tells us that the preimage of $Pol_3(n) \subset Arm_3(n)$ must be contained in the sphere $S^{4n-1}(\sqrt{2})$. Next, Theorem 49 tells us that a quaternionic vector in the preimage of $Pol_3(n)$ if and only if it can be expressed as $\vec{\mathbf{q}} = \vec{a}\mathbf{1} + \vec{b}\mathbf{j}$ for two complex vectors $\vec{a}, \vec{b}$ that are orthogonal and have the same length. From this we see that the two complex vectors must be unit length, and hence form an orthonormal 2-frame.

2. To see this, we must specify that the embedding of the cone sends the pair of an orthonormal 2-frame and a real number, $(\{\vec{a}, \vec{b}\}, r) \in V_2(\mathbb{C}^n) \times \mathbb{R}$ to the quaternionic vector $r\vec{a}\mathbf{1} + r\vec{b}\mathbf{j}$. Combining this with Theorem 49 and Lemma 40, the result follows.

3. This follows analogously to the case of $Pol_3(n)$, where we appeal to Proposition 47 to see that the preimage of $Pol_2(n)$ must be contained in the sphere $S^{2n-1}(\sqrt{2})$, and to Proposition 50 to see that the the embedded Stiefel manifold is precisely the preimage.

4. This follows directly from Proposition 50 and the planar analogy to Lemma 40 argued in the proof of Proposition 47.

Next, consider the structure of the preimage of a single polygon P. In the planar case, we see from Proposition 46 that the the j^{th} edge of P is the image under the squaring map

$(a_j, b_j) \mapsto (a_j^2 - b_j^2, 2a_jb_j)$. This map is a double cover of the complex plane except at the origin. More, the two individual preimages of a particular edge are related by negation. Replacing a_j, and b_j with $-a_j$ and $-b_j$ in the orthonormal 2-frame $\{\vec{a}, \vec{b}\}$ will not change the norms of the vectors, nor the orthogonality. Since we have this choice for each edge of positive length, we have precisely 2^{n-m} points that map to the polygon P.

Next, consider a spatial polygon Q. Here, we see from Proposition 36 that, provided it has positive length, the j^{th} edge of P is the image of precisely two points $\mathbf{q_j}, -\mathbf{q_j} \in \mathbb{IH}$, under the map $\text{Hopf}_{\mathbb{R}}|_{\mathbb{IH}}$. We then see in Proposition 38 that the other quaternions that map to this edge under the map $\text{Hopf}_{\mathbb{R}}$ will be precisely those that are related to $\pm q_j$ by left-multiplication by a unit length complex number. This shows us that the preimage of the j^{th} edge is an embedded copy of S^1.

Supposing that we have $\mathbf{q_j} = a_j\mathbf{1} + b_j\mathbf{j}$ and a unit length complex number z, we now need to examine the change in $\langle \vec{a}, \vec{b} \rangle$ and the norms of $\vec{a}$ and $\vec{b}$ when we replace a_j and b_j with za_j and zb_j. Since the inner product on $\mathbb{C}^n$ is defined as $\langle \vec{a}, \vec{b} \rangle = \sum_{i=1}^{n} a_i\overline{b_i}$, we can see that $(za_j)\overline{(zb_j)} = z\overline{z}a_j\overline{b_j} = a_jb_j$. Likewise, the norm is computed from the norms of the entries, which are unchanged by multiplication by a complex number of unit magnitude. This tells us that replacing the j^{th} entries of the orthonormal 2-frame $\{\vec{a}, \vec{b}\}$ with za_j and zb_j produces another orthonormal 2-frame. Putting this all together, we obtain the structure of an embedded $(n-m)$-dimensional torus in $\mathbb{H}^n$, with each product circle corresponding to a circle in a distinct coordinate. $\square$

Proposition 54. *The Stiefel manifold $V_k(\mathbb{R}^n)$ can be identified as the quotient of $O(n)$ by a subgroup isomorphic to $O(n-k)$.*

Proof. First we show that $O(n)$ acts transitively on $V_k(\mathbb{R}^n)$. To do this, we will first establish some notation:

- For a given subspace $V \subset \mathbb{R}^n$, denote by $V^\perp$ the orthogonal complement of V, $V^\perp = \{\vec{w} \in \mathbb{R}^n : \langle \vec{w}, \vec{v} \rangle = 0, \text{ for all } \vec{v} \in V\}$.

- For a given subspace V, denote by $V_\perp$ a matrix whose columns form an orthonormal basis for V.

- For a given matrix $M = [\vec{m}_1|\vec{m}_2|\cdots|\vec{m}_k]$, denote by $C(M)$ the column space of M, $C(M) = \{\sum_{i=1}^k c_i\vec{m}_i : c_i \in \mathbb{R}^n\}$.

We now note that we may identify $V_k(\mathbb{R}^n) \cong \{A \in \mathscr{M}_{n\times k}(\mathbb{R}) | A^\intercal A = I_k\}$. Under this identification, $O(n)$ acts upon $V_k(\mathbb{R}^n)$ by left-multiplication of matrices. Let $A, B \in V_k(\mathbb{R}^n)$, and construct a matrix of the form $Q = \begin{bmatrix} B & C(B)_\perp \end{bmatrix} \begin{bmatrix} A^\intercal \\ (C(A)_\perp)^\intercal \end{bmatrix}$. Such an Q will act on A as follows:

$$\begin{aligned} QA &= \begin{bmatrix} B & C(B)_\perp \end{bmatrix} \begin{bmatrix} A^\intercal \\ (C(A)_\perp)^\intercal \end{bmatrix} A \\ &= \begin{bmatrix} B & C(B)_\perp \end{bmatrix} \begin{bmatrix} I_k \\ O \end{bmatrix} \\ &= B. \end{aligned}$$

So we see that Q takes A to B.

If a given matrix is in the stabilizer of A, then it must fix the individual columns of A and also preserves the subspace $C(A)^\perp$. Fix some orthonormal basis for $C(A)^\perp$ to construct the matrix $A_\perp = (C(A)^\perp)_\perp$, whose columns form that orthonormal basis. Using these, we can construct the matrix $C_A = \begin{bmatrix} A & A_\perp \end{bmatrix}$. Suppose that $N \in O(n-k)$ and consider the matrix $N_A = C_A \begin{bmatrix} I_k & O \\ O & N \end{bmatrix} (C_A)^\intercal = \begin{bmatrix} A & A_\perp \end{bmatrix} \begin{bmatrix} I_k & O \\ O & N \end{bmatrix} \begin{bmatrix} A^\intercal \\ (A_\perp)^\intercal \end{bmatrix}$.

This matrix will act on A as follows:

$$\begin{aligned}
N_A A &= \begin{bmatrix} A & A_\perp \end{bmatrix} \begin{bmatrix} I_k & O \\ O & N \end{bmatrix} \begin{bmatrix} A^\intercal \\ (A_\perp)^\intercal \end{bmatrix} A \\
&= \begin{bmatrix} A & A_\perp \end{bmatrix} \begin{bmatrix} I_k & O \\ O & N \end{bmatrix} \begin{bmatrix} I_k \\ O \end{bmatrix} \\
&= \begin{bmatrix} A & A_\perp \end{bmatrix} \begin{bmatrix} I_k \\ O \end{bmatrix} \\
&= A.
\end{aligned}$$

We see then that N_A is in the stabilizer of A for any $N \in O(n-k)$. Next, take some M in the stabilizer of A and look at its conjugate by C_A under this group action, $C_A^{-1} M C_A$:

$$\begin{aligned}
C_A^{-1} M C_A &= C_A^\intercal M C_A \\
&= \begin{bmatrix} A^\intercal \\ (A_\perp)^\intercal \end{bmatrix} M \begin{bmatrix} A & A_\perp \end{bmatrix} \\
&= \begin{bmatrix} A^\intercal \\ (A_\perp)^\intercal \end{bmatrix} \begin{bmatrix} MA & MA_\perp \end{bmatrix} \\
&= \begin{bmatrix} A^\intercal \\ (A_\perp)^\intercal \end{bmatrix} \begin{bmatrix} A & MA_\perp \end{bmatrix} \\
&= \begin{bmatrix} I_k & O \\ O & (A_\perp)^\intercal M A_\perp \end{bmatrix}.
\end{aligned}$$

We claim that this lower $(n-k) \times (n-k)$ block $(A_\perp)^\intercal M A_\perp$ is an element of $O(n-k)$. Since the larger matrix, $C_A^{-1} M C_A$ is in $O(n)$, its columns form an orthonormal basis for $\mathbb{R}^n$. In particular, the last $n-k$ columns form an orthonormal set comprised of vectors whose first k entries are all zero. Dropping those first k entries will then have no effect on their pairwise inner products or on their individual norms. Hence this lower block has columns that form an orthonormal basis for $\mathbb{R}^{n-k}$. This allows us to see that the lower block is indeed an element of $O(n-k)$. We have constructed an explicit isomorphism between $O(n-k)$ and the stabilizer of A, namely $N \mapsto N_A$, and may then conclude that $V_k(\mathbb{R}^n) \cong O(n)/O(n-k)$. $\square$

Proposition 55. *The Stiefel manifold $V_k(\mathbb{C}^n)$ can be identified as the quotient of $U(n)$ by a subgroup isomorphic to $U(n-k)$.*

Proof. The proof of this mirrors the proof of the previous proposition. $\square$

Definition 56. The Grassmann manifold $G_k(\mathbb{F}^n)$ is the smooth, compact manifold of k-dimensional subspaces in $\mathbb{F}^n$.

There is an $O(k)$ action on $V_k(\mathbb{R}^n)$ which rotates the basis in the plane it spans and which we can see best as multiplication of matrix $A \in V_k(\mathbb{R}^n)$ on the right by the matrix $M \in O(k)$. This action defines an equivalence relation on $V_k(\mathbb{R}^n)$, allowing us to represent $G_k(\mathbb{R}^n) \cong V_k(\mathbb{R}^n)/O(k)$. Likewise, we can realize $G_k(\mathbb{C}^n) \cong V_k(\mathbb{C}^n)/U(k)$.

1.4 Measure

The goal of this section is to develop a measure on each of our polygon spaces. We have already shown the existence of surjective maps from well-studied manifolds to our polygon spaces. We will call these manifolds the model spaces for our polygon spaces and use the surjections to induce measures on our polygon spaces. Each model space carries a specific measure we will work with will that carries some additional properties beyond a generic measure which makes it a preferable measure for use in our examination of polygon spaces.

Theorem 57 (Haar's Theorem, [7]). *Let* $(G, *)$ *be a locally compact Hausdorff topological group. The there is, up to a positive multiplicative constant, a unique countably additive, nontrivial measure* μ *on the Borel subsets of* G*, called the Haar measure, satisfying the following four properties:*

- μ *is left-translation-invariant:* $\mu(gE) = \mu(E)$ *for every* $g \in G$ *and Borel set* E*.*
- μ *is finite on every compact set.*
- μ *is outer regular on Borel sets:* $\mu(E) = \inf\{\mu(U) : E \subseteq U, U \text{ open and Borel}\}$*.*
- μ *is inner regular on open Borel sets* E*:* $\mu(E) = \sup\{\mu(K) : K \subseteq E, K \text{ compact}\}$*.*

When G is compact, $\mu(G)$ is finite, and often given complete uniqueness by adding a fifth, normalizing condition that $\mu(G) = 1$. Since the orthogonal group $O(m)$ and the unitary group $U(m)$ are compact, we will let θ_m and ω_m be the respective normalized Haar measures on $O(m)$ and $U(m)$.

Recall that the sphere S^{m-1} admits a natural Borel measure called the spherical measure, σ^{m-1}, which we will also normalize to satisfy $\sigma^{m-1}(S^{m-1}) = 1$. This measure is nicely related to the Haar measure on the orthogonal group $O(m)$:

Definition 58. Given two measurable spaces, (X_1, Σ_1) and (X_2, Σ_2) with a measurable map $f : X_1 \to X_2$ and a measure $\mu : \Sigma_1 \to [0, \infty]$, we define the pushforward of μ by f, to be the measure $f_*(\mu) : \Sigma_2 \to [0, \infty]$ for which the measure of any $B \in \Sigma_2$, $(f_*(\mu))(B)$ is defined to be $\mu\left(f^{-1}(B)\right)$.

Proposition 59 (From [28] proposition 3.2.1 pg 91). *Fix a point* $s \in S^{N-1}$*. Define* $f : O(N) \to S^{N-1}$ *by* $f(g) = gs$*. Then the spherical measure* σ^{N-1} *is the same as the pushforward of the Haar measure* θ_{N-1} *by* f*.*

Definition 60. Given that $V_2(\mathbb{R}^n) \cong O(n)/O(n-k)$ (for $0 \leq k \leq n$), we will define the invariant measure θ_k on $V_k(\mathbb{R}^n)$ to be the resultant measure obtained by normalizing the measure obtained from pushing the Haar measure on $O(n)$ forward by the quotient map.

Likewise, we will define the invariant measure ω_k on $V_k(\mathbb{C}^n)$ to be the resultant measure obtained by normalizing the measure obtained from pushing the Haar measure on $U(n)$ forward by the quotient map.

The invariant measure is so named because it is invariant under the $O(n)$ or $U(n)$ action on $V_2(\mathbb{F}^n)$ given by left-multiplication, by which we mean that for any measurable subset $A \subseteq V_2(\mathbb{R}^n)$ and orthogonal matrix $g \in O(n)$, $\theta_k(A) = \theta_k(\{ga : a \in A\})$.

Definition 61. As seen in [5], we will define the symmetric measure, μ, for the following four spaces as follows:

- For $Arm_2(n)$, let μ be the pushforward of the spherical measure on $S^{2n-1}(\sqrt{2})$, normalized so that $\mu(Arm_2(n)) = 1$.
- For $Arm_3(n)$, let μ be the pushforward of the spherical measure on $S^{4n-1}(\sqrt{2})$, normalized so that $\mu(Arm_3(n)) = 1$.
- For $Pol_2(n)$, let μ be the pushforward of the Haar measure on $V_2(\mathbb{R}^n)$, normalized so that $\mu(Pol_2(n)) = 1$.
- For $Pol_3(n)$, let μ be the pushforward of the Haar measure on $V_2(\mathbb{C}^n)$, normalized so that $\mu(Pol_3(n)) = 1$.

A more intuitive way to think of an edge is as a length and direction. Direction of course only really makes sense for edges with non-zero length.

Definition 62. Given a non-empty subset $I \subseteq \{1, 2, ..., n\}$, define
$ZA_d^n(I) = \{\{\vec{e_1}, \ldots, \vec{e_n}\} \in Arm_d(n) : \vec{e_i} = \vec{0} \text{ for all } i \in I\}$. Likewise, define
$ZP_d^n(I) = ZA_d^n(I) \bigcap Pol_d(n)$.

Proposition 63. *The set $ZA^n_d(I)$ is a submanifold of $Arm_d(n)$ that is diffeomorphic to $Arm_d(n-m)$, where $m = \#(I)$.*

Proof. For notational convenience, order $I = \{i_1, i_2, \ldots, i_m\}$ so that $i_1 < i_2 < \cdots < i_m$, and order the set $J = \{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_m\}$ so that $j_1 < j_2 < \cdots < j_{n-m}$. We can then map $Arm_d(n-m)$ to $ZA^n_d(I) \subset Arm_d(n)$ by sending the arm $\{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_{n-m}}\}$ to the arm $\{\vec{f_1}, \vec{f_2}, \ldots, \vec{f_n}\}$ where $\vec{f_i} = \vec{0}$ for each $i \in I$, and $\vec{f_{j_k}} = \vec{e_k}$ for $k \in \{1, 2, \ldots, n-m\}$. Recalling that $Arm_d(n)$ is a subset of $\mathbb{R}^{dn}$, it is not hard to see this as the usual inclusion map of one such subset of $\mathbb{R}^{d(n-m)}$ into a subset of a subspace of $\mathbb{R}^{dn}$. □

Proposition 64. *Let $Z(d, n)$ denote the union of $ZA^n_d(I)$ taken over all subsets $I \subset \{1, 2, \ldots, n\}$ with $0 < \#(I) < n$. Then $\mu(Z(d, n)) = 0$.*

Proof. Since we have seen that each $ZA^n_d(I)$ is a submanifold of $Arm_d(n)$ of strictly lower dimension, we know that $\mu(ZA^n_d(I)) = 0$. Since $Z(d, n)$ is a finite union of measure zero sets, it too has measure zero. □

The set $Arm_d(n) \setminus Z(d, n)$ will then have a probability measure inherited from $Arm_d(n)$. No polygon in $Arm_d(n) \setminus Z(d, n)$ will have an edge of length zero, so we may define a pair of maps:

$$\Theta : Arm_d(n) \setminus Z(d, n) \to= \underbrace{S^{d-1} \times S^{d-1} \times \cdots \times S^{d-1}}_{n}$$

$$\Psi : Arm_d(n) \setminus Z(d, n) \to \Delta^{n-1} \subset \mathbb{R}^n$$

Here, Θ sends an arm to the list of the directions of each edge, Ψ sends an arm to the list of the lengths of each edge, and Δ^{n-1} is the scaled $(n-1)$-simplex given by
$\Delta^{n-1} = \{(\psi_1, \ldots, \psi_n) \in \mathbb{R}^n : \sum_{i=1}^{n} \psi_i = 2, \text{ and } \psi_i \geq 0 \text{ for all } i\}$. These maps will help us understand the distribution of the directions and lengths of the edge vectors.

Definition 65. Define the measure μ_Z on $\Delta^{n-1} \times \left(S^{d-1}\right)^n$ to be the pushforward of μ under the map (Ψ, Θ).

Proposition 66. *As a pushforward of the spherical measure, μ_Z is invariant under the S_n action which permutes simultaneously the order of the coordinates in the simplex and the product of spheres. It is also invariant under the action of rotating the spheres.*

Proof. For $d = 2$, set $k = 2$, and for $d = 3$, set $k = 4$. Consider a block matrix in the form of a permutation matrix where 1 stands for I_k and 0 stand for O_k. This matrix will permute the order of the n k-tuples of a point in the sphere S^{kn-1}. Notice that each row and each column of this matrix consists of a single non-zero entry. Moreover, that non-zero entry is always 1. The product of this matrix with its transpose will be the matrix whose (i, j)-entry is the dot product of the i^{th} and j^{th} rows. Since the only non-zero entry of the i^{th} row is the 1 located in a particular column, and likewise the only non-zero entry of the j^{th} row is the 1 located in a particular column, we see that this (i, j)-entry will be zero whenever $i \neq j$. Further, as the only non-zero entry is 1, and it occurs exactly once per row, the (i, i)-entry of the product will be 1. Hence, the permutation matrix is orthogonal, as its inverse is equal to its transpose. Proposition 66 tells us that the the spherical measure is invariant under the $O(kn)$ action. So the pushforward to $\Delta^{n-1} \times Z_d^n$ must be invariant under this action as well. Specifically, this action becomes the S_n action permuting simultaneously the coordinates in the simplex and the product of spheres.

Next consider a diagonal block matrix of the form:

$$\begin{bmatrix} R_1 & 0 & \dots & 0 \\ 0 & R_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_n \end{bmatrix},$$

where $R_i \in O(k)$. This matrix will apply the transformation of R_i to each of the i^{th} k-tuple.

In particular, $O(k)$ acts transitively on the set of k-tuples that share a common norm. Since the columns of this type of matrix are orthonormal, it is an orthogonal matrix. Hence we may use Proposition 59 to conclude that μ_Z will be invariant under this action of rotating the individual spheres.

$\square$

Note here, that the choice of R_i to obtain a desired rotation is not difficult to make. For $d = 2$, if we wish to rotate the S^1 by an angle θ, we select R_i to be the rotation in $O(2)$ of angle $\frac{\theta}{2}$. This follows from Proposition 46. For $d = 3$, we can see from Proposition 36 and Proposition 38 that we can easily find the entire preimage of a given edge vector. We then need only select one of the transformations in $O(4)$ that carries the current preimage, a great circle in S^3, to the desired one.

Corollary 67. *The distributions of directions of edges of polygonal arms in $Arm_d(n)$ are independently distributed uniformly on S^{d-1}.*

Proof. The invariance under permutations shows us that they are independently distributed, and, as seen on page 92 of [28] while discussing Proposition 59, invariance under rotations is enough to identify the measure as the uniform measure on S^{d-1}. $\square$

1.5 Expectations

While the sphere and the symmetric measure are already nice to work with, it will sometimes be convenient to drop the total length 2 condition to work with $\mathscr{A}_d(n)$. As we've seen in Propositions 39 and 47, we can model spaces $\mathscr{A}_d(n)$ by $\mathbb{C}^n$ and $\mathbb{H}^n$. Rather than work with the usual Lebesgue measure on $\mathbb{C}^n$ and $\mathbb{H}^n$, it would be nice to have a measure more closely related to the symmetic measure we have already established by pushing forward the spherical measure. Notice that we can think of $\mathbb{C}^n$ and $\mathbb{H}^n$ as the cones $\mathbb{C}^n = CS^{2n-1}$ and

$\mathbb{H}^n = CS^{4n-1}$. In doing so, it will make sense to have a product measure $\upsilon \times \sigma^{4n-1}$, for some finite measure υ on the interval $[0, \infty)$, as this will preserve the spherical symmetry.

Definition 68 (From [6])**.** The Hopf-Gaussian measure, γ_d^n, on $\mathscr{A}_d(n) \cong \mathbb{R}_{\geq 0} \times Arm_d(n)$ is defined by

$$\gamma_3^n(U) = \int_{\mathrm{Hopf}^{-1}(U)} \mathrm{d}\gamma^{4n}, \text{ for } U \subset \mathscr{A}_3(n),$$

where γ^{4n} is the standard Gaussian measure on $\mathbb{H}^n = \mathbb{R}^{4n}$, and

$$\gamma_2^n(U) = \int_{\mathrm{Hopf}^{-1}(U)} \mathrm{d}\gamma^{2n}, \text{ for } U \subset \mathscr{A}_2(n),$$

where γ^{2n} is the standard Gaussian measure on $\mathbb{C}^n = \mathbb{R}^{2n}$.

Proposition 69 (From [6])**.** *γ_d^n is the product measure $\gamma_d^n = \chi^2_{2^{d-1}n} \times \mu_d$, where $\chi^2_{2^{d-1}n}$ is the chi-squared distribution with $2^{d-1}n$ degrees of freedom on the interval $[0, \infty)$ and μ_d is the symmetric measure on $Arm_d(n)$.*

We now have two different ways to compute expectations of scalar-valued functions defined on arm space, so the first question one might ask is "When do they agree?"

Obviously not all the time, as for instance the expected total length of a polygonal arm of total length 2 is 2, while the expected length of a a polygonal arm sampled under the Hopf-Gaussian measure would be the expectation of the $\chi^2_{2^{d-1}n}$ parameter—namely $2^{d-1}n$. However, Cantarella et. al. showed the following:

Theorem 70 (From [6])**.** *Suppose that $F : \mathscr{A}_d(n) \to \mathbb{R}$ is a scale-invariant function. Then the expected value of F over $\mathscr{A}_d(n)$ with respect to the Hopf-Gaussian measure H is the same as the expected value of F over $Arm_d(n)$ with respect to the symmetric measure μ.*

Proof. The proof of this theorem can be found in [6]. $\square$

Proposition 71 (From [6])**.** *The probability distribution of the vector $\vec{r}$ joining the ends of a k-edge sub-arm of an arm in $\mathscr{A}_3(n)$ with the Hopf-Gaussian measure is spherically symmetric in $\mathbb{R}^3$ and given by the explicit formula:*

$$GS_k(\vec{r})\mathrm{d}Vol_{\vec{r}} = \frac{r^{k-3/2}K_{k-3/2}(r/2)}{2^{2k+2}\pi^{3/2}\Gamma(k)}\mathrm{d}Vol_{\vec{r}},$$

where $r = |\vec{r}|$, and $K_\nu(z)$ is the modified Bessel function of the second kind.

Proof. The proof of this proposition can be found in [6]. □

Bessel's differential equation, $z^2\dfrac{d^2w}{dz^2} + z\dfrac{dw}{dz} + (z^2 - \alpha^2)w = 0$, being a second-order differential equation has two linearly independent solutions. We call the first the Bessel function of the first kind, $J_\alpha(z) = \left(\frac{1}{2}z\right)^\alpha \sum_{k=0}^\infty (-1)^k \dfrac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(\alpha+k+1)}$. When $\alpha \in \mathbb{Z}$, this is an entire function. Otherwise, it has a branch point at $z = 0$. The second solution, the Bessel function of the second kind, is $Y_\alpha(z) = \dfrac{J_\alpha(z)\cos(\alpha\pi) - J_{-\alpha}(z)}{\sin(\alpha\pi)}$. When $\alpha \in \mathbb{Z}$ this is replaced by the limiting value $Y_n(z) = \dfrac{1}{\pi}\left.\dfrac{\partial J_\alpha(z)}{\partial\alpha}\right|_{\alpha=n} + \dfrac{(-1)^n}{\pi}\left.\dfrac{\partial J_\alpha(z)}{\partial\alpha}\right|_{\alpha=-n}$ (Section 10.2 of [32]).

The modified Bessel functions arise from the solution to the modified Bessel's equation obtained by replacing z with $\pm\mathbf{i}z$. Here, the first solution is given by $I_\alpha(z) = \left(\frac{1}{2}z\right)^\alpha \sum_{k=0}^\infty \dfrac{(\frac{1}{4}z^2)^k}{k!\Gamma(\alpha+k+1}$. The modified Bessel function of the second kind, $K_\alpha(z) = \frac{1}{2}\pi\dfrac{I_{-\alpha}(z) - I_\alpha(z)}{\sin(\alpha\pi)}$ has the defining property that it is asymptotic to $\sqrt{\pi/(2z)}e^{-z}$ as $z \to \infty$. As with the (unmodified) Bessel function of the second kind, when α is an integer we must take limiting values: $K_n(z) = \dfrac{(-1)^{n-1}}{2}\left(\left.\dfrac{\partial I_\alpha(z)}{\partial\alpha}\right|_{\alpha=n} + \left.\dfrac{\partial I_\alpha(z)}{\partial\alpha}\right|_{\alpha=-n}\right)$. Both of the modified Bessel functions are real-valued when $\alpha \in \mathbb{R}$ and are invariant under the involution $\alpha \mapsto -\alpha$. [32] (Section 10.25).

We will now build the analogous probability distribution function of the vector $\vec{r}$ joining the ends of a k-edge sub-arm of an arm in $\mathscr{A}_2(n)$ with the Hopf-Gaussian measure.

Proposition 72. *The probability distribution of the vector $\vec{r}$ joining the ends of a k-edge sub-arm of an arm in $\mathscr{A}_2(n)$ with the Hopf-Gaussian measure is spherically symmetric in $\mathbb{R}^2$ and given by the formula:*

$$GP_k(\vec{r})\,\mathrm{d}\,Vol_{\vec{r}} = \frac{r^{(k/2)-1}K_{(k/2)-1}\left(\dfrac{r}{2}\right)}{2^{k+1}\pi\Gamma\left(\dfrac{k}{2}\right)}\,\mathrm{d}\,Vol_{\vec{r}},$$

where $r = |\vec{r}|$.

Proof. Suppose that $\vec{z} = (z_1, \ldots, z_n) \in \mathbb{C}^n$ is sampled from the standard Gaussian distribution. This distribution is the same as the multivariate Gaussian on $\mathbb{R}^{2n}$ with mean 0 and covariance I_{2n}. Then the failure-to-close vector for the k-edge arm $\mathrm{Hopf}(\vec{z}) \in \mathscr{A}_2(n)$ is

$$\sum \mathrm{Hopf}(z_i) = \sum z_i^2 = \sum (a_i^2 - b_i^2, 2a_ib_i).$$

Since the a_i and b_i are chosen from standard Gaussian distributions, the distribution of the projection of the failure-to-close vector onto the first coordinate is the difference of two chi-squared distributions, each with k degrees of freedom. In [25], Johnson et al. showed that the density of a distribution of the form $Z = \chi_\nu^2\sigma_1^2 - \chi_\nu^2\sigma_2^2$, is given by:

$$q(r) = \frac{|1-c|^{m+1/2}|r|^m e^{-cr/b}}{\sqrt{\pi}2^m b^{m+1}\Gamma(m+\frac{1}{2})}K_m\left(\frac{r}{2}\right),$$

where, c is $\dfrac{\sigma_1^2-\sigma_2^2}{\sigma_1^2+\sigma_2^2}$, b is $4\dfrac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$, and m is $2\nu+1$. Here, we use the notation $\Gamma(z)$ to denote the gamma function, $\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}\,\mathrm{d}t$, known to be a meromoprhic function with no zeros and simple poles of residue $\dfrac{(-1)^n}{n!}$ at $z = -n$ (5.2.1 of [32]). In our case, this becomes:

$$\frac{2^{-k}(\sqrt{r})^{k-1}}{\sqrt{\pi}\Gamma(\frac{k}{2})}K_{\frac{k-1}{2}}\left(\frac{r}{2}\right).$$

In [29], Lord showed that when a spherically symmetric distribution on $\mathbb{R}^2$ has the property that the projection of $p(\vec{r})$ to any radial line has a probability distribution function of $p_1(\vec{r})$, then its probability distribution function is given by

$$p(r) = \frac{-1}{\pi} \int_r^\infty p_1'(t) \frac{1}{\sqrt{t^2 - r^2}}\, \mathrm{d}t.$$

We will now differentiate the distribution above, $p_1(t) = \dfrac{2^{-k}}{\sqrt{\pi}\Gamma\left(\frac{k}{2}\right)} t^{\frac{k-1}{2}} K_{\frac{k-1}{2}}\left(\frac{t}{2}\right)$:

$$\begin{aligned}
p_1'(t) &= \frac{2^{-k}}{\sqrt{\pi}\Gamma(\frac{k}{2})} \left(\left(\frac{d}{dt} t^{\frac{k-1}{2}} \right) K_{\frac{k-1}{2}} \left(\frac{t}{2}\right) + t^{\frac{k-1}{2}} \left(\frac{d}{dt} K_{\frac{k-1}{2}} \left(\frac{t}{2}\right) \right) \right) && (1.1) \\
&= \frac{2^{-k}}{\sqrt{\pi}\Gamma(\frac{k}{2})} \left(\frac{k-1}{2} t^{\frac{k-3}{2}} K_{\frac{k-1}{2}} \left(\frac{t}{2}\right) + t^{\frac{k-1}{2}} \left(-\frac{1}{2} K_{\frac{k-3}{2}} \left(\frac{t}{2}\right) - \frac{1}{2} K_{\frac{k+1}{2}} \left(\frac{t}{2}\right) \right) \frac{1}{2} \right) && (1.2) \\
&= \frac{2^{-k-1} t^{\frac{k-3}{2}}}{\sqrt{\pi}\Gamma(\frac{k}{2})} \left((k-1) K_{\frac{k-1}{2}} \left(\frac{t}{2}\right) - \frac{t}{2} K_{\frac{k-3}{2}} \left(\frac{t}{2}\right) - \frac{t}{2} K_{\frac{k+1}{2}} \left(\frac{t}{2}\right) \right) && (1.3) \\
&= \frac{2^{-k-1} t^{\frac{k-3}{2}}}{\sqrt{\pi}\Gamma(\frac{k}{2})} \left(\frac{2(\frac{k-1}{2})}{t/2} \left(\frac{t}{2}\right) K_{\frac{k-1}{2}} \left(\frac{t}{2}\right) - \frac{t}{2} K_{\frac{k-3}{2}} \left(\frac{t}{2}\right) - \frac{t}{2} K_{\frac{k+1}{2}} \left(\frac{t}{2}\right) \right) && (1.4) \\
&= \frac{2^{-k-1} t^{\frac{k-3}{2}}}{\sqrt{\pi}\Gamma(\frac{k}{2})} \left(\frac{t}{2} \left(K_{\frac{k+1}{2}} \left(\frac{t}{2}\right) - K_{\frac{k-3}{2}} \left(\frac{t}{2}\right) \right) - \frac{t}{2} K_{\frac{k-3}{2}} \left(\frac{t}{2}\right) - \frac{t}{2} K_{\frac{k+1}{2}} \left(\frac{t}{2}\right) \right) && (1.5) \\
&= -\frac{2^{-k-1} t^{\frac{k-1}{2}}}{\sqrt{\pi}\Gamma(\frac{k}{2})} K_{\frac{k-3}{2}} \left(\frac{t}{2}\right) && (1.6)
\end{aligned}$$

In line 1.1, we apply the product rule. In line 1.2, we apply the power rule, and a combination of the derivative identity 10.29.1.ii from the Digital Library of Mathematical Functions [32] with the chain rule. In line 1.3, we simplify by factoring out $\dfrac{r^{\frac{k-3}{2}}}{2}$. In line 1.4, we re-write the coefficient of the first modified Bessel function in preparation to apply the recurrence relation 10.29.1.i from [32] in line 1.5. Finally, in line 1.6 we simplify the expression by collecting like terms.

We now need to compute the integral

$$p(r) = \frac{1}{2^{k+1}\pi^{\frac{3}{2}}\Gamma\left(\frac{k}{2}\right)} \int_r^\infty t^{\frac{k-1}{2}} K_{\frac{k-3}{2}}\left(\frac{t}{2}\right) \frac{1}{\sqrt{t^2-r^2}} \,\mathrm{d}t.$$

Using the following lemma (Lemma 73), with $m = \frac{k-1}{2}$, $b = \frac{1}{2}$, and $p = -\frac{1}{2}$, we see that this evaluates to:

$$\begin{aligned} p(r) &= \frac{1}{2^{k+1}\pi^{\frac{3}{2}}\Gamma\left(\frac{k}{2}\right)} r^{\frac{k-1}{2}-\frac{1}{2}} K_{\frac{k-1}{2}-\frac{1}{2}}\left(\frac{r}{2}\right)\Gamma\left(1-\frac{1}{2}\right) \\ &= \frac{1}{2^{k+1}\pi\Gamma\left(\frac{k}{2}\right)} r^{\frac{k}{2}-1} K_{\frac{k}{2}-1}\left(\frac{r}{2}\right) \end{aligned}$$

□

Lemma 73.

$$\int_r^\infty t^m\, K_{m-1}(bt)\,(t^2-r^2)^p \,\mathrm{d}t = 2^p\, b^{-(p+1)}\, r^{m+p}\, K_{m+p}(br)\,\Gamma(1+p),$$

provided that $b > 0$*,* $r > 0$*,* $m > -1$*, and* $p > -1$*.*

Proof. First, we make the substitution $s = \dfrac{t^2}{r^2}$:

$$\frac{1}{2} r^{2p+m+1} \int_1^\infty s^{\frac{m-1}{2}} K_{m-1}\left(br\sqrt{s}\right)(s-1)^p \,\mathrm{d}s.$$

Next, we apply the integral identity 6.592.12 from [19]:

$$\int_1^\infty x^{-\frac{1}{2}\nu}(x-1)^{\mu-1}K_\nu(a\sqrt{x})\mathrm{d}x = \Gamma(\mu)2^\mu a^{-\mu}K_{\nu-\mu}(a),$$

which holds for any $a, \nu, \mu \in \mathbb{C}$ such that $\mathrm{Re}(a) > 0$, and $\mathrm{Re}(\mu) > 0$. We have also observed that $K_\nu(a) = K_{-\nu}(a)$. This allows us to evaluate the integral, as we have $\nu = 1 - m$, $a = br > 0$, and $\mu = p + 1 > 0$.

$$\frac{1}{2} r^{2p+m+1} \Gamma(p+1) 2^{p+1} (br)^{-p-1} K_{(1-m)-(p+1)}(br) = 2^p b^{-(p+1)} r^{m+p} K_{(m+p)}(br) \Gamma(p+1).$$

□

Now that we have the probability distribution function we may use this to set up an integral for the expected value over $\mathscr{A}_2(n)$ of any function of a sequence of consecutive edges.

Proposition 74. *The codimension 3 Hausdorff measure of $\mathscr{P}_3(n)$ in $\mathscr{A}_3(n)$ is the value of $GS_n(\vec{0})$ which is given by $CS_n = \dfrac{\Gamma(n - \frac{3}{2})}{64\sqrt{\pi}\Gamma(n)}$.*

The codimension 2 Hausdorff measure of $\mathscr{P}_2(n)$ in $\mathscr{A}_2(n)$ is the value of $GP_n(\vec{0})$ which is given by $CP_n = \dfrac{1}{8(n-2)\pi}$.

Proof. The proof of the first claim is given in [6], so we will only give the proof of the second claim.

We saw in Proposition 72 that $GP_n(\vec{r}) = \dfrac{1}{2^{n+1}\pi\Gamma\left(\frac{n}{2}\right)} r^{\frac{n}{2}-1} K_{\frac{n}{2}-1}\left(\frac{r}{2}\right)$. Since the Bessel function has a pole at 0, we will need to take the limiting value. From 10.30.2 of [32], we know that, for fixed ν, as $r \to 0$, $K_\nu(r) \sim \frac{1}{2}\Gamma(\nu)\left(\frac{1}{2}r\right)^{-\nu}$. This tells us that:

$$\begin{aligned}
\lim_{\|r\| \to 0^+} GP_n(\vec{r}) &= \lim_{\|r\| \to 0^+} \frac{1}{2^{n+1}\pi\Gamma\left(\frac{n}{2}\right)} r^{\frac{n}{2}-1} K_{\frac{n}{2}-1}\left(\frac{r}{2}\right) \\
&= \lim_{\|r\| \to 0^+} \frac{1}{2^{n+2}\pi\Gamma\left(\frac{n}{2}\right)} r^{\frac{n}{2}-1} \Gamma\left(\frac{n}{2} - 1\right)\left(\frac{1}{4}r\right)^{-(\frac{n}{2}-1)} \\
&= \lim_{\|r\| \to 0^+} \frac{1}{2^4\pi} \frac{\Gamma\left(\frac{n}{2} - 1\right)}{\Gamma\left(\frac{n}{2}\right)} \\
&= \frac{1}{2^4\pi} \frac{\Gamma\left(\frac{n}{2} - 1\right)}{\left(\frac{n}{2} - 1\right)\Gamma\left(\frac{n}{2} - 1\right)} \\
&= \frac{1}{8\pi\,(n-2)}.
\end{aligned}$$

$\square$

Theorem 75. *Let F be a function defined on k consecutive edges. Set $s = r_1 + r_2 + \cdots + r_k$ and $z = |\vec{e_1} + \vec{e_2} + \cdots + \vec{e_k}|$. Then the expectation of F over $\mathscr{P}_2(n)$ is given by the integral:*

$$\frac{n-2}{2^{n+k-2}\pi^k\Gamma(\frac{n-k}{2})}\underbrace{\int_0^{2\pi}\int_0^{\infty}\cdots\int_0^{2\pi}\int_0^{\infty}}_{k} e^s K_{\frac{n-k}{2}-1}\left(\frac{z}{2}\right) z^{\frac{n-k}{2}-1} F(\vec{e_1},\ldots,\vec{e_k})\mathrm{d}r_1\mathrm{d}\theta_1\ldots\mathrm{d}r_k\theta_k.$$

Proof. For the consecutive sequence of edges $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}$ to belong to a closed polygon, the remaining $n-k$ edge vectors must sum to $-(\vec{e_1} + \vec{e_2} + \cdots + \vec{e_k})$. As explained in section 5.1 of [12], the probability distribution must then be given by:

$$P(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}) = \frac{GP_1(\vec{e_1})GP_1(\vec{e_2})\ldots GP_1(\vec{e_k})GP_{n-k}(-\vec{e_1} - \vec{e_2} - \cdots - \vec{e_k})}{CP_n}.$$

Writing this in polar coordinates, $\vec{e_i} = (r_i, \theta_i)$, the result then follows from simplifying the expectation integral and volume form as written in these coordinates. $\square$

Corollary 76. *Let $f(\vec{e_1}, \vec{e_2})$ be a function defined on a pair of consecutive edges. Then the expectation of f over $\mathscr{P}_2(n)$ is given by the integral:*

$$\frac{n-2}{2^n\pi^2\Gamma\left(\frac{n}{2}-1\right)}\int_0^{2\pi}\int_0^{2\pi}\int_0^{\infty}\int_0^{\infty} e^{-\frac{1}{2}(r_1+r_2)} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-2} f(\vec{e_1}, \vec{e_2})\mathrm{d}r_1\mathrm{d}r_2\mathrm{d}\theta_1\mathrm{d}\theta_2,$$

where $z = |\vec{e_1} + \vec{e_2}| = \sqrt{r_1^2 + r_2^2 + 2r_1r_2\cos(\theta_2 - \theta_1)}$.

1.6 Curvature

In this section we will discuss a specific application of Proposition 76. First we will introduce the notion of curvature for a polygon and compare the performance between direct sampling and numerical approximation of the form given in Proposition 76.

Definition 77. Given two consecutive edges e_i and e_{i+1} of a polygonal arm in $\mathbb{R}^d$, they must lie in a plane. Define the turning angle θ_i to be the angle of rotation in that plane from the direction of e_i to the direction of e_{i+1} (with the convention that the turning angle is between 0 and π).

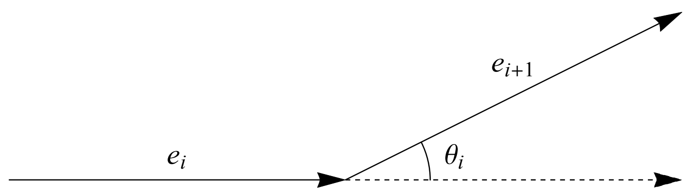

Figure 1.3: Turning Angle

Definition 78. Define the total curvature of a polygonal arm to be the sum of the turning angles.

Proposition 79. *The expectation of any turning angle for a polygonal arm sampled from* $Arm_d(n)$ *or* $\mathscr{A}_d(n)$ *is* $\frac{\pi}{2}$. *The expectation of the total curvature of a polygonal arm sampled from either* $Arm_d(n)$ *or* $\mathscr{A}_d(n)$ *is* $(n-1)\frac{\pi}{2}$.

Proof. Given the invariance of the measure under permutations of edges, we see that the expectation of θ_i will be the same as the expectation of θ_1. In the plane spanned by e_i and e_{i+1}, we can express the edges in polar coordinates as $e_i = (r_i, t_i) \in \mathbb{R}_{\geq 0} \times S^1$. In this way, the turning angle becomes simply the absolute value of the difference in their S^1 coordinates. Since, as we have previously mentioned, the directions are independent identically distributed according to the spherical measure, and the edge lengths are independently distributed, we may compute the expected value of θ_1 as:

$$\int_{S^{d-1}\times S^{d-1}} \min(|t_1 - t_2|, 2\pi - |t_1 - t_2|)(\sigma^{d-1})^2.$$

First split this into an iterated integral, and consider first the case $d = 2$:

$$\frac{1}{4\pi^2}\int_{S^1}\left(\int_{S^1}\min(|t_1-t_2|, 2\pi-|t_1-t_2|)\,\mathrm{d}t_2\right)\mathrm{d}t_1.$$

By rotating the inner S^1 by an angle t_1, we see that the expectation becomes

$$\frac{1}{4\pi^2}\int_{S^1}\left(\int_{S^1}\min(t_2, 2\pi-t_2)\,\mathrm{d}t_2\right)\mathrm{d}t_1.$$

From here, splitting the inner integral into two, we have:

$$\begin{aligned}\frac{1}{4\pi^2}\int_{S^1}\left(\int_{t_2=0}^{\pi} t_2\,\mathrm{d}t_2+\int_{t_2=\pi}^{2\pi} 2\pi - t_2\,\mathrm{d}t_2\right)\mathrm{d}t_1 &= \frac{1}{4\pi^2}\int_{S^1}\left(\frac{\pi^2}{2}+\frac{\pi^2}{2}\right)\mathrm{d}t_1\\ &= \frac{1}{4}\int_{S^1}\mathrm{d}t_1\\ &= \frac{\pi}{2}\end{aligned}$$

For the case $d = 3$, after splitting the integral up we have:

$$\frac{1}{16\pi^2}\int_{S^2}\left(\int_{s=0}^{\pi}\int_{\phi=0}^{2\pi}\min(|t_1-t_2|, 2\pi-|t_1-t_2|)\sin(s)\,\mathrm{d}\phi\mathrm{d}s\right)\mathrm{d}\sigma^2.$$

Next, use a change of variables on the inner S^2 that rotates it so that the plane spanned by $\{e_1, e_2\}$ is the xy-plane with e_1 lying on the positive x-axis.

$$\begin{aligned}\frac{1}{16\pi^2}\int_{S^2}\left(\int_{s=0}^{\pi}\int_{\phi=-\pi}^{\pi}|\phi|\sin(s)\,\mathrm{d}\phi\mathrm{d}s\right)\mathrm{d}\sigma^2 &= \frac{1}{8\pi^2}\int_{S^2}\left(\int_{s=0}^{\pi}\int_{\phi=0}^{\pi}\phi\sin(s)\,\mathrm{d}\phi\mathrm{d}s\right)\mathrm{d}\sigma^2\\ &= \frac{1}{16}\int_{S^2}\left(\int_{s=0}^{\pi}\sin(s)\,\mathrm{d}s\right)\mathrm{d}\sigma^2\\ &= \frac{1}{8}\int_{S^2}\mathrm{d}\sigma^2\\ &= \frac{\pi}{2}\end{aligned}$$

Next, recall that for integrable random variables, the expectation of a sum is the sum of the expectations, even when they are not independent (Section 5.1 of [12]). We can conclude that the expectation of total curvature for a polygonal arm sampled from either $Arm_d(n)$ or $\mathscr{A}_d(n)$ to be $n\frac{\pi}{2}$. $\square$

Since the turning angle does not depend on the orientation of the pair $\vec{e_1}$ and $\vec{e_2}$, we may use Proposition 66 to rotate the configuration so that $\vec{e_1}$ lies on the positive x-axis. By combining this with Proposition 76 and Theorem 70, we may express the expectation of the turning angle of a closed planar polygon as a 4-dimensional integral, and then immediately integrate out along the direction of $\vec{e_1}$. This gives us:

$$\theta(n) = \int_0^\pi \int_0^\infty \int_0^\infty \frac{n-2}{2^{n-2}\pi\Gamma\left(\frac{n}{2}-1\right)} e^{-\frac{1}{2}(r_1+r_2)} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-2} \psi \,\mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}\psi,$$

where again we have that $z = |\vec{e_1} + \vec{e_2}| = \sqrt{r_1^2 + r_2^2 + 2r_1 r_2 \cos(\psi)}$. To compute this numerically, it will be convenient to have a bounded domain.

Theorem 80. *The expected turning angle of a polygon in $Pol_2(n)$, $\theta(n)$ is given by the integral:*

$$\frac{n-2}{2^{n-2}\pi\Gamma\left(\frac{n}{2}-1\right)} \int_0^\pi \int_0^1 \int_0^1 \left(f(r_1, r_2, \psi) + \frac{f(\frac{1}{r_1}, r_2, \psi)}{r_1^2} + \frac{f(r_1, \frac{1}{r_2}, \psi)}{r_2^2} + \frac{f(\frac{1}{r_1}, \frac{1}{r_2}, \psi)}{r_1^2 r_2^2} \right) \psi \mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}\psi,$$

where $f(r_1, r_2, \psi) = e^{-\frac{1}{2}(r_1+r_2)} K_{\frac{n}{2}-2}\left(\frac{\sqrt{r_1^2+r_2^2+2r_1r_2\cos(\psi)}}{2}\right) (r_1^2 + r_2^2 + 2r_1 r_2 \cos(\psi))^{\frac{n-1}{2}}$.

Proof. Notice that we may write any convergent integral of the form

$$\begin{aligned}\int_0^\infty f(x)\,\mathrm{d}x &= \int_0^1 f(x)\,\mathrm{d}x + \int_1^\infty f(x)\,\mathrm{d}x \\ &= \int_0^1 f(x)\,\mathrm{d}x + \int_0^1 f\left(y^{-1}\right) y^{-2}\,\mathrm{d}y \\ &= \int_0^1 f(x) + f(x^{-1})x^{-2}\,\mathrm{d}x.\end{aligned}$$

The result then follows from applying this trick to the integral given above for both the r_1 and r_2 variables. □

While we have not yet successfully computed this integral in the general setting, we have reduced the numerical problem from computing an integral over the $2n-3$ dimensional space $Pol_2(n)$ to a more efficiently computed 3-dimensional integral over a bounded domain.

Proposition 81. *The expected turning angle of a polygon in $Pol_2(n)$, $\theta(n)$ is given by the integral:*

$$\theta(n) = \pi - \frac{n-2}{2^{n-2}\pi\Gamma\left(\frac{n}{2}-1\right)} \int_0^\infty \int_1^\infty e^{-\frac{sz}{2}} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-1} \left(\frac{2a}{\sqrt{s^2-1}}\Phi\left(a^2, 2, \frac{1}{2}\right)\right) \mathrm{d}s\mathrm{d}z,$$

where $a = s - \sqrt{s^2-1}$, and Φ is the Lerch Transcendent given by $\Phi(a,b,c) = \sum_{k=0}^{\infty} \frac{a^k}{(c+k)^b}$ (25.14.1 of [32]).

Proof. This is simply a lengthy computation and so we shall leave the full proof for the Appendix A. □

Notice, however, that the presence of this infinite sum, prevents this expression from being numerically preferable to the one we have established in Theorem 80. We include it solely as a demonstration of how involved the computation is to carry out exactly.

Proposition 82 (From [6]). *Let $\theta(n)$ and $\kappa(n)$ be the expectations over $Pol_2(n)$ of turning angle and total curvature. Then the expectation of $\theta(n)$ of a closed n-gon approaches $\frac{\pi}{2}$, and the limit $\lim_{n\to\infty} \kappa(n) - \frac{n\pi}{2} = \frac{2}{\pi}$.*

Proof. A proof of this is found in [6], but let us attempt to to see this from our integral. We can express the turning angle as:

$$\theta(n) = \frac{2^{-n}(n-2)}{\pi\Gamma\left(\frac{n}{2}-1\right)} \int_0^{2\pi}\int_0^{\infty}\int_0^{\infty} z^{\frac{n}{2}-2} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) e^{-\frac{1}{2}(r_1+r_2)} \operatorname{Min}(\psi, 2\pi-\psi)\, \mathrm{d}r_1\, \mathrm{d}r_2\, \mathrm{d}\psi,$$

where $z = \sqrt{r_1^2 + r_2^2 + 2r_1r_2\cos(\psi)}$. Let $f(r_1, r_2, \psi, n) = \frac{2^{-n}(n-2)}{\pi\Gamma\left(\frac{n}{2}-1\right)} z^{\frac{n}{2}-2} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) e^{-\frac{1}{2}(r_1+r_2)}$ and $g(\psi) = \operatorname{Min}(\psi, 2\pi-\psi)$, so that we have $\theta(n) = \int_0^{2\pi}\int_0^{\infty}\int_0^{\infty} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}r_1\, \mathrm{d}r_2\, \mathrm{d}\psi$. First, notice that $\cos(\psi) = \cos(2\pi-\psi)$, which tells us $f(r_1, r_2, \psi, n) = f(r_1, r_2, 2\pi-\psi, n)$, and that $2\pi - (2\pi - \psi) = \psi$, so $g(\psi) = g(2\pi-\psi)$. We may now split the integral with respect to ψ into:

$$\begin{aligned}
\int_0^{2\pi} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}\psi &= \int_0^{\pi} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}\psi + \int_{\pi}^{2\pi} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}\psi \\
&= \int_0^{\pi} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}\psi + \int_{\pi}^{2\pi} f(r_1, r_2, 2\pi-\psi, n) g(2\pi-\psi)\, \mathrm{d}\psi \\
&= \int_0^{\pi} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}\psi + \int_{\pi}^{0} -f(r_1, r_2, \phi, n) g(\phi)\, \mathrm{d}\phi \\
&= 2\int_0^{\pi} f(r_1, r_2, \psi, n) g(\psi)\, \mathrm{d}\psi \\
&= 2\int_0^{\pi} f(r_1, r_2, \psi, n) \psi\, \mathrm{d}\psi
\end{aligned}$$

Next, from 10.41.2 of [32] we know that for large z, $K_\nu(z) \sim \sqrt{\dfrac{\pi}{2\nu}}\left(\dfrac{ez}{2\nu}\right)^{-\nu}$, so in the limit, we have:

$$\begin{aligned}
K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) &\sim \sqrt{\frac{\pi}{2\nu}}\left(\frac{ez}{2\nu}\right)^{-\nu} \\
&= \sqrt{\frac{\pi}{n-4}}\left(\frac{e\frac{z}{2}}{n-4}\right)^{2-\frac{n}{2}} \\
&= 2^{\frac{n}{2}-1}e^{2-\frac{n}{2}}\sqrt{\pi}(n-4)^{\frac{n-5}{2}}z^{2-\frac{n}{2}}.
\end{aligned}$$

This tells us that, in the limit, we have:

$$\begin{aligned}
\theta(n) &= \frac{2^{1-n}(n-2)}{\pi\Gamma\left(\frac{n}{2}-1\right)}\int_0^\pi\int_0^\infty\int_0^\infty z^{\frac{n}{2}-2}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)e^{-\frac{1}{2}(r_1+r_2)}\psi\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}\psi \\
&= \frac{2^{1-n}(n-2)}{\pi\Gamma\left(\frac{n}{2}-1\right)}\int_0^\pi\int_0^\infty\int_0^\infty z^{\frac{n}{2}-2}2^{\frac{n}{2}-1}e^{2-\frac{n}{2}}\sqrt{\pi}(n-4)^{\frac{n-5}{2}}z^{2-\frac{n}{2}}e^{-\frac{1}{2}(r_1+r_2)}\psi\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}\psi \\
&= \frac{2^{1-n}(n-2)}{\pi\Gamma\left(\frac{n}{2}-1\right)}\int_0^\pi\int_0^\infty\int_0^\infty z^{\frac{n}{2}-2}2^{\frac{n}{2}-1}e^{2-\frac{n}{2}}\sqrt{\pi}(n-4)^{\frac{n-5}{2}}z^{2-\frac{n}{2}}e^{-\frac{1}{2}(r_1+r_2)}\psi\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}\psi \\
&= \frac{2^{-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)}{\sqrt{\pi}\Gamma\left(\frac{n}{2}-1\right)}\int_0^\pi\int_0^\infty\int_0^\infty e^{-\frac{1}{2}(r_1+r_2)}\psi\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}\psi \\
&= \frac{2^{-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)}{\sqrt{\pi}\Gamma\left(\frac{n}{2}-1\right)}\int_0^\pi\int_0^\infty\int_0^\infty e^{-\frac{1}{2}(r_1+r_2)}\psi\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}\psi \\
&= \frac{2^{-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)}{\sqrt{\pi}\Gamma\left(\frac{n}{2}-1\right)}(2\pi^2) \\
&= \frac{2^{1-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)\pi^{\frac{3}{2}}}{\Gamma\left(\frac{n}{2}-1\right)}
\end{aligned}$$

If we now use Stirling's Formula (5.11.3 of [32]), we see that

$\Gamma\left(\frac{n}{2}-1\right) \sim \sqrt{2\pi\left(\frac{n}{2}-2\right)}\left(\frac{\frac{n}{2}-2}{e}\right)^{\frac{n}{2}-2} = (2e)^{2-\frac{n}{2}}(n-4)^{\frac{n-3}{2}}\sqrt{\pi}$. This lets us see that:

$$
\begin{aligned}
\theta(n) &= \frac{2^{1-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)\pi^{\frac{3}{2}}}{\Gamma\left(\frac{n}{2}-1\right)} \\
&\sim \frac{2^{1-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)\pi^{\frac{3}{2}}}{(2e)^{2-\frac{n}{2}}(n-4)^{\frac{n-3}{2}}\sqrt{\pi}} \\
&= \left(\frac{n-2}{n-4}\right)\frac{\pi}{2} \\
&= \left(1+\frac{2}{n-4}\right)\frac{\pi}{2}.
\end{aligned}
$$

From here, we see that indeed, $\lim_{n\to\infty}\theta(n) = \frac{\pi}{2}$. From our approximations, however, it appears that:

$$
\begin{aligned}
\lim_{n\to\infty}\kappa(n) - n\frac{\pi}{2} &= \lim_{n\to\infty} n\theta(n) - n\frac{\pi}{2} \\
&= \lim_{n\to\infty} n\left(\theta(n)-\frac{\pi}{2}\right) \\
&= \lim_{n\to\infty} n\frac{\pi}{2}\left(1+\frac{2}{n-4}-1\right) \\
&= \lim_{n\to\infty} \frac{\pi}{2}\left(\frac{2n}{n-4}\right) \\
&= \pi.
\end{aligned}
$$

In the future, we plan to take better care with the asymptotic approximation of the Bessel Function, as numeric experiments indicate that our approximation at that stage leads to an error in line with:

$$
n\left(\frac{2^{1-\frac{n}{2}}e^{2-\frac{n}{2}}(n-4)^{\frac{n-5}{2}}(n-2)\pi^{\frac{3}{2}}}{\Gamma\left(\frac{n}{2}-1\right)}-\frac{\pi}{2}\right) \simeq 2.88.
$$

□

As an example, for n of the form $5 + 10m$ for $1 \leq m \leq 100$, we numerically computed the difference between the expected total curvature of a polygon in $Pol_2(n)$ and $n\frac{\pi}{2}$ by using the integral in Theorem 80 with an adaptive Monte-Carlo algorithm on one million points. We additionally recorded both the error estimates for the integral, and the time it took to compute. We then sampled polygons directly from $Pol_2(n)$ for the same amount of time, and computed the difference and the error estimates. Below are the plots that illustrate the convergence to this asymptotic bound. Notice that the maximum error from the integrals was 1.7079×10^{-4}, while the maximum error from the samples was 3.92231×10^{-3}. Indeed, of all n, the closest the gap between the errors from the integral and the sample errors occurred for $n = 15$, where the sample error was only 9.8459 times as large as the integral's error.

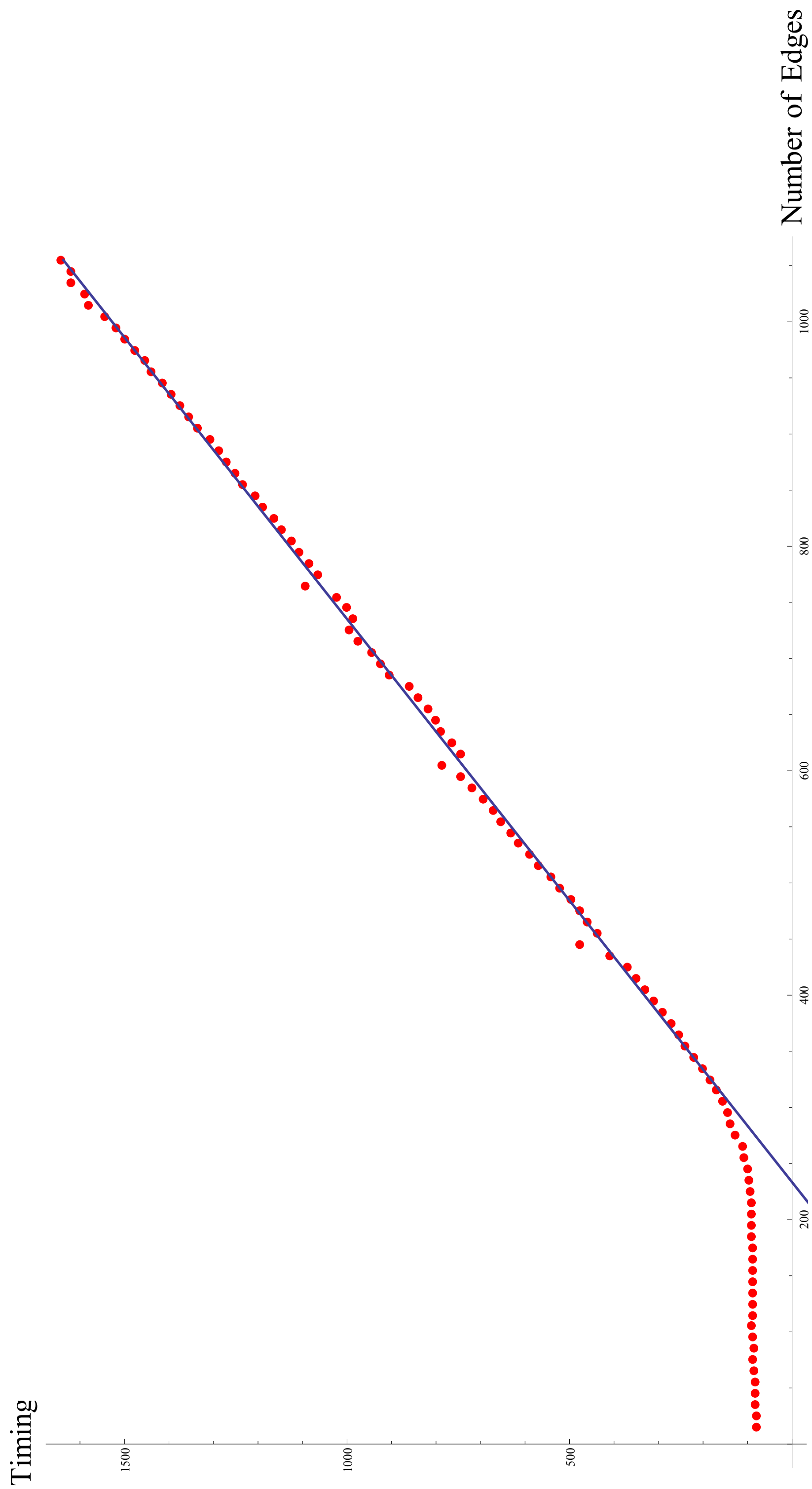

Figure 1.4: Plot of Timing

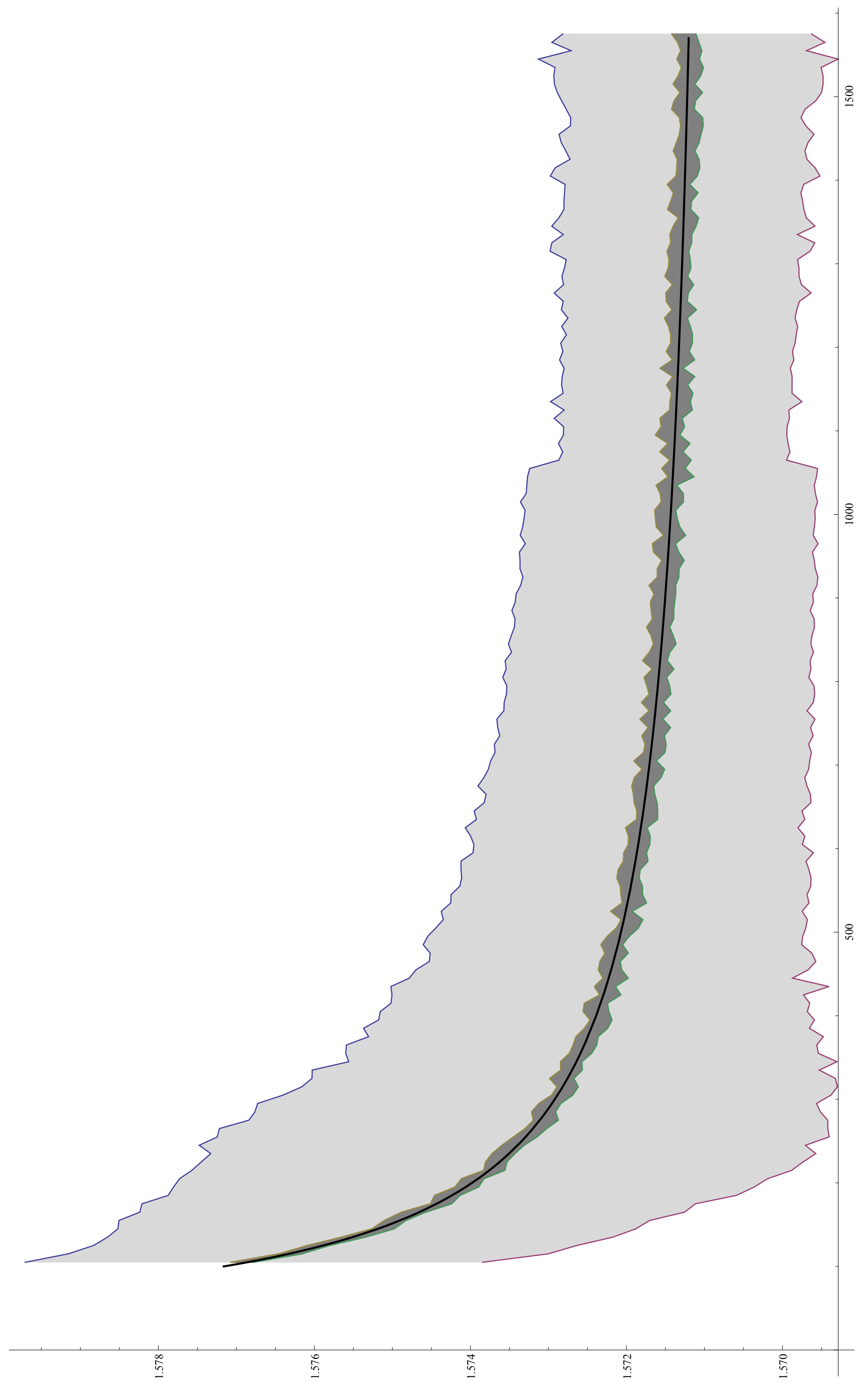

Figure 1.5: Comparing the error bars for the expected turning angle derived from numeric integration and direct sampling to the known asymptote. The black line is the asymptotic value, the light grey is the region between the sample average ± sample standard error, and the dark grey is the region between the numeric approximation of the integral ± the reported error.

Chapter 2

Planar Polygons

Recall that we have placed a measure on $Pol_2(n)$ and $Arm_2(n)$ by pushing forward the measure on $V_2(\mathbb{R}^n)$ and S^{2n-1}. As such, if we wanted to sample a random planar polygon uniformly under this measure, we could instead sample a point of $V_2(\mathbb{R}^n)$ and look at its image under the Hopf map. Likewise, we can sample a planar polygonal arm by sampling uniformly on the sphere and applying the Hopf map. Given that the integral we've proposed for finding total curvature is difficult to evaluate exactly, let us instead attempt to bound it.

2.1 Variation Distance

Given the intuition that the closure condition on polygons becomes less imposing as the number of edges grows, one expects that for any given configuration of k edges, where $k << n$, the probability of finding a this configuration present in a random closed polygon sampled from $Pol_2(n)$ should approach the probability of finding the configuration in a random arm sampled from $Arm_2(n)$. However, before this connection to Stiefel manifolds was made, this conjecture had not seen a rigorous proof. Let us correct this now.

Definition 83. Given two probability measures P and Q defined on a sigma-algebra $\mathscr{F}$ of the sample space Ω, define the total variation distance between P and Q to be

$$\|P - Q\|_{TV} = 2 \sup_{A \in \mathscr{F}} |P(A) - Q(A)|.$$

Informally, we can think of this as the largest possible difference between the probabilities that P and Q assign the same event.

Lemma 84 (From [9]). *If P and Q are absolutely continuous with respect to a reference measure ρ having densities p and q, then $\|P - Q\|_{TV} = \int |p - q| \mathrm{d}\rho$.*

To loosely see this connection, consider the set $B = \{x : p(x) \geq q(x)\}$. Then for any $A \in \mathscr{F}$, we have that $P(A) - Q(A) \leq P(A \cap B) - Q(A \cap B)$, as the difference in probability will not decrease by ignoring those elements for which $p(x) - q(x) < 0$. Likewise, $P(A \cap B) - Q(A \cap B) \leq P(B) - Q(B)$, as including more elements of B cannot decrease the difference. By similar reasoning, $Q(A) - P(A) \leq Q(B^c) - P(B^c)$. Hence,

$$2 \sup_{A \in \mathscr{F}} |P(A) - Q(A)| = [P(B) - Q(B)] + [Q(B^c) - P(B^c)] = \int |p - q| \mathrm{d}\rho.$$

Next, notice that we may sample an n-edge polygon from the symmetric measure on $Pol_2(n)$ by applying the map F_2 to a 2-frame sampled from the Haar measure on $V_2(\mathbb{R}^n)$. This 2-frame may in turn may be obtained by sampling a matrix from the Haar measure on $O(n)$ and taking the first two columns. Likewise, we may sample an n-edge arm from the symmetric measure on $Arm_2(n)$ by applying p to a point sampled from the spherical measure on $S^{2n-1}(\sqrt{2}) \subset \mathbb{R}^{2n}$, where $F_2 = p \circ i$ for the embedding of $V_2(\mathbb{R}^n)$ into $S^{2n-1}(\sqrt{2})$, and p is the map which sends the vector

$(x_1, y_1, x_2, y_2, \ldots, x_n, y_n)$ to $((x_n + y_n\mathbf{i})^2, (x_n + y_n\mathbf{i})^2, \ldots, (x_n + y_n\mathbf{i})^2) \in \mathbb{C}^n$.

Since these both of these maps, F_2 and p, are constructed by applying the map $z \mapsto z^2$ to pairs of real coordinates seen as the real and imaginary parts of a single complex coordinate, if we are interested in the distribution of the first few edges of polygons sampled from the symmetric measure, we need only focus on the distributions of the first few coordinates of $\{\vec{v} = (x_1, y_1, x_2, y_2, \ldots, x_n, y_n) \in \mathbb{R}^{2n} : \|v\|^2 = 2\}$ under the embeddings of $V_2(\mathbb{R}^n)$ and $S^{2n-1}(\sqrt{2})$ into $\mathbb{R}^{2n}$.

Persi Diaconis has published much on the topic of how small collections of coordinates of certain high-dimensional random variables behave like independent Gaussians. Specifically we will use the following two theorems:

Theorem 85. [8] *Suppose that Z is the $r \times s$ upper block of a random matrix U which is uniform on $O(n)$, implying that it has mean 0 and covariance of $\frac{1}{n} I_r \otimes I_s$. Let X be an rs multivariate normal distribution with the same mean and covariance. Then, provided that $r+s+2 < n$, the total variation distance between the law of Z and the law of X is bounded by $B(r,s;n) = 2\left(\left(1 - \frac{r+s+2}{n}\right)^{-c} - 1\right)$, where $c = \frac{t^2}{2}$ and $t =$min(r,s).*

Here, $A \otimes B$ is the Kronecker product of A and B, given by:

Definition 86. Where $A = (a_{i,j})$ is an $m \times n$ matrix and $B = (b_{i,j})$ is a $p \times q$ matrix we define the Kronecker product $A \otimes B$ to be the $mp \times nq$ matrix, given in block form as

$$\begin{bmatrix} a_{1,1}B & \ldots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \ldots & a_{m,n}B \end{bmatrix}$$

Theorem 87. [9] *Let $Q_{n,r,k}$ be the law of $(\xi_1, \ldots, \xi_k)$ when $(\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_n)$ is uniformly distributed over the surface of the sphere $\left\{\xi : \sum_{i=1}^{n} \xi_i^2 = r^2\right\}$. Let P_σ^k be the law of $\sigma\zeta_1, \ldots, \sigma\zeta_k$ where the ζ are independent standard normals. Then the total variation distance between $Q_{n,r,k}$ and $P_{r/\sqrt{n}}^k$ is bounded by $2\frac{k+3}{n-k-3}$, for $1 \leq k \leq n-4$.*

Equipped with these bounds on total variation distance between the distribution of certain multivariate normal distributions and small collections of coordinates in high-dimensional spheres and orthogonal matrices, we will now bound the total variation distance between the distribution of small collections of edges in high-dimensional polygons sampled from the respective symmetric measures on $Pol_2(n)$ and $Arm_2(n)$.

Theorem 88. *Let $P(k,n)$ be the law of the first k-edged segment of a random n-edged closed polygon sampled under the symmetric measure on $Pol_2(n)$, and $A(k,n)$ be the law of the first k-edged segment of a random n-edged arm sampled under the symmetric measure on $Arm_2(n)$. If $1 \leq k \leq n-2$, then we have that the total variation between $P(k,n)$ and $A(k,n)$ is bounded above by $\mathscr{B}_2(k,n) = 2\left(\frac{2k+3}{2n-2k-3} + \frac{(2n-k-4)(k+4)}{(n-k-4)^2}\right)$.*

Proof. First, notice that $LP(k,n)$ is given by the law of Z, the $k \times 2$ upper left block of a matrix sampled uniformly on $O(n)$. Likewise, $LA(k,n)$ would be the law of the first $2k$ coordinates of a point sampled from the sphere $\left\{\xi : \sum_{i=1}^{2n} \xi_i^2 = (\sqrt{2})^2\right\}$. To use Theorems 85 and 87, we will need X, the $2k$ multivariate normal distribution with mean 0 and covariance matrix $\frac{1}{n}I_{2k}$, and $P^{2k}_{\sqrt{2}/\sqrt{2n}}$, the law of $\frac{1}{\sqrt{n}}\zeta_1, \ldots, \frac{1}{\sqrt{n}}\zeta_{2k}$ with the ζ_i independent standard normals. Next, notice here that this $P^{2k}_{1/\sqrt{n}}$ is a multivariate distribution with mean 0 and covariance given by $\left(\frac{1}{\sqrt{n}}I_{2k}\right)\left(\frac{1}{\sqrt{n}}I_{2k}\right)^{\intercal} = \frac{1}{n}I_{2k}$. So we see that X and $P^{2k}_{1/\sqrt{n}}$ are multivariate normal distributions with the same mean and covariance. Notice that they are the same multivariate normal distribution, often denoted by $N\left(\mathbf{0}, \frac{1}{n}I_{2k}\right)$. Since total variation is a norm on measures, it satisfies the triangle inequality. We then have that:

$$\begin{aligned}
\|LP(k,n) - LA(k,n)\| &= \|LP(k,n) - N\left(\mathbf{0}, \frac{1}{n}I_2k\right) + N\left(\mathbf{0}, \frac{1}{n}I_2k\right) - LA(k,n)\| \\
&\leq \|LP(k,n) - N\left(\mathbf{0}, \frac{1}{n}I_2k\right)\| + \|N\left(\mathbf{0}, \frac{1}{n}I_2k\right) - LA(k,n)\|
\end{aligned}$$

$$
\begin{aligned}
&= \|Z - X\| + \|P^{2k}_{1/\sqrt{n}} - Q_{n,\sqrt{2},2k}\| \\
&\le 2\left(\left(1 - \frac{k+4}{n}\right)^{-2} - 1\right) + 2\left(\frac{2k+3}{2n-2k-3}\right) \\
&= 2\left(\left(\frac{n-k-4}{n}\right)^{-2} - 1\right) + 2\left(\frac{2k+3}{2n-2k-3}\right) \\
&= 2\left(\frac{n^2}{(n-k-4)^2} - \frac{n^2 - 2n(k+4) + (k+4)^2}{(n-k-4)^2} + \frac{2k+3}{2n-2k-3}\right) \\
&= 2\left(\frac{2n(k+4) - (k+4)^2}{(n-k-4)^2} + \frac{2k+3}{2n-2k-3}\right) \\
&= 2\left(\frac{(2n-k-4)(k+4)}{(n-k-4)^2} + \frac{2k+3}{2n-2k-3}\right).
\end{aligned}
$$

□

Proposition 89. *$\mathscr{B}_2(k,n)$ is asymptotic to* $\dfrac{6k+19}{n}$.

Proof. Combining this expression, and using the notation of $o(n^p)$ to represent a quantity, $q(n)$, that satisfies the limiting condition $\lim_{n\to\infty} \dfrac{q(n)}{n^p} = 0$, we have:

$$
\begin{aligned}
\mathscr{B}_2(k,n) &= 2\left(\frac{(2n-k-4)(k+4)}{(n-k-4)^2} + \frac{2k+3}{2n-2k-3}\right) \\
&= 2\left(\frac{(2n-k-4)(k+4)(2n-2k-3)}{(n-k-4)^2(2n-2k-3)} + \frac{(2k+3)(n-k-4)^2}{(n-k-4)^2(2n-2k-3)}\right) \\
&= 2\left(\frac{4(k+4)n^2 + (2k+3)n^2 + o(n^2)}{(n-k-4)^2(2n-2k-3)}\right) \\
&= 2\left(\frac{(6k+19)n^2 + o(n^2)}{2n^3 + o(n^3)}\right) \\
&= \frac{6k+19}{n} + o(n^{-1}).
\end{aligned}
$$

□

Before moving on, let us take a moment to discuss this bound in more detail. For one thing, if one fixes k, then $\mathscr{B}_2(k,n) \sim \dfrac{6k+19}{n}$ will limit to 0 as $n \to \infty$. It should be pointed out that $\mathscr{B}_2(k,n)$ is decreasing to this asymptotic: for any $k < \dfrac{n}{2}$, we will have

$\mathscr{B}_2(k,n) > \dfrac{6k+19}{n}$. Nonetheless, if we wanted to let k grow with n, then as long as k is $o(n)$ [for example $k = \lambda n^p$ for any $p \in [0,1)$ and $\lambda \in (0,\infty)$], we have that $\lim_{n\to\infty} \mathscr{B}_2(k,n) = 0$. This is sharp in the sense that if we write $k = \alpha n$, then we can see that:

$$\begin{aligned}
\lim_{n\to\infty} \mathscr{B}_2(\alpha n, n) &= \lim_{n\to\infty} 2\left(\frac{(2n-\alpha n-4)(\alpha n+4)}{(n-\alpha n-4)^2} + \frac{2\alpha n+3}{2n-2\alpha n-3}\right) \\
&= \lim_{n\to\infty} 2\left(\frac{((2-\alpha)n-4)(\alpha n+4)}{((1-\alpha)n-4)^2} + \frac{2\alpha n+3}{2(1-\alpha)n-3}\right) \\
&= \lim_{n\to\infty} 2\left(\frac{(2-\alpha)\alpha n^2-\alpha n+4(2-\alpha n)-16}{(1-\alpha)^2n^2-8(1-\alpha)n+16} + \frac{2\alpha n+3}{2(1-\alpha)n-3}\right) \\
&= 2\left(\frac{(2-\alpha)\alpha}{(1-\alpha)^2} + \frac{\alpha}{1-\alpha}\right) \\
&= 2\alpha\left(\frac{2-\alpha}{(1-\alpha)^2} + \frac{1-\alpha}{(1-\alpha)^2}\right) \\
&= \frac{2\alpha(3-2\alpha)}{(1-\alpha)^2}.
\end{aligned}$$

It is not hard to see that for $\alpha > \dfrac{4-\sqrt{11}}{5} \simeq 0.136675$, this limit, $\dfrac{2\alpha(3-2\alpha)}{(1-\alpha)^2}$, is greater than 1. This tells us that as long as, in the limit, $k <$ 13% of n, we have some information about the distribution of k-edged segments in $Pol_2(n)$ by virtue of our knowledge of the distribution of k-edged segments in $Arm_2(n)$. To get a better sense of this bound, consider the following table of values of $\mathscr{B}_2(k,n)$:

		n			
		300	500	1,000	10,000
k	2	0.106074	0.0629767	0.0312423	0.00310241
	3	0.127162	0.0753618	0.0373375	0.00370335
	10	0.280319	0.163969	0.0804659	0.00791443
	20	0.517348	0.296629	0.143515	0.0139439

Definition 90. Call a function $f : Arm_d(n) \to \mathbb{R}$ a k-edged locally defined function if $f([\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}, \vec{u_1}, \ldots, \vec{u_{n-k}}]) = f([\vec{e_1}, \vec{e_2}, \ldots, \vec{e_k}, \vec{v_1}, \ldots, \vec{v_{n-k}}])$ for all $\vec{e_i}, \vec{u_j}, \vec{v_j} \in \mathbb{R}^d$, $i \in \{1, 2, \ldots, k\}, j \in \{1, 2, \ldots, n-k\}$.

Theorem 91. *Let f be an essentially bounded, k-edged locally defined function. Then the expectation of f over $Pol_2(n)$ may be approximated by the expectation of f over $Arm_2(n)$ to within $M\mathscr{B}_2(k,n)$, where M is a bound for f almost everywhere.*

Proof. The expectation of f over $Arm_2(n)$, $E_{Arm_2(n)}(f)$ is given by $\int_{Arm_2(n)} f\mu_A$, where μ_A is the symmetric measure on $Arm_2(n)$. Likewise, the expectation of f over $Pol_2(n)$, $E_{Pol_2(n)}(f)$ is given by $\int_{Pol_2(n)} f\nu_P$, where ν_P is the symmetric measure on $Arm_2(n)$. Since f is a k-edged locally defined function, we may integrate out the last $n-k$ edges to obtain $E_{Arm_2(n)} = \int_{\mathscr{A}_2(k)} f(q)\mu_n^k$, where μ_n^k is the law of the first k edges of a polygon sampled from the symmetric measure on $Arm_2(n)$. Similarly, we can see that $E_{Pol_2(n)} = \int_{\mathscr{A}_2(k)} f(q)\nu_n^k$, where ν_n^k is the law of the first k edges of a polygon sampled from the symmetric measure on $Pol_2(n)$. We may therefore write:

Let g be the function that takes an ordered pair of vectors in $\mathbb{R}^n$ to the value of q determined by the first k edges of the polygon determined by Hopf of the pair. Then we have that the expectation of q, $E_{pol}(q)$ is the result of integrating f against the masure $\mu = LP(k,n)$ on the space of configurations of k edges. Likewise, we have the measure $\nu = LA(k,n)$. As such, we can write:

$$
\begin{aligned}
|E_{Pol_2(n)}(q) - E_{Arm_2(n)}(q)| &= |\int f\nu_n^k - \int f\mu_n^k| \\
&= |\int f(\nu_n^k - \mu_n^k)| \\
&\leq \|f\|_\infty \|\nu_n^k - \mu\|_{TV} \\
&\leq \|f\|_\infty \mathscr{B}_2(k,n).
\end{aligned}
$$

□

Corollary 92. *Let f be an essentially bounded, k-edged locally defined function. Let $E_p(n)$ stand for the expectation of f over $Pol_2(n)$, and $E_a(n)$ stand for the expectation of f over $Arm_2(n)$. Further, let $\widetilde{E_p}(n)$ stand for the expectation of the sum of f over a polygon in $Pol_2(n)$, by which we mean the expectation of the quantity $\sum_{i=1}^{n} f(e_i, e_{i+1}, \ldots, e_{i+k})$, where the indices are taken modulo n. Additionally, let $\widetilde{E_a}(n)$ stand for the expectation of the sum of f over a polygon in $Arm_2(n)$. If $\lim_{n\to\infty} nE_a(n) = \infty$, then $\lim_{n\to\infty} \frac{E_p(n)}{E_a(n)} = 1$ and $\lim_{n\to\infty} \frac{\widetilde{E_p}(n)}{\widetilde{E_a}(n)} = 1$.*

Proof. From Theorem 91, we see that $E_a(n) - M\mathscr{B}_2(k,n) \leq E_p(n) \leq E_a(n) + M\mathscr{B}_2(k,n)$. Dividing through by $E_a(n)$, this becomes $1 - M\frac{\mathscr{B}_2(k,n)}{E_a(n)} \leq \frac{E_p(n)}{E_a(n)} \leq 1 + M\frac{\mathscr{B}_2(k,n)}{E_a(n)}$. From Proposition 89, we see that $\mathscr{B}_2(k,n)$ is asymptotic to $\frac{6k+19}{n}$. This, in addition to our assumption on $\lim_{n\to\infty} nE_a(n)$, tells us that $\lim_{n\to\infty} M\frac{\mathscr{B}_2(k,n)}{E_a(n)} = \lim_{n\to\infty} M\frac{6k+19}{nE_a(n)} = 0$. The first result then follows from the Squeeze Theorem [39].

In the second case, note that from the invariance under permutations, we have that $\widetilde{E_p}(n) = nE_p(n)$ and $\widetilde{E_a}(n) = (n-k-1)E_a(n)$. This produces the inequality $\widetilde{E_a}(n) - nM\mathscr{B}_2(k,n) \leq \widetilde{E_p}(n) \leq \widetilde{E_a}(n) + nM\mathscr{B}_2(k,n)$, which we can divide through to produce: $1 - M\frac{\mathscr{B}_2(k,n)}{E_a(n)}\left(\frac{n}{n-k-1}\right) \leq \frac{\widetilde{E_p}(n)}{\widetilde{E_p}(n)} \leq 1 + M\frac{\mathscr{B}_2(k,n)}{E_a(n)}\left(\frac{n}{n-k-1}\right)$. The result then follows from the Squeeze Theorem and our earlier observation that $M\frac{\mathscr{B}_2(k,n)}{E_a(n)} \to 0$ as $n \to \infty$. $\square$

2.2 Curvature

Let us return to the question of the total curvature of planar polygons. As stated before, this is defined as the sum of the turning angles, and when we sample under the symmetric measure, each turning angle has the same expectation. Combining this with the fact that an expectation of a sum is the sum of the expectations (even for highly correlated data) as cited before from Section 5.1 of [12], we see that the expectation of total curvature will be n

times the expectation of a turning angle. For arms, we have already seen that the turning angle has expected value of $\frac{\pi}{2}$. Using Theorem 91, we see that $|E_{Pol_2(n)}(\theta) - \frac{\pi}{2}| \leq \pi\mathscr{B}_2(2,n)$. Moreover, we know that a closed polygon will have a higher expected turning angle than a polygonal arm. As such, we see that we can bound the expectation of the total curvature of a closed planar polygon as:

$$0 \leq E_{pol}(\kappa) - n\frac{\pi}{2} \leq 2n\pi\left(\frac{7}{2n-7} + \frac{12(n-3)}{(n-6)^2}\right)$$

Of particular interest, we see from taking the limit of this inequality, that the expectation of total curvature of planar polygons lies between $n\frac{\pi}{2}$ and $31\pi + n\frac{\pi}{2} + O\left(\frac{1}{n}\right)$. Of course we already have a trivial upper bound of $n\pi$, but the bound we show is better, provided that $n > 69$.

Even though it has already been shown in [6] that $\frac{E_{Pol_2(n)}(\kappa)}{E_{Arm_2(n)}(\kappa)} \to 1$, our corollary here shows this not to be an artifact of total curvature, but of the proximity in distribution between pairs of edges in open polygonal chains and pairs of edges in closed polygonal chains. Let us now look at the variance of total curvature.

First note that we have a trivial upper bound on the variance: Fix some $n \in \mathbb{N}$. Note that the largest total curvature possible for an n-edged closed polygon is $n\pi$, occurring if the entire polygon lies on a line, with edge directions reversing at each vertex. Likewise, the smallest total curvature possible is 2π, occurring if the entire polygon lies on a line with a single vertex having edges with distinct directions. In the most extreme example, were it the case that all of the n-edged closed polygons were partitioned into such extremes, with $\frac{n-4}{2n-4} \times 100\%$ having total curvature $n\pi$ and the remaining $\frac{n}{2n-4} \times 100\%$ having total curvature 2π, then the expectation of total curvature would be $\frac{\pi}{2}$ and the variance of the total curvature of $\pi^2\left(\frac{n^2-4n}{4}\right)$. Chebyshev's inequality (Theorem 93 below, cited as a common corollary to Markov's Inequality, page 91 of [38]), would then tell us that the probability that a sampled polygon has total curvature between $\frac{n\pi}{2} \pm k\pi \frac{\sqrt{n(n-4)}}{2}$ is greater than or equal to $1 - \frac{1}{k^2}$. As usual, this is only relevant for $k > 1$, but in such a case, $k\pi \frac{\sqrt{n(n-4)}}{2} \simeq kn\frac{\pi}{2}$, so we have gained almost no additional insight.

Theorem 93 (Chebyshev's Inequality, from [38])**.** *Let X be a random variable with finite expected value μ and finite non-zero variance σ^2. Then for any real number $k > 0$,* $\Pr(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$.

Proposition 94. *The variance of total curvature of a random polygon sampled under the symmetric measure on $Pol_2(n)$ is bounded by*

$$M = \pi^2 \left(n\mathscr{B}_2(2,n) + 2n\mathscr{B}_2(3,n) + (n^2 - 3n)\mathscr{B}_2(4,n) \right) - n^2 \left(\pi\epsilon_n + \epsilon_n^2 \right),$$

where $e_n = E_{Pol_2(n)}[\theta_1] - \frac{\pi}{2}$ is surplus of the expectation of the turning angle of a polygon over $Pol_2(n)$ over $\frac{\pi}{2}$.

Corollary 95. *The variance of total curvature of a random polygon sampled under the symmetric measure on $Pol_2(n)$ is bounded above by $(n\pi)^2\mathscr{B}_2(4,n) \simeq 43n\pi^2$.*

Proof of Proposition. We know that the covariance of a pair can be computed as either $\mathrm{Cov}(\theta_i, \theta_j) = E[(\theta_i - E[\theta_i])(\theta_j - E[\theta_j])] = E[\theta_i\theta_j] - E[\theta_i]E[\theta_j]$. We already have established the bounds that $\frac{\pi}{2} \leq E(\theta_1) \leq \frac{\pi}{2} + \pi\mathscr{B}_2(2,n)$. For convenience, let $t_n = E(\theta_1)$ and define $\epsilon_n = t_n - \frac{\pi}{2}$, so that $\epsilon_n > 0$ and $\epsilon_n \to 0$ as $n \to \infty$.

We may partition the pairs (θ_i, θ_j) into three categories: (1) $j \equiv i \bmod n$, (2) $j \equiv i \pm 1 \bmod n$, and (3) $j \equiv i \pm k \bmod n$ for $1 < k < \frac{n}{2}$. By the symmetry of the measure, the covariance of any pair will be the same as the covariance of any other pair from the same category. More, we see that in each category, the covariance is the integral of an essentially bounded function determined by a set of consecutive edges (a pair of edges in category 1, a triple of edges in category 2, and four edges in category 3).

Recall then that the variance of a sum is equal to the sum of the covariance of the pairs (corollary 8 on page 64 of [15]). This tells us that the variance of total curvature may be partitioned into the sum:

$$
\begin{aligned}
Var\left(\sum_{i=1}^{n}\theta_i\right) &= \left(\sum_{i=1}^{n}\mathrm{Cov}(\theta_i,\theta_i)\right) + 2\left(\sum_{i=1}^{n}\mathrm{Cov}(\theta_i,\theta_{i+1})\right) + 2\left(\sum_{i=1}^{n-2}\sum_{k=2}^{n-i}\mathrm{Cov}(\theta_i,\theta_{i+k})\right)\\
&= \left(\sum_{i=1}^{n}\mathrm{Cov}(\theta_1,\theta_1)\right) + 2\left(\sum_{i=1}^{n}\mathrm{Cov}(\theta_1,\theta_2)\right) + 2\left(\sum_{i=1}^{n-2}\sum_{k=2}^{n-i}\mathrm{Cov}(\theta_1,\theta_3)\right)\\
&= n\,\mathrm{Cov}(\theta_1,\theta_1) + 2n\,\mathrm{Cov}(\theta_1,\theta_2) + n(n-3)\,\mathrm{Cov}(\theta_1,\theta_3).
\end{aligned}
$$

By choosing to compute $\mathrm{Cov}(\theta_i,\theta_j) = E[\theta_i\theta_j] - E[\theta_i]E[\theta_j]$, and recalling our definition that $t_n = E[\theta_i]$, we may express this as:

$$
\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n}\theta_i\right) &= n(E[\theta_1^2]-t_n^2) + 2n(E[\theta_1\theta_2]-t_n^2) + (n^2-3n)(E[\theta_1\theta_3]-t_n^2)\\
&= nE[\theta_1^2] + 2nE[\theta_1\theta_2] + (n^2-3n)E[\theta_1\theta_3] - (nt_n)^2.
\end{aligned}
$$

We can compute $E[\theta_1^2]$ as the integral of a scale-invariant function determined by a pair of edges that is essentially bounded by π^2. This means that we may use Theorem 91 to conclude that $|E_{Pol_2(n)}[\theta_1^2] - E_{Arm_2(n)}[\theta_1^2]| \leq \pi^2\mathscr{B}_2(2,n)$. A simple calculation shows that $E_{Arm_2(n)}[\theta_1^2] = \dfrac{\pi^2}{4}$, so we have that $E_{Pol_2(n)}[\theta_1^2] \leq \pi^2\left(\mathscr{B}_2(2,n)+\dfrac{1}{4}\right)$.

Likewise, $E[\theta_1\theta_2]$ and $E[\theta_1\theta_3]$ are computed as the integral of scale-invariant functions determined by three and four edges respectively. The independence of edge directions in $Arm_2(n)$ tells us that $E_{Arm_2(n)}[\theta_1\theta_2] = E_{Arm_2(n)}[\theta_1]E_{Arm_2(n)}[\theta_2] = \dfrac{\pi^2}{4}$. So we see that $E_{Pol_2(n)}[\theta_1]E_{Pol_2(n)}[\theta_3] \leq \pi^2\left(\mathscr{B}_2(3,n)+\dfrac{1}{4}\right)$ and $E_{Pol_2(n)}[\theta_1]E_{Pol_2(n)}[\theta_2] \leq \pi^2\left(\mathscr{B}_2(4,n)+\dfrac{1}{4}\right)$.

This allows us to place an upper bound on the variance of total curvature for $Pol_2(n)$ as follows:

$$\begin{aligned}
\text{Var}\left(\sum_{i=1}^{n}\theta_i\right) &= nE[\theta_1^2] + 2nE[\theta_1\theta_2] + (n^2-3n)E[\theta_1\theta_3] - (nt_n)^2 \\
&\leq n\pi^2\mathscr{B}_2(2,n) + 2n\pi^2\mathscr{B}_2(3,n) + (n^2-3n)\pi^2\mathscr{B}_2(4,n) + \frac{n^2\pi^2}{4} - n^2(\frac{\pi}{2}+\epsilon_n)^2 \\
&\leq \pi^2\left(n\mathscr{B}_2(2,n) + 2n\mathscr{B}_2(3,n) + (n^2-3n)\mathscr{B}_2(4,n)\right) - n^2\left(\pi\epsilon_n + \epsilon_n^2\right) \\
&\leq \pi^2 n^2\mathscr{B}_2(4,n)
\end{aligned}$$

Next, we claim that $\mathscr{B}_2(k,n)$ is increasing in k. Recall that $\mathscr{B}_2(k,n) = 2\left(\frac{(2n-k-4)(k+4)}{(n-k-4)^2} + \frac{2k+3}{2n-2k-3}\right)$. The second summand is clearly increasing, with k, as the denominator is decreasing while the numerator is increasing. The first summand likewise has a decreasing denominator, and the numerator, $(2n-(k+4))(k+4)$ is quadratic in k with negative concavity. Since the critical point of this quadratic occurs at $k = n-4$ (which is also the largest k for which the bound holds), we see that the numerator of the first summand is also increasing.

We can now establish a larger bound by replacing $\mathscr{B}_2(2,n)$ and $\mathscr{B}_2(3,n)$ with $\mathscr{B}_2(3,n)$. We then obtain an even larger bound by ignoring the $-n(\pi\epsilon_n + \epsilon_n^2)$. $\square$

Chebyshev's inequality tells us that the probability of a polygon having total curvature κ inside the range of $[nt_n - \lambda\sqrt{\text{Var}}, nt_n + \lambda\sqrt{\text{Var}}]$ is at least $1-\frac{1}{\lambda^2}$. By the above corollary, we can write extend this to a slightly larger, but easier to work with interval by replacing $\sqrt{\text{Var}}$ with $n\pi\sqrt{\mathscr{B}_2(4,n)}$. This interval is then $[n\left(\frac{\pi}{2} + \epsilon_n - \lambda\pi\sqrt{\mathscr{B}_2(4,n)}\right), n\left(\frac{\pi}{2} + \epsilon_n + \lambda\sqrt{\mathscr{B}_2(4,n)}\right)]$. We know from Proposition 82 that ϵ_n is asymptotic to $\frac{2}{n\pi}$. We also know from Proposition 89 that $\mathscr{B}_2(4,n)$ is asymptotic to, and less than, $\frac{43}{n}$, so we see that, asymptotically, $\epsilon_n < \mathscr{B}_2(4,n)$. Since we only obtain useful information when $\lambda > 1$, this interval may be augmented to $\left[n\pi\left(\frac{\pi}{2}-\frac{7\lambda}{\sqrt{n}}\right), n\pi\left(\frac{\pi}{2}+2\frac{7\lambda}{\sqrt{n}}\right)\right]$. Notice here, that the length of the interval is $21\pi\sqrt{n}$. So the length of this interval is growing at a rate of $O(\sqrt{n})$.

Of course we already have the trivial bounds that all planar polygons in $Pol_2(n)$ will have total curvature between 2π and $n\pi$, so let us first check that these bounds are better than that. By setting $\lambda = \sqrt{2}$, we can say that at most $\frac{1}{2}$ of the polygons in $Pol_2(n)$, lie outside our given bounds, and that the lower will be larger than 2π when $n \geq 48$, while the upper will be smaller than $n\pi$ when $n \geq 159$. Past those marks, our bounds from variance become more useful than the trivial bounds.

Before moving on, we would like to point out that this analysis is adaptable for any essentially bounded k-edged locally defined function. By which we mean that the interested reader will be able to verify that the arguments leading to Proposition 94 and Corollary 95 could be slightly adjusted to say that the variance of the sum of f over all n runs of consecutive k-edges in a polygon is bounded by the quantity $(nM)^2\mathscr{B}_2(2k, n)$, where M is a bound for f almost everywhere.

2.3 Intersections

We now turn to the topic of self intersections. As with curvature, finding the expectation of the total number self-intersections is a bit large of a task, so we will instead focus on bounding it. If we partition the polygon into triples of consecutive edges, we can see that the total number of self-intersections of the polygon is certainly larger than the total number of triples that contain a self-intersection. Recall then that for polygonal arms, the edges are all independently sampled, and then scaled down to the appropriate length. As such, the expected number of triples that contain a self-intersection in arms of $3n$ edges would be bounded by n times the expected number of self-intersections in an arm of 3 edges.

If a configuration of three edges is to intersect itself, then the third edge must intersect the first edge. Label these edges e_1, e_2, e_3, and their endpoints v_0, v_1, v_2, and v_3, where e_i has endpoints v_{i-1} and v_i. Further, let us denote by r_i and θ_i the length and direction of e_i. For convenience, rotate, and reflect if necessary, so that $\theta_1 = 0$ and $\theta_2 \in [0, \pi]$

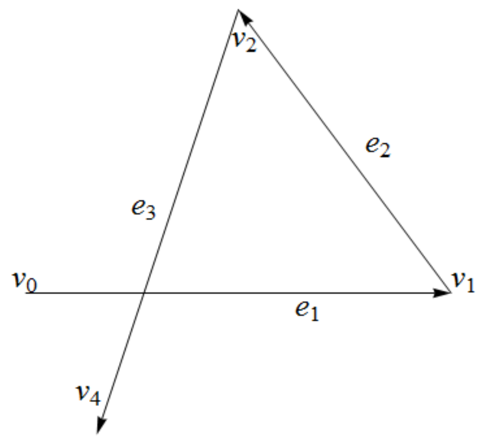

Figure 2.1: Intersecting 3 Edged Arm

Proposition 96. *Let σ be the sign function, the function* $\sigma(x) = \begin{cases} |x|x^{-1}, & x > 0 \\ 0, & x = 0 \end{cases}$. *Then the function I on $\mathbb{R}^3_{>0} \times [0, \pi] \times [0, 2\pi)$ given by*

$$\begin{aligned} 4I(r_1, r_2, r_3, \theta_2, \theta_3) = 1 &- \sigma[r_2 \sin(\theta_2) + r_3 \sin(\theta_3)] \\ &- \sigma[\sin(\theta_2 - \theta_3)(r_2 \sin(\theta_2 - \theta_3) - r_1 \sin(\theta_3))] \\ &+ \sigma[(r_2 \sin(\theta_2) + r_3 \sin(\theta_3))(r_2 \sin(\theta_2 - \theta_3) - r_1 \sin(\theta_3)) \sin(\theta_2 - \theta_3)]), \end{aligned}$$

returns 1 if the corresponding three edge polygonal arm with edges $e_1 = (r_1, 0)$, $e_2 = (r_2, \theta_2)$, $e_3 = (r_3, \theta_3)$ contains a self-intersection and 0 otherwise.

Moreover, the function $I(r_1, r_2, r_3, \theta_2, \theta_3)$ is the characteristic function of the the set

$$I^s = \left\{ (r_1, r_2, r_3, \theta_2, \theta_3) : r_3 > -r_2 \frac{\sin(\theta_2)}{\sin(\theta_3)} \,\&\, r_1 > r_2 \frac{\sin(\theta_2 - \theta_3)}{\sin(\theta_3)} \right\}$$

Proof. First, observe that two line segments intersect when the each of the two half-planes created by extending one segment to a full line contains the endpoints of the other segment, and that condition holds when reversing the roles of the segments. For two points in the plane, $x, y \in \mathbb{R}^2$, let $L(x, y; p)$ be a function representing the line between x and y, so that the line is the locus of points $\{p \in \mathbb{R}^2 : L(x, y; p) = 0\}$. Then one can determine which half-plane a point p lies on by checking the sign of $L(x, y; p)$. Writing the criteria for intersections in terms of the sign of $L(x, y; p)$ then gives us:
$I(r_1, r_2, r_3, \theta_2, \theta_3) = \frac{1}{4}(1-\sigma(L_1(v_2))\sigma(L_1(v_3)))(1-\sigma(L_3(v_0)\sigma(L_3(v_1))))$, where v_0 is the origin, v_1 is the point $(r_1, 0)$, v_2 is the point $v_1 + r_2(\cos(\theta_2), \sin(\theta_2)$, v_3 is the point $v_2 + r_3(\cos(\theta_3), \sin(\theta_3))$, $L_1(p) = L(v_0, v_1; p)$ and $L_2(p) = L(v_2, v_3; p)$. It is clear from this expression of $I(r_1, r_2, r_3, \theta_2, \theta_3)$ that the image of I is precisely the two point set $\{0, 1\}$, so I will be the characteristic function of some set.

We can expand this formula and simplify to arrive at the proposed simplified form. The computation of this involves many lines that must split in unappealing ways, so it has been moved to Appendix B for the fastidious reader's enjoyment. Picking up from there, we have that

$$\begin{aligned}
4I(r_1, r_2, r_3, \theta_2, \theta_3) = 1 & \\
& - \sigma[r_2 \sin(\theta_2) + r_3 \sin(\theta_3)] \\
& - \sigma[\sin(\theta_2 - \theta_3)(r_2 \sin(\theta_2 - \theta_3) - r_1 \sin(\theta_3))] \\
& + \sigma[(r_2 \sin(\theta_2) + r_3 \sin(\theta_3))(r_2 \sin(\theta_2 - \theta_3) - r_1 \sin(\theta_3)) \sin(\theta_2 - \theta_3)]),
\end{aligned}$$

Next, let us consider the requirements for $I(r_1, r_2, r_3, \theta_2, \theta_3)$ to be 1. We need both $(r_2 \sin(\theta_2) + r_3 \sin(\theta_3))$ and $\sin(\theta_2 - \theta_3)(r_2 \sin(\theta_2 - \theta_3) - r_1 \sin(\theta_3)))$ to be negative and their product to be positive. Of course the product of any two negative numbers is positive, so we need only look at the the middle two summands of I. showing the For this first to be negative, it must be the case that $r_3 \sin(\theta_3) < -r_2 \sin(\theta_2)$. Since $\theta_2 \in [0, \pi]$, we know that $\sin(\theta_2) \geq 0$, so we require $\sin(\theta_3) < 0$, telling us that we must have that $\theta_3 \in (\pi, 2\pi)$. With $\sin(\theta_3) \neq 0$, let us write this requirement as $r_3 > -r_2 \dfrac{\sin(\theta_2)}{\sin(\theta_3)}$. For the second inequality, we need $(\sin(\theta_2 - \theta_3)(r_2 \sin(\theta_2 - \theta_3) - r_1 \sin(\theta_3)))$, to be negative. In the case that $\sin(\theta_2 - \theta_3)$ is positive, we need $\sin(\theta_3)$ positive, which cannot happen as we already required $\theta_3 \in (\pi, 2\pi)$. This leaves us with the case that $\sin(\theta_2 - \theta_3)$ is negative. In that case, we need $r_2 \sin(\theta_2 - \theta_3) > r_1 \sin(\theta_3)$, so dividing by this negative quantity gives us $r_1 > r_2 \dfrac{\sin(\theta_2 - \theta_3)}{\sin(\theta_3)}$, as desired. □

Proposition 97. *The expected number of self-intersections in a 3-edged planar polygonal arm is within the range of* $0.05246379 \pm 0.00006397$.

Proof. Since this is a scale invariant quantity, Theorem 70 tells us that it has the same expected value over $\mathscr{A}_2(3)$ as it does over $Arm_2(3)$. Next, to integrate I over $\mathscr{A}_2(3)$, notice that I is the characteristic function of a region, so its expectation is the same as the measure of that region. We have then that the number of self-intersections is given by the integral:

$$\int_0^\infty \int_\pi^{2\pi} \int_0^\pi \int_{r_2 \sin(\theta_2 - \theta_3)/\sin(\theta_3)}^\infty \int_{-r_2 \sin(\theta_2)/\sin(\theta_3)}^\infty G_1(r_1) G_1(r_2) G_1(r_3) r_1 r_2 r_3 \, \mathrm{d}r_3 \mathrm{d}r_1 \mathrm{d}\theta_2 \mathrm{d}\theta_3 \mathrm{d}r_2.$$

Working this integral out by hand may prove tedious, so we will instead supply the numeric approximation of the integral. Using an adaptive Monte Carlo algorithm on one million points yields 0.05246379 with an error bound estimate of 0.00006397. □

As such, we find that a given arm with $3n$ edges is expected to have more than $0.05246379n$ self-intersections, as that is the number of expected self-intersections formed by an edge $\vec{e_k}$ with the edge $\vec{e_{k+2}}$, where $k \equiv 1 \pmod 3$. Utilizing Theorem 91, we can then establish bounds on the probability of finding a self-intersection formed by the first three edges of a closed polygon of $0.05246379 \pm \mathscr{B}_2(3, n)$. As before, the symmetric nature of the symmetric measure then lets us say that the expected number of self-intersections occurring from a triple of edges of the form $\vec{e_k}, \vec{e_{k+1}}, \vec{e_{k+2}}$ with $k \equiv 1 \pmod 3$ in a $3n$-edged polygon is bounded by $n(0.05246379 \pm \mathscr{B}_2(3,n))$. Looking at the asymptotic value, we see that this is limiting to $0.05246379n \pm \frac{37}{3}$.

Since this is only a bound on the expectation, we should point out that the example case in which $(05.246379 + \frac{18.5}{n})\%$ of $3n$-gons have the property that, under this partition, all n of the 3-edge configurations have an intersection, while the remaining polygons all have no intersections of this type would also achieve this asymptotic bound. We could, as before, turn to a variance argument to uncover more information. However, there is another perspective we can take that is worth investigating.

2.4 Second Bound for Intersections

Recall that if an event occurs with probability p, and we have r independent samples, then the probability of the event occurring at least once is given by $1-(1-p)^r$. In particular, this tells us that that the probability of an arm with $3n$ edges to have at least one self-intersection arising from our 3-edge configuration is $1 - 0.947536^n$. We know then, that the probability of the event occurring at least once is approaching 1, which is something we could not immediately say previously.

Now consider the probability of at least one self-intersection arising from a configuration of 3 consecutive edges occurring in the first r sets of 3 edges in a closed polygon with n edges. This quantity then depends only on the first $3r$ edges, and is represented by a characteristic function of a measurable set. As such, invoking Theorem 91, tells us that the probability, $p_n(r)$, of finding at least one self-intersection from a 3-edge configuration is bounded as:

$$|p_n(r) - (1 - 0.947536^r)| \leq \mathscr{B}_2(3r, n) = 2\left(\frac{12r+6}{2n-6r-3} - \frac{(2n-3r-4)(3r+4)}{(n-3r-4)^2}\right)$$

Since we are only particularly interested in bounding from below, we will only care about the part of the absolute value that gives $p_n(r) \geq (1 - 0.947536^r) - \mathscr{B}_2(3r, n)$. Set $q_n(r)$ to be this lower bound. We are now faced with the natural question: When is $q_n(r)$ the maximized?

This is a straight-forward optimization problem on the closed domain of $r \in \left[1, \frac{n-4}{3}\right)$. Letting $\alpha = 0.947536$ and taking the derivative with respect to r, will give us, after a little simplification:

$$q_n'(r) = 12n\left(\frac{-n}{(n-3r-4)^3} - \frac{2}{(2n-6r-3)^2}\right) - \alpha^r \ln(\alpha)$$

Since $0 < \alpha < 1$, we have that $-\alpha^r \ln(\alpha) > 0$. On the other hand, since we know that $n > 0$ and $r < \frac{n-4}{3}$, we see that both $\frac{-n}{(n-3r-4)^3} < 0$ and $\frac{-2}{(2n-6r-3)^2} < 0$. Looking at the second derivative, we have that:

$$q_n''(r) = 36n\left(\frac{-3n}{(n-3r-4)^4} - \frac{8}{(2n-6r-3)^3}\right) - \alpha^r \ln(\alpha)^2.$$

Notice here that this will be strictly negative, as it is the sum of 3 strictly negative terms: The first being a negative number divided by an even power of some quantity, the second being a negative 8 divided by a power of the positive quantity $2n - 6r - 3 > 2n - 2(n-4) - 3 = 5$, and the third being the product of a power of the positive constant α an an even power of $\ln(\alpha)$. This tells us that $q_n(r)$ is concave down, and may admit at most one critical point, which would be a global maximum. At the lower endpoint, $r = 1$, we see that $q_n(1) > 0$ (provided that $n > 705$), while $q_n\left(\frac{3n}{20}\right) < 0$, as we have pointed out prior that $\mathscr{B}_2(k, n) > 1$ for $k > 14\%$ of n. Hence, we know that a global maximum is achieved. Unfortunately, solving directly for this critical point, is non-trivial. If we allow ourselves the approximation $\mathscr{B}_2(k, n) \geq \frac{6k+19}{n}$, we see that $q_n(r) \geq 1 - \alpha^r - \frac{18r+19}{n}$, we can observe that the only critical point of this lower bound of $q_n(r)$ occurs at

$\dfrac{\ln(18) - \ln(-n\ln(\alpha))}{\ln(\alpha)} \simeq -107.8334 + 18.5562\ln(n)$. Notice below that this curve, which is positive for $n > 334$ and greater than 2 for $n > 372$ provides a lower bound for the critical point of $q_n(r)$. Applying logarithmic regression to the values of the critical points for n from 242 to 10000 provides the other curve shown, $-107.8334 + 18.5562\ln(n)$.

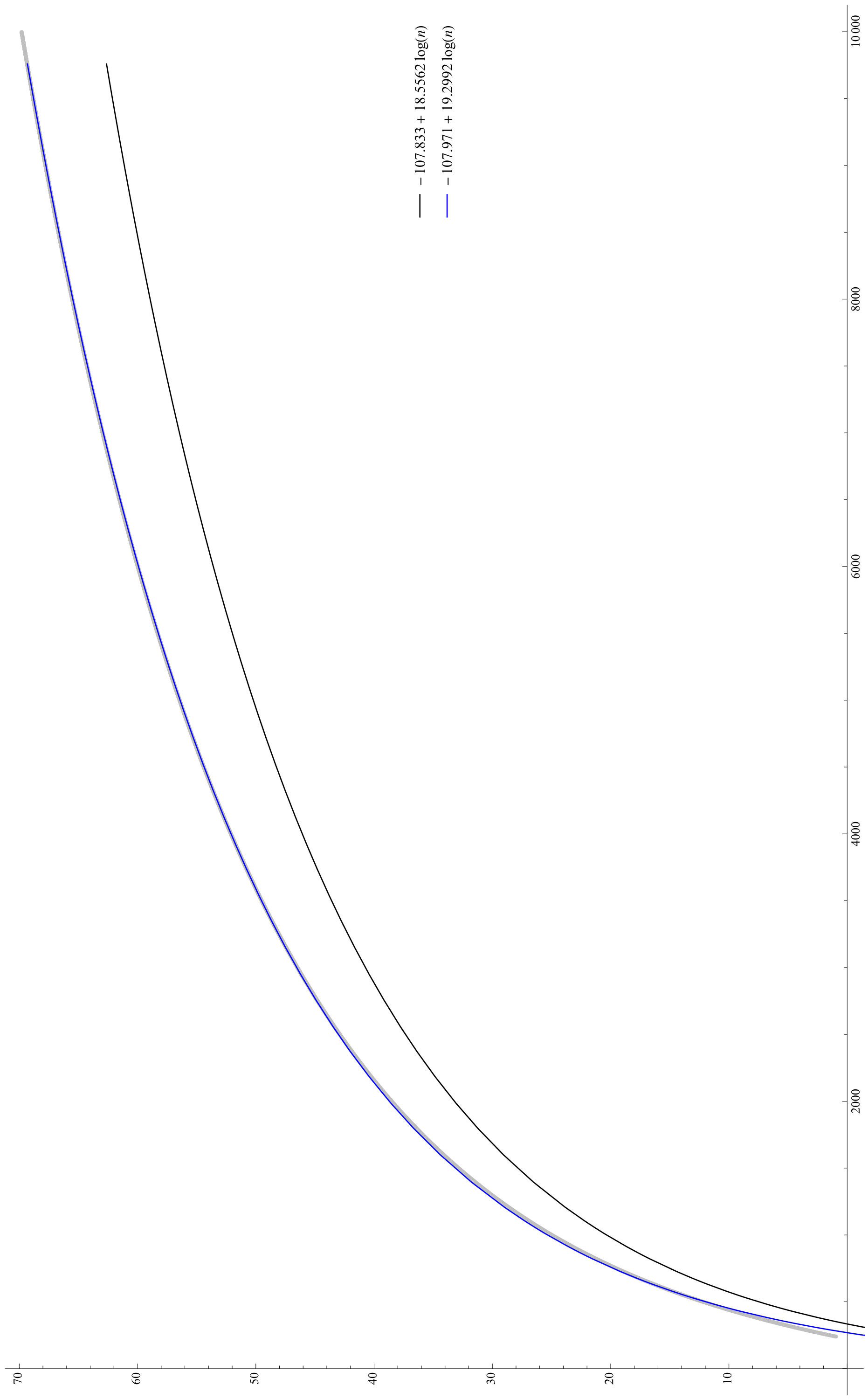

Figure 2.2: Comparing the critical point growth. The gray dots are the numerical approximations for each n, the blue curve is the logarithmic regression, and the black curve is the critical point of the simplified lower bound.

Chapter 3

Space Polygons

Our next goal will be to extend this approach to space polygons. Much of the ground work has been laid out in chapter 2, but we will first need to develop a bound on the total variation between the appropriate multivariate Gaussian, and the upper $r \times s$ block of a matrix sampled uniformly on $U(n)$, the set of unitary $n \times n$ matrices.

Recall that we see that we may sample an n-edge polygon from the symmetric measure on $Arm_3(n)$ by applying the map h to a point sampled from the spherical measure on the sphere of radius $\sqrt{2}$ in $\mathbb{H}^n$, where h first applies the Hopf map, $\mathbf{q} \mapsto \overline{\mathbf{q}}\mathbf{i}\mathbf{q}$, to each coordinate, and then identifies the resulting purely imaginary quaternion with an edge vector in $\mathbb{R}^3$. Likewise, we may sample a closed n-edged polygon under the symmetric measure on $Pol_3(n)$ by applying the map h to the image of a point sampled from the Haar measure on $V_2(\mathbb{C}^n)$ under the embedding $\{\vec{a}, \vec{b}\} \mapsto \vec{a}\mathbf{1} + \vec{b}\mathbf{j}$.

Next, recall that to sample a point on the sphere, it suffices to sample $2n$ complex independent and identically distributed normal random variables and then scale the resultant vector to length of $\sqrt{2}$ [31]. Likewise, we may sample an orthonormal 2-frame from the Haar measure on $V_2(\mathbb{C}^n)$ by applying the Gram-Schmidt orthonormalization procedure to a pair of points sampled from the spherical measure on $S^{2n-1} \subset \mathbb{C}^n$ (pg. 29 [5]). We needn't work out the probability density function for k-edged sub-arms, as it is given in [6], but one still has the intuition that k-edge segments of arms are sampled very closely in total variation to a multi-variate normal distribution. Unfortunately, we will need to establish a total variation bound on the upper $k \times 2$ block of a random unitary matrix to establish an analogue to Theorem 88 for $Pol_3(n)$ and $Arm_3(n)$.

3.1 Variation Bound

Many sources in the literature claim their results for real variates can be extended or adapted to complex or even quaternionic variates, but leave such extensions to the interested reader (see for example [8]). Here we will provide those adaptations to produce the analogue to the variation bound given in [8] for unitary matrices. We wish to stress at this time that the following is simply an adaptation of what is currently found in the literature to our particular problem. After developing the result needed we will return to our own work with polygons. Lastly, throughout this section we will let $\mathscr{L}(*)$ denote "the law of $*$," as is the convention in many of the references.

Definition 98. Given a subspace M of $\mathbb{C}^n$, the compact subgroup $U_n(M) \subset U(n)$ is defined by $U_n(M) = \{g \in U(n) | gx = x \, \text{for all } x \in M\}$.

Definition 99. Since $U_n(M)$ is compact, we may pushforward the Haar measure on $U(n)$ to and then normalize it to produce a measure ν_M on $U_n(M)$.

Definition 100. We say that U is uniform on $U_n(M)$ if it is a random element with law ν_M.

Definition 101. Let P be the orthogonal projection onto the m-dimensional subspace $M \subset \mathbb{C}^n$ and set $Q = I - P$ to be the orthogonal projection onto $M^\perp$. Let r be no larger than $n - m$ and let α be a complex matrix of size $r \times n$. Define $A(M, \alpha) = \alpha Q \alpha^*$. Further, since Q is Hermitian, we see that $A(M, \alpha)$ will be Hermitian. The Spectral Theorem ([14]) then tells us that there exists a unitary matrix U_A and a real diagonal matrix D such that $A(M, \alpha) = U^* D U$. Since $A(M, \alpha)$ is positive semi-definite, we know that all elements of D are non-negative, so it makes since to define the matrix $D^{1/2}$ to be the matrix whose $(i, j)-$entry is the positive square root of the $(i, j)-$entry of D. We then define

$A^{1/2}(M, \alpha) := U^* D^{1/2} U$. In particular, note that

$$A^{1/2}(M, \alpha) A^{1/2}(M, \alpha) = U^* D^{1/2} U U^* D^{1/2} U = U^* D^{1/2} I D^{1/2} U = U^* D U = A.$$

Lemma 102. *Fix an m-dimensional subspace $M \subset \mathbb{C}^n$, and let P be the projection matrix for M. Let U be uniformly distributed on $U(n-m)$ and let Z be the upper left $r \times s$ corner block of U. Let α be an $r \times n$ complex matrix and let β be an $s \times n$ complex matrix, where r and s are no larger than $n-m$. For $A = A(M, \alpha)$ and $B = B(M, \beta)$, and the variate $V = \alpha U \beta^*$, we have $\mathscr{L}(V) = \mathscr{L}(A^{1/2} Z B^{1/2} + \alpha P \beta^*)$. are*

Proof. First, notice that for any m-dimensional subspace M, and any $\Gamma \in U(n)$, the subgroup $U_n(\Gamma M) = \{g \in U(n) : gx = x \text{ for all } x \in \Gamma M\}$, is equal to the subgroup $\Gamma U_n(M) \Gamma^*$. To see this, note that if $g \in U_n(M)$ and $x \in \Gamma M$, then there is a unique $y \in M$ so that $x = \Gamma y$. Then $(\Gamma g \Gamma^*) x = (\Gamma g) y = \Gamma y = x$, so $\Gamma g \Gamma^* \in U_n(\Gamma M)$. Further, if $h \in U_n(\Gamma M)$, and $y \in M$, then $h \Gamma y = \Gamma y$. Multiplying on the left by Γ^* then shows us that $\Gamma^* h \Gamma y = y$, so $\Gamma^* h \Gamma \in U_n(M)$. This then tells us that $h \in \Gamma U_n(M) \Gamma^*$. Together, these give us the relationships that $\Gamma U_n(M) \Gamma^* \subseteq U_n(\Gamma M) \subseteq \Gamma U_n(M) \Gamma^*$ as desired.

Next, since $U_n(\Gamma M) = \Gamma U_n(M) \Gamma^*$, it suffices to establish the lemma in the case where $M = M_0 = \left\{ \vec{z} \in \mathbb{C}^n : \vec{z} = \begin{bmatrix} \vec{x} \\ \vec{0} \end{bmatrix}, \vec{x} \in \mathbb{C}^m \right\}$.

For M_0, it is clear that

$U_n(M_0) = \left\{ g \in O_n \, : \, g = \begin{bmatrix} I_m & O \\ O & h \end{bmatrix}, \, h \in U(n-m) \right\}$. Hence, if U is uniform on $U(n-m)$, then $U_0 = \begin{bmatrix} I_m & O \\ O & U \end{bmatrix}$ is uniform on $U_n(M_0)$.

Set $P_0 = \begin{bmatrix} I_m & O \\ O & O \end{bmatrix}$, the orthogonal projection onto M_0, and set

$Q_0 = I - P_0 = \begin{bmatrix} O & O \\ O & I_{n-m} \end{bmatrix}$. We can write $Q_0 = C_0 C_0^*$, where C_0 is the $n \times (n-m)$ matrix $\begin{bmatrix} O \\ I_{n-m} \end{bmatrix}$. Then for any $V = \alpha U_0 \beta^*$, we have that

$$
\begin{aligned}
V &= \alpha I_n U_0 I_n \beta^* && (3.1)\\
&= \alpha (P_0 + Q_0) U_0 (P_0 + Q_0) \beta^* && (3.2)\\
&= \alpha (P_0 U_0 + Q_0 U_0)(P_0 + Q_0) \beta^* && (3.3)\\
&= \alpha (P_0 U_0 P_0 + Q_0 U_0 P_0 + P_0 U_0 Q_0 + Q_0 U_0 Q_0) \beta^* && (3.4)\\
&= \alpha (P_0 P_0 U_0 + Q_0 P_0 U_0 + U_0 P_0 Q_0 + Q_0 U_0 Q_0) \beta^* && (3.5)\\
&= \alpha (P_0 + O U_0 + U_0 O + Q_0 U_0 Q_0) \beta^* && (3.6)\\
&= \alpha (P_0 + Q_0 U_0 Q_0) \beta^* && (3.7)\\
&= \alpha P_0 \beta^* + \alpha Q_0 U_0 Q_0 \beta^*. && (3.8)
\end{aligned}
$$

In 3.1 we use the identity that $U_0 = I_n U_0 I_n$ and in 3.2 the identity that $I_n = P_0 + Q_0$. Lines 3.3 and 3.4 follow from the distributive property. Line 3.5 comes from the identity that $P_0 U_0 = U_0 P_0 = P_0$. Line 3.6 follows from the identity that $P_0 Q_0 = Q_0 P_0 = O$. We have then that $V = \alpha Q_0 U_0 Q_0 \beta^* + \alpha P_0 \beta^* = \alpha C_0 C_0^* U_0 C_0 C_0^* \beta^* + \alpha P_0 \beta^* = \gamma U \delta^* + \alpha P_0 \beta^*$, where $\gamma = \alpha C_0$ and $\delta = \beta C_0$ and we have used the fact that $C_0^* U_0 C_0 = C_0^* \begin{bmatrix} O \\ U \end{bmatrix} = U$. Now notice that we have $A_0 = \gamma\gamma^* = \alpha Q_0 \alpha^* = A(M_0, \alpha)$, and $B_0 = \delta\delta^* = \beta Q_0 \beta^* = A(M_0, \beta)$. This allows us to write γ and δ in their polar decompositions [16], as $\gamma = A_0^{1/2} \begin{bmatrix} I_r & O \end{bmatrix} \psi_1$, and $\delta = B_0^{1/2} \begin{bmatrix} I_s & O \end{bmatrix} \psi_2$, where $\psi_1, \psi_2 \in U(n-m)$. Recalling that U is uniform on $U(n-m)$, and is thus sampled from the Haar measure, we see that $\mathscr{L}(U) = \mathscr{L}(\psi_1 U \psi_2^*)$ which gives us:

$$
\begin{aligned}
\mathscr{L}(V) &= \mathscr{L}(\alpha Q_0 U_0 Q_0 \beta^* + \alpha P_0 \beta^*) \\
&= \mathscr{L}\left(A_0^{1/2} \begin{bmatrix} I_r & O \end{bmatrix} \psi_1 U \psi_2^* \begin{bmatrix} I_s \\ O \end{bmatrix} B_0^{1/2} + \alpha P_0 \beta^* \right) \\
&= \mathscr{L}\left((A_0^{1/2} \begin{bmatrix} I_r & O \end{bmatrix} U \begin{bmatrix} I_s \\ O \end{bmatrix} B_0^{1/2} + \alpha P_0 \beta^* \right) \\
&= \mathscr{L}(A_0^{1/2} Z B_0^{1/2} + \alpha P_0 \beta^*)
\end{aligned}
$$

Where $Z = \begin{bmatrix} I_r & O \end{bmatrix} U \begin{bmatrix} I_s \\ O \end{bmatrix}$ is the $r \times s$ upper left block of U, as desired. $\square$

We have already established a view of the Stiefel manifold $V_q(\mathbb{C}^n)$ as the set of all $n \times q$ complex matrices A that satisfy $A^* A = I_q$, and the fact that that if Γ is uniform on $U(n)$ then $\Gamma_1 = \Gamma \begin{bmatrix} I_q \\ O \end{bmatrix}$, is uniform on $V_q(\mathbb{C}^n)$.

Definition 103. For a compact group G acting on a measurable space $\mathscr{Y}$, a function $\tau : \mathscr{Y} \to \mathscr{Z}$ is called a maximal invariant function under G if (1) $\tau(gy) = \tau(y)$ for all $y \in \mathscr{Y}$ and $g \in G$ and (2) for any pair of points $y_1, y_2 \in \mathscr{Y}$ such that $\tau(y_1) = \tau(y_2)$, there exists some $g \in \mathscr{Y}$ such that $g y_1 = y_2$.

Proposition 104 (From [13]). *Suppose that G is a compact group that acts on a measurable space $\mathscr{Y}$. Let $\tau : \mathscr{Y} \to \mathscr{Z}$ be a maximal invariant function, and let $Z_i = \tau(Y_i)$ where $\mathscr{L}(Y_i) = P_i$ for $i = 1, 2$ are two G-invariant distributions. If $\mathscr{L}(Z_1) = \mathscr{L}(Z_2)$, then $P_1 = P_2$.*

Next, for $q \leq p$, partition $\Gamma_1 = \begin{bmatrix} \Delta \\ \Psi \end{bmatrix}$, where Δ is $p \times q$ and Ψ is $(n-p) \times q$. Additionally, let $\mathbb{L}_{q,n}$ be the space of all $n \times q$ complex matrices of rank q, and note that $V_q(\mathbb{C}^n) \subsetneq \mathbb{L}_{q,n}$.

Proposition 105. *Suppose $X \in \mathbb{L}_{q,n}$ has a left $U(n)$-invariant distribution. Let $\phi : \mathbb{L}_{q,n} \to V_q(\mathbb{C}^n)$ satisfy $\phi(gx) = g\phi(x)$ for all $x \in \mathbb{L}_{q,n}$ and $g \in U(n)$, which is to say that ϕ is an equivariant map. Then $\mathbb{L}(\phi(X)) = \mathbb{L}(\Gamma_1)$. In other words, the image of any invariant distribution under an equivariant map is the Haar measure on $V_q(\mathbb{C}^n)$.*

Proof. From the uniqueness of the uniform distribution on $V_q(\mathbb{C}^n)$, it suffices to show that $\mathscr{L}(g\phi(X)) = \mathscr{L}(\phi(X))$ for $g \in U(n)$.We have from assumption on ϕ that $\mathscr{L}(g\phi(X)) = \mathscr{L}(\phi(gX))$ and from left $U(n)$-invariance that $\mathscr{L}(\phi(gX)) = \mathscr{L}(\phi(X))$ $\square$

Notice here that a particular such ϕ is given by $\phi(x) = x(x^*x)^{-1/2}$, (the unitary matrix of the polar decomposition of the matrix x, as seen in Lemma 2.1 of [22]), as we have that $\phi(gx) = gx((gx)^*gx)^{-1/2} = gx(x^*g^*gx)^{-1/2} = gx(x^*x)^{-1/2}$.

Proposition 106. *Let $X \in \mathbb{L}_{q,n}$ and partition it into $X = \begin{bmatrix} Y \\ Z \end{bmatrix}$, $Y : p \times q$, $Z : (n-p) \times q$. Then $\mathscr{L}(\Delta) = \mathscr{L}(Y(Y^*Y + Z^*Z)^{-1/2})$, where again Δ is the top $p \times q$ block of Γ_1.*

Proof. We have then, that $X^*X = \begin{bmatrix} Y^* & Z^* \end{bmatrix} \begin{bmatrix} Y \\ Z \end{bmatrix} = Y^*Y + Z^*Z$, so the matrix $Y(Y^*Y + Z^*Z)^{-1/2}$ is the upper $p \times q$ block of $X(X^*X)^{-1/2}$. The result then follows from the previous proposition. $\square$

Proposition 107. *Let $U \in \mathbb{L}_{p,n}$ and partition it into $U = \begin{bmatrix} V \\ W \end{bmatrix}$, $V : q \times p$, $W : (n-q) \times p$. Then $\mathscr{L}(\Delta^*) = \mathscr{L}(V(V^*V + W^*W)^{-1/2})$ and $\mathscr{L}(\Delta) = \mathscr{L}((V^*V + W^*W)^{-1/2}V^*)$.*

Proof. By mirroring the previous proof, we see that $\mathscr{L}(\Delta^*) = \mathscr{L}(V(V^*V + W^*W)^{-1/2})$. Since $V^*V + W^*W$ is Hermitian, so too is its square root. This tells us that $(V(V^*V + W^*W)^{-1/2})^* = (V^*V + W^*W)^{-1/2})V^*$, so we can conclude that $\mathscr{L}(\Delta) = \mathscr{L}((V^*V + W^*W)^{-1/2}V^*)$. □

We know have the tools needed to find explicitly the density of these distributions. First, we will define the densities we will be using:

Definition 108 (From [17])**.** For a matrix distribution Y, whose n rows are independent and identically distributed p-variate complex Gaussian random variables with covariance matrix Σ. Then the distribution of $Y^*Y = \sum_{k=1}^{n} Y_i Y_i^*$, has the probability density function given by:

$$p_W(A) = \frac{\det(A)^{n-p}}{\pi^{\frac{1}{2}p(p-1)}\Gamma(n)\cdots\Gamma(n-p+1)\det(\Sigma)^n} e^{-\operatorname{tr}(\Sigma^{-1}A)},$$

defined on the set of Hermitian positive semi-definite $p \times p$ matrices. This distribution is known as the Complex Wishart distribution and we will denote it as $\mathscr{CW}(p, n, \Sigma)$

Next, we point out that in [11] matrices with the above distribution are said to have the complex matrix variate gamma distribution $\mathscr{CG}_p(n, \Sigma)$.

Definition 109 (From [11])**.** Define the complex matrix variate beta type I distribution as follows: for $A \sim \mathscr{CG}_m(a, I_m) = \mathscr{CW}(m, a, I_m)$ and $B \sim \mathscr{CG}_m(b, I_m) = \mathscr{CW}(m, b, I_m)$, the distribution of either of

$$(1)\ U = (A+B)^{-1/2}A((A+B)^{-1/2}))$$

$$(2)\ V = A^{1/2}(A+B)^{-1}(A^{1/2}).$$

Further, the density function of this distribution, denoted as $\mathscr{CBI}_m(U; a, b)$, is given by:

$$p_B(M) = \frac{\mathscr{C}\Gamma_m(a+b)}{\mathscr{C}\Gamma_m(a)\mathscr{C}\Gamma_m(b)} \det(M)^{a-m}\det(I_m - M)^{b-m},$$

defined on the set of $m \times m$ Hermitian positive semi-definite matrices M, where $\mathscr{C}\Gamma_m(a)$ stands for $\pi^{m(m-1)/2} \prod_{j=1}^{m} \Gamma(a-j+1)$.

Proposition 110. *$\mathscr{L}(\Delta^*\Delta) = \mathscr{CBI}_q(p, n-p)$ and $\mathscr{L}(\Delta\Delta^*) = \mathscr{CBI}_p(q, n-q)$. Further, $\mathscr{L}(\Delta^*\Delta)$ has a density given by:*

$$p(\Delta^*\Delta) = \frac{\mathscr{C}\Gamma_q(n)}{\mathscr{C}\Gamma_q(p)\mathscr{C}\Gamma_q(n-p)} \det(\Delta^*\Delta)^{p-q} \det(I_q - \Delta^*\Delta)^{n-p-q}$$

Proof. Let X be distributed as $N(0, I_n \otimes I_q)$, and be partitioned as $X = \begin{bmatrix} Y \\ Z \end{bmatrix}$.

From our definition of the Complex Wishart distribution,
$\mathscr{L}(Y^*Y) = \mathscr{CW}(q, p, I_q) = \mathscr{CG}_q(p, I_q)$ and $\mathscr{L}(Z^*Z) = \mathscr{CW}(q, n-p, I_q) = \mathscr{CG}_q(n-p, I_q)$. Next, we see from Proposition 107 that
$\mathscr{L}(\Delta\Delta^*) = \mathscr{L}(((Y^*Y + Z^*Z)^{-1/2})Y^*Y(Y^*Y + Z^*Z)^{-1/2})$. Finally, from the definition of the complex matrix variate beta type I distribution, since this is in the form
$U = (A+B)^{-1/2}A((A+B)^{-1/2}))$ for $A = Y^*Y \sim \mathscr{CG}_q(p, I_q)$ and $B = Z^*Z \sim \mathscr{CG}_q(n-p, I_q)$, we see that $\Delta\Delta^*$ has a distribution of type $\mathscr{CBI}_p(q, n-q)$. Likewise, we see that
$\Delta^*\Delta \sim \mathscr{CBI}_q(p, n-p)$ and a density function given by:

$$\begin{aligned} p(\Delta^*\Delta) &= \frac{\mathscr{C}\Gamma_q(p+n-p)}{\mathscr{C}\Gamma_q(p)\mathscr{C}\Gamma_q(n-p)} \det(\Delta^*\Delta)^{p-q} \det(I_q - \Delta^*\Delta)^{n-p-q} \\ &= \frac{\mathscr{C}\Gamma_q(n)}{\mathscr{C}\Gamma_q(p)\mathscr{C}\Gamma_q(n-p)} \det(\Delta^*\Delta)^{p-q} \det(I_q - \Delta^*\Delta)^{n-p-q} \end{aligned}$$

□

Theorem 111 (From [37])**.** *For a complex matrix M of size $p \times q$, if the density of M depends only on the matrix $B = M^*M$, by a function $f(B)$, then the density of $B = M^*M$ is given by* $\dfrac{f(B)\det(B)^{p-q}\pi^{q(p-(1/2)(q-1))}}{\prod_{j=1}^{q}\Gamma(p-j+1)}$

Proof. This is the main theorem of [37], so the proof is omitted here. □

We know have the tools needed to determine the probability density function of Δ:

Theorem 112. *For the the upper $p \times q$ block of Γ_1, called Δ, the density of Δ is given by $f(\Delta) = c_1 |I_q - \Delta^* \Delta|^{n-p-q}$, where c is the constant given by*

$$c_1 = \pi^{qp} \prod_{j=1}^{q} \left(\frac{\Gamma(n-j+1)}{\Gamma(n-p-j+1)} \right).$$

Proof. It follows from the Proposition 110 that $\Delta^*\Delta$ has a density given by $\mathscr{CBI}_q(p, n-p)$. First, we have that the distribution of Δ is invariant under the action of $U(p)$ given by left multiplication, $\Delta \to g\Delta, g \in U(p)$. Second, we have a maximal invariant given by $\tau(\Delta) = \Delta^*\Delta$. Let Ψ be the random matrix variate with density given by f. $U(p)$ acts on Ψ, and the density of the maximal invariant $\tau(\Psi)$ is then calculated from Theorem 111 as

$$\begin{aligned}
h(\Psi^*\Psi) &= \frac{c_1 \det(I_q - \Psi^*\Psi)^{n-p-q} \det(\Psi^*\Psi)^{p-q} \pi^{q(p-(1/2)(q-1))}}{\prod_{j=1}^{q} \Gamma(p-j+1)} \\
&= \pi^{qp} \prod_{j=1}^{q} \left(\frac{\Gamma(n-j+1)}{\Gamma(n-p-j+1)} \right) \frac{\det(I_q - \Psi^*\Psi)^{n-p-q} \det(\Psi^*\Psi)^{p-q} \pi^{q(p-(1/2)(q-1))}}{\prod_{j=1}^{q} \Gamma(p-j+1)} \\
&= \pi^{-q(q-1)/2} \prod_{j=1}^{q} \frac{\Gamma(n-j+1)}{\Gamma(p-j+1)\Gamma(n-p-j+1)} \det(\Psi^*\Psi)^{p-q} \det(I_q - \Psi^*\Psi)^{n-p-q},
\end{aligned}$$

This calculation shows that $\mathscr{L}(\Psi^*\Psi) = \mathscr{CBI}_q(p, n-p)$, so we see that $\mathscr{L}(\Psi^*\Psi) = \mathscr{L}(\Delta^*\Delta)$. Since we can see that the distribution of Ψ is invariant under the group action of $U(p)$, it follows from Proposition 104 that $\mathscr{L}(\Psi) = \mathscr{L}(\Delta)$. Hence, f must be the density of Δ. $\square$

Theorem 113. *Let Z be the upper left $r \times s$ block of a random matrix U which is uniform on $U(n)$, so that it has density given by Theorem 112. Further, we have that $EZ = O \in \mathscr{M}_{r,s}(\mathbb{C})$ and $\mathrm{Cov}(Z) = n^{-1} I_r \otimes I_s$, so we shall take X to be a random matrix* *with the $r \times s$ complex multivariate Gaussian distribution with the same mean and covariace. Then, provided that $r+s+2 < n$, the variation distance between $\mathscr{L}(Z)$ and $\mathscr{L}(X)$ is bounded above by $B(r,s;n) := 2\left(\left(1 - \dfrac{r+s}{n}\right)^{-t^2} - 1\right)$, where $t = \min(r,s)$.*

Proof. Setting $\mathscr{L}(X) = P_1$ and $\mathscr{L}(Z) = P_2$, let us start with the case of $s \leq r$. The density f_1 of P_1 is given by $f(x) = \frac{1}{\pi^{rs}} e^{-tr(x^*x)}$ [17]. The density f_2 of P_2 is given in Theorem 112. Since these are functions of x^*x and z^*z respectively, the variation distance is equal to the variation distance between the distributions of x^*x and z^*z. x^*x has, in accordance with the definition above, the complex Wishart distribution $\mathscr{CW}\left(s, r, \frac{1}{n} I_s\right)$, and hence a density given by

$$
\begin{aligned}
f(v) &= \frac{\det(v)^{r-s}}{\pi^{\frac{1}{2}s(s-1)}\Gamma(r)\cdots\Gamma(r-s+1)\det(\frac{1}{n}I_s)^r} e^{-\operatorname{tr}((\frac{1}{n}I_s)^{-1}v)} \\
&= \frac{\det(v)^{r-s}}{\pi^{\frac{1}{2}s(s-1)}\Gamma(r)\cdots\Gamma(r-s+1)n^{-sr}} e^{-n\operatorname{tr}(v)} \\
&= \det(v)^{r-s} e^{-n\operatorname{tr}(v)} \pi^{-\frac{1}{2}s(s-1)} \frac{n^{rs}}{\prod_{j=1}^{s}\Gamma(r-j+1)},
\end{aligned}
$$

defined on the set of $s \times s$ Hermitian, positive-definite matrices. The density of z^*z we have seen in Proposition 110 to be given by

$$
\begin{aligned}
g(v) &= \frac{\mathscr{C}\Gamma_s(n)}{\mathscr{C}\Gamma_s(r)\mathscr{C}\Gamma_s(n-r)} \det(v)^{r-s}\det(I_s - v)^{n-r-s} \\
&= \frac{\pi^{\frac{1}{2}s(s-1)}\prod_{j=1}^{s}\Gamma(n-j+1)}{(\pi^{\frac{1}{2}s(s-1)}\prod_{j=1}^{s}\Gamma(r-j+1))(\pi^{\frac{1}{2}s(s-1)}\prod_{j=1}^{s}\Gamma(n-r-j+1))} \det(v)^{r-s}\det(I_s - v)^{n-r-s} \\
&= \det(v)^{r-s}\det(I_s - v)^{n-r-s}\pi^{-\frac{1}{2}s(s-1)} \prod_{j=1}^{s} \frac{\Gamma(n-j+1)}{\Gamma(r-j+1)\Gamma(n-r-j+1)},
\end{aligned}
$$

defined on the set of matrices with v and $I - v$ positive definite. By an alternate characterization of total variation (seen in [8]), we see that the total variation distance is given by $\delta_{r,s,n} := \int |g(v) - f(v)| dv = 2\int_E \left(\frac{g(v)}{f(v)} - 1\right) f(v) dv$, where E is the set of $s \times s$ positive definite matrices on which $g(v) > f(v)$. As we will be using it often, let us now simplify the expression $\dfrac{g(v)}{f(v)}$:

$$\begin{aligned}
\frac{g(v)}{f(v)} &= \frac{\det(v)^{r-s}\det(I_s-v)^{n-r-s}\pi^{-\frac{1}{2}s(s-1)}\prod_{j=1}^{s}\dfrac{\Gamma(n-j+1)}{\Gamma(r-j+1)\Gamma(n-r-j+1)}}{\det(v)^{r-s}e^{-n\,\mathrm{tr}(v)}\pi^{-\frac{1}{2}s(s-1)}\dfrac{n^{rs}}{\prod_{j=1}^{s}\Gamma(r-j+1)}} \\
&= \frac{\det(I_s-v)^{n-r-s}}{e^{-n\,\mathrm{tr}(v)}n^{rs}}\prod_{j=1}^{s}\frac{\Gamma(n-j+1)}{\Gamma(n-r-j+1)} \\
&= \det(I_s-v)^{n-r-s}e^{n\,\mathrm{tr}(v)}\prod_{j=1}^{s}\frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)}
\end{aligned}$$

Hence, $\delta_{r,s,n} \leq 2\sup_{v\in E}\left(\frac{g(v)}{f(v)}-1\right)$. Set $M_{r,s,n} := \sup_{v\in E}\left(\frac{g(v)}{f(v)}-1\right)$, so that $\delta_{r,s,n} \leq 2\, M_{r,s,n}$. Differentiation shows that the maximum of $\left(\frac{g(v)}{f(v)}-1\right)$ is attained uniquely for $v = \dfrac{r+s}{n}I_s$:

Let us first write $\dfrac{g(v)}{f(v)} = c\det(I_s-v)^{n-r-s}e^{n\,\mathrm{tr}(v)}$, with $c = \prod_{j=1}^{s}\dfrac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)}$ independent of v. Next, computing the derivative with respect to v, we will look at first at the partials from the entries off the diagonal, and secondly at the entries of the diagonal.

Case 1: $(i \neq j)$ In this case, we first note that $\dfrac{\partial}{\partial v_{i,j}}e^{n\,\mathrm{tr}(v)} = 0$, as the trace depends only on the diagonal. This tells us that $\dfrac{\partial}{\partial v_{i,j}}\dfrac{g(v)}{f(v)} = ce^{n\,\mathrm{tr}(v)}\dfrac{\partial}{\partial v_{i,j}}(\det(I_s-v))^{n-r-s}$. Applying the Power Rule and Chain Rule ([39]), we see that $\dfrac{\partial}{\partial v_{i,j}}(\det(I_s-v))^{n-r-s} = (n-r-s)(\det(I_s-v))^{n-r-s-1}\dfrac{\partial}{\partial v_{i,j}}\det(I_s-v)$. Next, we see from 2.1.1 of [34] that $\dfrac{\partial}{\partial v_{i,j}}\det(I_s-v) = \det(I_s-v)\,\mathrm{tr}\left((I_s-v)^{-1}\dfrac{\partial}{\partial v_{i,j}}(I_s-v)\right)$. Here, $\dfrac{\partial}{\partial v_{i,j}}(I_s-v)$ is a matrix whose only non-zero entry the (i,j)-entry, which is a -1. Hence, we see that that $\mathrm{tr}\left((I_s-v)^{-1}\dfrac{\partial}{\partial v_{i,j}}(I_s-v)\right)$ is (j,i)-entry of $-(I_s-v)^{-1}$. Recall that for an invertible matrix M, $M^{-1} = \dfrac{1}{\det(M)}\,\mathrm{adj}(M) = \dfrac{1}{\det(M)}\,\mathrm{C}(M)^{\intercal}$, where $\mathrm{adj}(M)$ is the adjoint matrix and $\mathrm{C}(M)$ is the cofactor matrix (3.1.2 and 3.1.4 of [34]). Therefore, we see that the

(j,i)-entry of $-(I_s - v)^{-1}$ is the (i,j)-entry of $\dfrac{-1}{\det(I_s - v)}\,\mathrm{C}(I_s - v)$. We may then conclude that $\dfrac{\partial}{\partial v_{i,j}}\dfrac{g(v)}{f(v)} = -ce^{n\,\mathrm{tr}(v)}(n-r-s)(\det(I_s - v))^{n-r-s-1}\,\mathrm{C}(I_s - v)_{\{i,j\}}$. We can then see that this will only be zero when $\mathrm{C}(I_s - v)_{\{i,j\}}$ is zero, as the first three terms are all positive, and the determinant term is non-zero as $g(v)$ is only defined on the set of matrices with both v and $I_s - v$ positive definite.

Case 2: $(i = j)$. We first apply the Product Rule to see that

$$\frac{\partial}{\partial v_{i,i}}\frac{g(v)}{f(v)} = c\left(e^{n\,\mathrm{tr}(v)}\left(\frac{\partial}{\partial v_{i,i}}(\det(I_s - v))^{n-r-s}\right) + (\det(I_s - v))^{n-r-s}\left(\frac{\partial}{\partial v_{i,i}}e^{n\,\mathrm{tr}(v)}\right)\right).$$

We have already computed the partial derivative of the power of the determinant. In the second term, we see from a quick application of the chain rule that $\dfrac{\partial}{\partial v_{i,i}}e^{n\,\mathrm{tr}(v)} = ne^{n\,\mathrm{tr}(v)}$. We may then conclude that

$$\frac{\partial}{\partial v_{i,i}}\frac{g(v)}{f(v)} = ce^{n\,\mathrm{tr}(v)}(\det(I_s - v))^{n-r-s-1}\left(-(n-r-s)\,\mathrm{C}(I_s - v)_{\{i,i\}} + n\det(I_s - v)\right).$$

This will be zero only when $n\det(I_s - v) = (n-r-s)\,\mathrm{C}(I_s - v)_{\{i,i\}}$.

We have now classified the critical point of $\dfrac{g(v)}{f(v)}$ to be any matrix v for which the (i,j) cofactor of $(I_s - v)$ is given by the equation $\dfrac{n}{n-r-s}\det(I_s - v)\delta_{i,j}$, where $\delta_{i,j}$ is the Kronecker delta. We have already seen how to express the inverse of a matrix in terms of the determinant and the cofactor matrix, so since we know all of the cofactors of $I_s - v$, we know the inverse of $I_s - v$. Specifically, $\left[(I_s - v)^{-1}_{i,j}\right] = \dfrac{1}{\det(I_s - v)}\left[\dfrac{n}{n-r-s}\det(I_s - v)\delta_{j,i}\right]$. Observe that the matrix on the right-hand side of the equation is simply the identity matrix scaled by $\dfrac{n}{n-r-s}$. Inverting both sides produces $I_s - v = \dfrac{n-r-s}{n}I_s$, so we see that $v = \left(1 - \dfrac{n-r-s}{n}\right)I_s = \dfrac{r+s}{n}I_s$.

Now that we see that this is the only critical point, we will show that it produces a maximum. All of the following properties are given in [4]. First, recall that a critical point of a concave function must be a maximum. Second, note that if $\phi(x)$ is convex, then so too are $\alpha\phi(x)$, $\phi(x+t)$, and $\phi(Ax)$ for any $\alpha > 0, t \in \mathbb{R}^M$, and $M \times M$ matrix A and $-\phi(x)$ is concave. Fourth, we know that the sum, product, and composition of convex functions are convex. From this last property, we see that a concave function pre-composed with a convex function is concave and the product of a convex function and a concave function is concave. Using these properties, it is easy to see that the trace of a matrix is convex, as it is the sum of the projections to the the diagonal elements. Likewise, from the fact that $\frac{d^2}{d^2x}e^{\alpha x} = \alpha^2 e^{\alpha x}$, we know that $e^{\alpha x}$ is convex, showing that $ce^{n\,\mathrm{tr}(v)}$ is convex. Now, we need only show that $\det(v)$ is concave to show the concavity of $\dfrac{f(v)}{g(v)} - 1$, which is given as Example 3.39 of [4].

Hence, we know that

$$
\begin{aligned}
M_{r,s,n} + 1 &= \frac{g((r+s)n^{-1}I_s)}{f((r+s)n^{-1}I_s)} \\
&= \det(I_s - (r+s)n^{-1}I_s)^{n-r-s} e^{n\,\mathrm{tr}((r+s)n^{-1}I_s)} \prod_{j=1}^{s} \frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)} \\
&= \det\left(\left(1 - \frac{r+s}{n}\right) I_s\right)^{n-r-s} e^{ns(\frac{r+s}{n})} \prod_{j=1}^{s} \frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)} \\
&= \left(1 - \frac{r+s}{n}\right)^{s(n-r-s)} e^{s(r+s)} \prod_{j=1}^{s} \frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)} \\
&= \prod_{j=1}^{s} \left(\frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)} \left(1 - \frac{r+s}{n}\right)^{n-r-s} e^{r+s} \right)
\end{aligned}
$$

We would now like to write this in terms of logarithms. To do this, we first observe that

$$
\begin{aligned}
-n\int_0^t \ln(1-x)\,\mathrm{d}x &= -n\left((x-1)\ln(1-x) - x\right)_{x=0}^{x=t} \\
&= -n((t-1)\ln(1-t) - t) \\
&= nt + (n - nt)\ln(1-t).
\end{aligned}
$$

Setting $t = \frac{r+s}{n}$ gives us $-n\int_0^{\frac{r+s}{n}} \ln(1-x)\,\mathrm{d}x = (r+s) + (n-r-s)\log\left(1 - \dfrac{r+s}{n}\right)$.

Next, set

$$A_j = \ln\left(\frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)}\right) - n\int_0^{(r+s)/n} \ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right),$$

we can write $M_{r,s,n}+1 = \prod_{j=1}^{s} e^{A_j}$. Now let us write A_j in a more pliable form by noting that

$$\begin{aligned}
\log\left(\frac{\Gamma(n-j+1)}{n^r\Gamma(n-r-j+1)}\right) &= \ln(\Gamma(n-j+1)) - \ln(\Gamma(n-r-j+1)) - \ln(n^r) &(3.9)\\
&= \left(\sum_{i=1}^{n-j}\ln(i)\right) - \left(\sum_{i=1}^{n-r-j}\ln(i)\right) - \left(\sum_{i=1}^{r}\ln(n)\right) &(3.10)\\
&= \left(\sum_{i=n-j-r+1}^{n-j}\ln(i)\right) - \left(\sum_{i=1}^{r}\ln(n)\right) &(3.11)\\
&= \left(\sum_{k=1}^{r}\ln(n-j-k+1)\right) - \left(\sum_{i=1}^{r}\ln(n)\right) &(3.12)\\
&= \sum_{i=1}^{r}\ln\left(\frac{n-j-i+1}{n}\right) &(3.13)\\
&= \sum_{i=1}^{r}\ln\left(1-\frac{j+i-1}{n}\right). &(3.14)
\end{aligned}$$

In line 3.10, we have used the fact that for $x \in \mathbb{N}$, $\Gamma(x) = \prod_{i=1}^{x-1} i$. In line 3.12, we introduce the change of indices $k = (n-j+1)-i$, which ranges from 1 when $i = n-j$ to r when $i = n-j-r+1$.

This lets us simplify A_j into the form:

$$A_j = \left(\sum_{i=1}^{r}\log\left(1-\frac{j+i-1}{n}\right)\right) - n\int_0^{(r+s)/n}\ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right).$$

Writing A_j in this way as sum of three quantities, it is easy to see that $A_j \le A_1$ for all $j = 1, 2, \ldots, s$: Only the first depends on j, and as j increases, $1 - \dfrac{j+i-1}{n}$ is decreasing, so that $\log\left(1-\dfrac{j+i-1}{n}\right)$ is decreasing. This allows us to to bound

$M_{r,s,n} + 1 \leq \prod_{j=1}^{s} e^{A_1} = e^{sA_1}$. Next, we claim that $-\log(1-x)$ is an increasing convex function on $[0,1)$. To see this, first, we note that the first derivative, $\dfrac{1}{1-x}$, is strictly positive for all $x \in [0,1)$, while the second derivative, $\dfrac{-1}{(1-x)^2}$, is strictly negative. Next, recall that the graph of a convex function $h(x)$ on any interval $[a,b]$ lies below the graph of the secant line from $(a, f(a))$ to $(b, f(b))$. Next, let $l_{[a,b]}(x) = \dfrac{\ln(1-a) - \ln(1-b)}{b-a}(x-a) - \ln(1-a)$ be the function whose graph is the secant line of $-\ln(1-x)$ from $(a, -\ln(1-a))$ to $(b, -ln(1-b))$. We then have the inequality that $0 \leq -\ln(1-x) \leq l_{[a,b]}(x)$ for any

$0 \leq a < x < b < 1$. In particular, monotonicity of integration tells us then that

$0 \leq -\int_a^b \log(1-x)\mathrm{d}x \leq \int_a^b l_{[a,b]}(x)\mathrm{d}x = \dfrac{b-a}{2}(-\ln(1-b) - \ln(1-a))$.

Setting $a = \dfrac{i-1}{n}$ and $b = \dfrac{i}{n}$, we then have that

$-n\int_{(i-1)/n}^{i/n} \log(1-x)\mathrm{d}x \leq \dfrac{1}{2}\left(-\log(1-\frac{i}{n}) - \log(1-\frac{i-1}{n})\right)$. Which we can write as

$\dfrac{1}{2}\log(1-\dfrac{i}{n}) \leq n\int_{(i-1)/n}^{i/n} \log(1-x)\mathrm{d}x - \dfrac{1}{2}\log(1-\frac{i-1}{n})$. We now have the tools to bound A_1 nicely:

$$A_1 = \left(\sum_{i=1}^{r} \ln\left(1-\frac{i}{n}\right)\right) - n\int_0^{(r+s)/n} \ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right) \tag{3.15}$$

$$= \left(2\sum_{i=1}^{r} \frac{1}{2}\ln\left(1-\frac{i}{n}\right)\right) - n\int_0^{(r+s)/n} \ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right) \tag{3.16}$$

$$\leq \left(\sum_{i=1}^{r} \frac{1}{2}\ln\left(1-\frac{i}{n}\right)\right) + \left(\sum_{i=1}^{r} n\int_{(i-1)/n}^{i/n} \ln(1-x)\mathrm{d}x - \frac{1}{2}\ln\left(1-\frac{i-1}{n}\right)\right) \tag{3.17}$$

$$- n\int_0^{(r+s)/n} \ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right) \tag{3.18}$$

$$= \left(\sum_{i=1}^{r} \frac{1}{2}\ln\left(1-\frac{i}{n}\right) - \frac{1}{2}\ln\left(1-\frac{i-1}{n}\right)\right) + \left(\sum_{i=1}^{r} n\int_{(i-1)/n}^{i/n} \ln(1-x)\mathrm{d}x\right) \tag{3.19}$$

$$- n\int_0^{(r+s)/n} \ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right) \tag{3.20}$$

$$= \frac{1}{2}\ln\left(1-\frac{r}{n}\right) - n\int_{r/n}^{(r+s)/n} \ln(1-x)\mathrm{d}x + \ln\left(1-\frac{r+s}{n}\right) \tag{3.21}$$

$$\leq \frac{1}{2}\ln\left(1-\frac{r}{n}\right) - \frac{s+1}{2}\left(\ln\left(1-\frac{r}{n}\right) + \ln\left(1-\frac{r+s}{n}\right)\right) + \ln\left(1-\frac{r+s}{n}\right) \tag{3.22}$$

$$= -\frac{s}{2}\ln\left(1-\frac{r}{n}\right) - \frac{s-1}{2}\ln\left(1-\frac{r+s}{n}\right) \tag{3.23}$$

$$\leq -\left(\frac{s}{2}+\frac{s-1}{2}\right)\ln\left(1-\frac{r+s}{n}\right) \tag{3.24}$$

$$\leq -s\ln\left(1-\frac{r+s}{n}\right) \tag{3.25}$$

In lines 3.17-3.18, we have applied the bound we obtained form the convexity argument to one of the sums of $\frac{1}{2}\ln\left(1-\frac{i}{n}\right)$. In lines 3.19-3.20, we combine the sums of the logarithms, in preparation to evaluate the single telescoping sum in lines 3.21. In lines 3.22, we use again the convexity argument to bound the integral by the sum of two logarithms before collecting terms in 3.23. In line 3.24, we use the fact that $-\ln(1-x)$ is increasing. Finally, in line 3.25, since $-\ln\left(1-\frac{r+s}{n}\right) > 0$, we use the slightly simpler upper bound for $\left(s-\frac{1}{2}\right)$.

We then have that $M_{r,s,n} + 1 \leq e^{-s^2\ln(1-(r+s)/n)} = \left(1-\frac{r+s}{n}\right)^{-s^2}$. Hence, we have that $\delta_{r,s,n} \leq 2\left(\left(1-\frac{r+s}{n}\right)^{-s^2} - 1\right)$. To finish the proof, in the case that $r \leq s$, we repeat these arguments with their roles reversed. This brings us to the promised form: $\delta_{r,s,n} \leq 2\left(\left(1-\frac{r+s}{n}\right)^{-(\min(r,s)^2)} - 1\right) = B(r,s;n)$. □

Now that we are done with the detour through random matrix theory, let us return our focus to random polygons.

Theorem 114. *Let f be an essentially bounded k-edged locally defined function. Then the expectation of f over $Pol_3(n)$ may be approximated by the expectation of f over $Arm_3(n)$ to within $M\mathscr{B}_3(k,n)$, where M is a bound for f almost everywhere, and*

$$\mathscr{B}_3(k,n) := B(k,2;n) = 2\left(\frac{4k+3}{4n-4k-3} + \frac{n^4}{(n-k-2)^4} - 1\right)$$

Theorem 115. *Let f be an essentially bounded, k-edged locally defined function. Then the expectation of f over $Pol_2(n)$ may be approximated by the expectation of f over $Arm_2(n)$ to within $M\mathscr{B}_2(k,n)$, where M is a bound for f almost everywhere.*

Corollary 116. *Let q be an essentially bounded, locally measured quantity of a polygonal chain. Let $E_p(n)$ stand for the expectation of q over $Pol_3(n)$, and $E_a(n)$ stand for the expectation of q over $Arm_3(n)$. If $nE_a(n) \to \infty$, then $\dfrac{E_p(n)}{E_a(n)} \to 1$.*

Moreover, the expectation of the average of q over the polygon, $\widetilde{E_p}(n)$ and the expectation of the sum of q over the polygon, $\overline{E_p}(n)$ also satisfy $\dfrac{\widetilde{E_p}(n)}{\widetilde{E_a}(n)} \to 1$ and $\dfrac{\overline{E_p}(n)}{\overline{E_a}(n)} \to 1$.

Proof. The proofs of these mirror those given in for Theorem 88, Theorem 91 and Corollary 92, where we replace $V_2(\mathbb{R}^n)$ with $V_2(\mathbb{C}^n)$, S^{2n} with S^{4n}, and our variation bounds from the shared, close-proximity multivariate Gaussian, come from Theorem 113 and Theorem 87 respectively. $\square$

Looking at this bound, we see that, as with the planar case, it is limiting to 0 at a rate of $O(n)$. Specifically, it is asymptotic to $\lim_{n\to\infty} n\mathscr{B}_3(k,n) = 10k + \dfrac{35}{2}$, and we additionally have again that $\mathscr{B}_3(k,n) < \dfrac{10k+17.5}{n}$, provided that the bound is useful ($\mathscr{B}_2(k,n)$ is greater than 2 for $k > \frac{n}{5}$). When $k = o(n^p)$ with $0 < p < 1$, this is limiting to 0. On the other hand, when $k = \alpha n$, $\mathscr{B}_3(\alpha n, n)$ is limiting to $\dfrac{2}{1-\alpha} + \dfrac{2}{(1-\alpha)^4} - 2$, which is greater than 1 for $\alpha > 0.08235533$. In other words, provided that the the number of edges k is less than 8% of n, we are able the distributions of k-edged segments coming from $Arm_3(n)$ are close in total variation to those coming from $Pol_3(n)$.

3.2 Torsion

In [6], we can see that the analogous integral to find the expected total curvature with respect to the symmetric measure on $Pol_3(n)$ and the Hopf-Gaussian measure on $\mathscr{P}_3(n)$ is much tamer than for planar polygons. Indeed, it is explicitly computed to give $E(\kappa; Pol_3(n), \sigma) = \frac{\pi}{2}n + \frac{\pi}{4}\frac{2n}{2n-3}$. Let us now then attempt to solve the problem of finding the expected total torsion.

Definition 117. For a polygon in $\mathbb{R}^d$, we define the torsion angle (sometimes called the dihedral angle) at an edge e_i by the following procedure: Let p_i be the plane which is normal to e_i at v_i. Project edges e_{i-1} and e_{i+1} to p_i along v_i to get a 2-edge planar polygonal arm in p_i with middle vertex v_i. The torsion angle is then defined as the angle between these edges, with the convention that we take its value in the range $(-\pi, \pi]$.

Proposition 118. *The distribution of the torsion angle for arms is the same as the distribution of $\pi - \theta$, where θ is the polar angle in the spherical coordinates of the edges.*

Proof. Write $e_i = (r_i, \theta_i, \phi_i)$ in spherical coordinates. Rotate the configuration so that e_i is on the z-axis. If we then further rotate so that e_{i-1} has no y-component, we can see that the projections to p_i (the xy-plane) form a planar 2-edge arm that runs along the negative x-axis, then turns to form an edge given in polar coordinates as $(\tilde{r}_{i+1}, \theta_{i+1})$. As such, the torsion angle of the rotated configuration will be given by $\pi - \theta_{i+1}$. Since the distribution of arms is invariant under the $SO(3)$ action on $\mathbb{R}^3$, the result follows. $\square$

Proposition 119. *The expectation of the torsion angle of a polygonal arm sampled under the symmetric measure on $Arm_3(n)$ or $\mathscr{A}_3(n)$ is* 0.

Proof. Similar to how we found the expectation of curvature, since we know that the symmetric measure is expressible as a product measure on $\mathbb{R}^n \times (S^2)^n$, with the spherical measure on the individual copies of S^2, we see that the distribution of the polar angles will be uniform on $[0, 2\pi)$, so the expectation of $\pi - \theta$ will be 0. $\square$

For polygons, we have an integral even more imposing than the one for planar polygon's curvature. So this is an excellent opportunity to use the total variation bound. Using

Theorem 114, we see that $|E(\tau_i) - 0| \leq \pi\mathscr{B}_3(3,n)$. This give us bounds on total torsion of $-n\pi\mathscr{B}_3(3,n) \leq E(\tau) \leq n\pi\mathscr{B}_3(3,n)$. This is limiting to the range of $[-55.5\pi, 55.5\pi]$. However, unlike total curvature, the expectation of total torsion over $Arm_3(n)$ is 0 for any n. As such, $\lim_{n\to\infty} nE_{Arm_3(n)}(\tau) = 0$, so we may not apply Corollary 116. Nonetheless, we may hope to glean useful information by considering the variance.

Proposition 120. *Where τ_i is the torsion angle at edge e_i of a polygon sampled under the symmetric measure on $Arm_3(n)$ or $\mathscr{A}_3(n)$, we have that* $\operatorname{Cov}(\tau_i, \tau_j) = \delta_{i,j}\frac{1}{3}\pi^2$.

Proof. We can easily see from the independence of directions and our earlier description of the dihedral angle, that the covariance of any distinct pair of dihedral angles will be 0. So let us focus on $\operatorname{Cov}(\tau_i, \tau_i)$ where we have:

$$\begin{aligned}
\operatorname{Cov}(\tau_i, \tau)i) &= \operatorname{Var}(\tau_i) \\
&= \int_{Arm_3(n)} (\tau_i - 0)^2 \, \mathrm{d}\sigma \\
&= \frac{1}{2\pi}\int_0^{2\pi} (\pi - \theta_i)^2 \, \mathrm{d}\theta_i \\
&= \frac{-1}{6\pi}(\pi - \theta_i)^3 \Big|_{\theta_i=0}^{\theta_i=2\pi} \\
&= \frac{1}{6\pi}(2\pi^3) \\
&= \frac{1}{3}\pi^3.
\end{aligned}$$

□

For an explicit example, notice that this means the variance of total torsion for an open polygonal arm is $\frac{n}{3}\pi^2$, which pairs with Chebyshev's inequality to tell us that we should expect less than one-third of all open polygonal arms to have total torsion with absolute value greater than $\pi\sqrt{n}$. Given this, we sampled 100,000 polygons from $Arm_3(100)$ and found that 91,626 had total torsion less than $\pi\sqrt{100}$. Further, they had a mean of -0.030508 and a variance of only 1.257 larger than $\frac{100}{3}\pi^2$.

Proposition 121. *Where τ_i is the torsion angle at edge e_i of a closed polygon sampled under the symmetric measure on $Pol_3(n)$ or $\mathscr{P}_3(n)$, we have that the variance of total torsion $\tau = \sum_{i=1}^n \tau_i$ is bounded above by $\frac{n}{3}\pi^2 + n^2\pi^2\mathscr{B}_3(6,n)$.*

Proof. The $U(n)$ invariance will again give us that we should partition the pairs of torsion angles into: (A) (τ_i, τ_i), (B) $(\tau_i, \tau_{i\pm1})$, (C) $(\tau_i, \tau_{i\pm2})$ and (D) all others. Within these categories, those in (A) have covariance equal to $\mathrm{Cov}_p(\tau_1, \tau_1)$, those in (B) will match $\mathrm{Cov}_p(\tau_1, \tau_2)$, those in (C) will match $\mathrm{Cov}_p(\tau_1, \tau_3)$ and those in (D) will match $\mathrm{Cov}_p(\tau_1, \tau_4)$. Breaking the variance apart, we have:

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^n \tau_i\right) &= \sum_{i=1}^n \sum_{j=1}^n \mathrm{Cov}_p(\tau_i, \tau_j) \\
&= n\,\mathrm{Cov}_p(\tau_1,\tau_1) + 2n\,\mathrm{Cov}_p(\tau_1,\tau_2) + 2n\,\mathrm{Cov}_p(\tau_1,\tau_3) + (n^2-5n)\,\mathrm{Cov}_p(\tau_1,\tau_4) \\
&= nE_p(\tau_1^2) + 2nE_p(\tau_1\tau_2) + 2nE_p(\tau_1\tau_3) + (n^2-5n)E_p(\tau_1\tau_4) - n^2E[\tau_1]^2.
\end{aligned}$$

Here, we see that both $E_p(\tau_1)$ and $E_p(\tau_1^2)$ are obtained as the integral of an essentially bounded 3-edge locally determined function, and similarly we need 4 edges for $E_p(\tau_1\tau_2)$, 5 for $E_p(\tau_1\tau_3)$ and 6 for $E_p(\tau_1\tau_4)$. We have seen that, over arms, $E(\tau_i, \tau_j) = \delta_{i,j}\frac{1}{3}\pi^2$, so we may bound $E_p(\tau_1^2) \leq \frac{1}{3}\pi^2 + \pi^2\mathscr{B}_3(6,n)$, and $E_p(\tau_i\tau_j) \leq 0 + \mathscr{B}_3(6,n)$ for $i < j$, by using the fact that, for fixed n, $\mathscr{B}_3(k,n)$ is an increasing function of k. To see this fact, recall that we have $\mathscr{B}_3(k,n) = 2\left(\frac{4k+3}{4n-4k-3} + \frac{n^4}{(n-k-3)^4} - 1\right)$, written as the sum of three quantities, only the first two of which depend on k. In the sum, the first summand has an increasing numerator and decreasing denominator as k increases, while the second has constant numerator and decreasing denominator. This shows us that $B_3(k,n)$ is increasing in k (within its domain). This leaves us with:

$$\begin{aligned}
\operatorname{Var}\left(\sum_{i=1}^{n}\tau_i\right) &= nE_p(\tau_1^2) + 2nE_p(\tau_1\tau_2) + 2nE_p(\tau_1\tau_3) + (n^2-5n)E_p(\tau_1\tau_4) - n^2E[\tau_1]^2 \\
&\leq \frac{n}{3}\pi^2 + n^2\pi^2\mathscr{B}_3(6,n) - n^2E_p[\tau_1^2] \\
&\leq \frac{n}{3}\pi^2 + n^2\pi^2\mathscr{B}_3(6,n)
\end{aligned}$$

□

This bound is asymptotically bounded by $86\pi^2 n$, which we can pair with Chebyshev's Inequality and our earlier observation about the bounds on expected total torsion to see that, for large n, we expect that at least $\left(1-\dfrac{1}{\lambda^2}\right)100\%$ of polygons in $Pol_3(n)$ have total torsion in the range of $\pm\pi\left(55.5+\lambda\sqrt{86n}\right)$. As before, we would like to point out that this range becomes better than the trivial bounds on total torsion in the case of $\lambda=\sqrt{2}$ for $n>272$. Unfortunately, as we pointed out just after Proposition 121, the total torsion of a polygon sampled from $Pol_3(n)$ is bounded by $\pm n\pi\mathscr{B}_3(3,n)\simeq\pm55.5\pi$, and we can check that $\pi^2n^2\mathscr{B}_3(6,n)>55.5\pi$ for $n>8$, so our variance bound is just too high to be used profitably with Chebyshev's Inequality.

As an example, we sampled 100,000 100-edged closed polygons and computed their total torsion. We found, of course, that all 100,000 had total torsion between the range of $[-55.5\pi, 55.5\pi]$. More, they had a sample mean of 1.2629 and a sample variance of 327.96.

3.3 Local Knotting

Given that random walks have a lot of application to polymer physics, and that there is a lot of interest in that field concerning the knot type of a particular random walk, let us now move to the topic of knotting. The theme so far has been to look locally and make approximations on closed random polygons by comparing them to open random polygons. To continue this, rather than looking at the knot type of the entire random walk, we will look at local knotting.

Recall that a knot sum $K_1 \# K_2$ is not a well-defined operation (pg. 40 of [35] gives an example of the ambiguity with the square and granny knot), but it is still the case that if one of the summands is non-trivial, then the knot sum is non-trivial (pg. 281 of [35]). It is therefore tempting to think that if we can show random walks are locally knotted then we have shown that they are globally knotted. Unfortunately, this does not follow immediately, as there is always the possibility of retracing the steps involved in producing a local knot. Nonetheless, the likelihood of the presence of local knots is still an interesting question.

Definition 122. A ball-arc pair is a pair (B, α), where B is a ball and α is an arc, such that $\alpha \subseteq B$ and $\alpha \cap \partial B = \partial\alpha$.

One motivation to consider a ball-arc pair (B, α) is that it gives us a canonical way to turn the open arc α into a closed knot K_α. To do this, we simply connect the boundary of α by using a path along the boundary of B. This motivates the following new term:

Definition 123. A k-edged polygonal chain is knotable if the first and last endpoints are exterior to the convex hull of the middle $k-2$ endpoints.

To turn a knotable polygonal chain into a knot, we can then join the endpoints by any simple curve that remains exterior to the convex hull of the polygonal chain. Notice here that this also fits nicely with our previous observation: if a knotable polygonal chain forms a non-trivial knot, then any closed polygon that contains the chain and additionally avoids the convex hull of the chain will also be knotted.

As a final example of the de Finetti style result, we present the results of an experiment in which we sampled 100,000 polygons from each of $Arm_3(1000)$ and $Pol_3(1000)$. For each polygon sampled, we then determined whether the first 10-edged segment formed a knotable chain and the resulting knot-type when it is knotable. Since both the expectation that the segment is knottable and the expectation that the segment is knottable and forms a knot of type K can be viewed as the expectation of a Boolean expression and relies only on the 10-edges in question, we may use Theorem 114 to say that the difference in the expectations over open polygons and the expectations over closed polygons is bounded absolutely by $\mathscr{B}_3(10, 1000) \approx 0.129204$.

For the convenience of the reader, here is explicit pseudocode for this experiment, which assumes the existence of a function Gaussian which gives a random value sampled from a standard Gaussian distribution and $QHull$ [3].

ComplexDot(V, W)

    ▷ Compute the Hermitian dot product of two complex n-vectors

    for $ind = 1$ **to** n

        do $Dot{+} = V[ind] * \mathrm{Conj}(W[ind])$

    return Dot

Normalize(V, W)

    ▷ Normalize a complex n-vector to unit length.

    for $ind = 1$ **to** n

        do $UnitV[ind] = V[ind]/\mathrm{Sqrt}(\mathrm{ComplexDot}((V, V))$

    return UnitV

HopfMap(a, b)

$\triangleright$ Compute the vector in $\mathbb{R}^3$ given by the Hopf map applied to the quaternion $a + b\mathbf{j}$.

return $(a * \text{Conj}(a) - b * \text{Conj}(b), 2\text{Re}(a * \text{Conj}(b)), 2\text{Im}(a * \text{Conj}(b)))$

Random-Open-Polygon(n)

$\triangleright$ Produce edge vectors for random open space polygon of length 2.

$\triangleright$ 1. Generate a vector with Gaussian coordinates.

for $ind = 1$ **to** $2n$

do $A[ind] = Gaussian() + I * Gaussian()$

$\triangleright$ 2. Re-scale the vector

for $ind = 1$ **to** $2n$

do $SA[ind] = \text{SQRT}(2)A[ind]/\text{SQRT}(\text{ComplexDot}(A, A))$

$\triangleright$ 3. Apply the coordinate-wise Hopf map

for $ind = 1$ **to** n

do $Edge[ind] = \text{HopfMap}(SA[2 * ind - 1], SA[2 * ind])$

return Edge

Random-Closed-Polygon(n)

$\triangleright$ Produce edge vectors for random closed space polygon of length 2.

$\triangleright$ 1. Generate a frame with Gaussian Coordinate

for $ind = 1$ **to** n

do $A[ind] = Gaussian() + I * Gaussian()$

$B[ind] = Gaussian() + I * Gaussian()$

$\triangleright$ 2. Perform Gram-Schmidt to get $FrameA$ and $FrameB$

for $ind = 1$ **to** $2n$

do $FrameA[ind] = A[ind]$

$FrameB[ind] = B[ind] - (\text{ComplexDot}(B, A)/\text{ComplexDot}(A, A))A[ind]$

$FrameA = \text{Normalize}(FrameA)$

$FrameB = \text{Normalize}(FrameB)$

$\triangleright$ 3. Apply the coordinate-wise Hopf map

for $ind = 1$ **to** n

do $Edge[ind] = \text{HopfMap}(FrameA[ind], FrameB[ind])$

return Edge

GenerateVertices($Edge$)

$\triangleright$ Generates the vertices of a polygonal chain from its edge vectors.

Vert[0]=(0,0,0)

for $ind = 1$ **to** n

do $Vert[ind + 1] = Vert[ind] + Edge[ind]$

return Vert

IsKnotable($Vert$)

$\triangleright$ Tests to see if the given list of vertices form a knotable polygonal chain.

$\triangleright$ 1. Create the Convex Hull of the first and last $n - 1$ vertices.

for $ind = 1$ **to** $n - 1$

do $First[ind] = Vert[ind]$

$Last[ind] = Vert[ind + 1]$

$FirstHull = \text{qconvex}(First)$

$LastHull = \text{qconvex}(Last)$

$\triangleright$ 2. If the first and last vertex belong to the convex hull, then they are exterior to the convex hull of the middle $n - 2$ verticies.

If $Vert[1]$ is in $FirstHull$

then $FirstBoolean$=**True**

else $FirstBoolean$=**False**

If $Vert[n]$ is in $LastHull$

then $LastBoolean$=**True**

else $LastBoolean$=**False**

If $FirstBoolean$ **and** $LastBoolean$

then $Knottable$=**True**

else $Knottable$=**False**

return Knottable

Center($Vert$)

$\triangleright$ Finds the center of mass of a polygonal chain.

for $ind = 1$ **to** n

do $Center+ = V[ind]/n$

return $Center$

KnotChain($Vert$)

$\triangleright$ Closes a knotable arc of total length less than 2.

$\triangleright$ 1. Find the center of mass and directions from the center of mass to the ends of the segment.

$Center =$ Center($Vert$)

$FirstEdge = Vert[1] - Center$

$FirstEdgeUnit =$ Normalize($FirstEdge$)

$LastEdge = Vert[n] - Center$

$LastEdgeUnit =$ Normalize($LastEdge$)

$\triangleright$ 2. Create an arc external to the convex hull of the segment.

$NormalDirection =$ CrossProduct($FirstEdgeUnit, LastEdgeUnit$)

$NewVert[1] = Vert[n]$

$NewVert[2] = Vert[n] + 2 * LastEdgeUnit$

$NewVert[3] = NewVert[2] + 2 * NormalDirection$

$NewVert[5] = Vert[1] + 2 * FirstEdgeUnit$

$NewVert[4] = NewVert[5] + 2 * NormalDirection$

$\triangleright$ 3. Append the original segment to form a closed polygon.

for $ind = 1$ **to** n

do $NewVert[ind + 5] = Vert[ind]$

return $NewVert$

After finding those samples whose first 10-edge segment was knottable, their HOMFLYPT Polynomial and knot type were computed using $Vecttools$ [1].

Category	Total Found In $Arm_3(1000)$	Total Found in $Pol_3(1000)$
Knotable	55180	55255
Unknot	55105	55178
Knotted	75	77
Trefoil (3_1)	71	74
Figure Eight Knot (4_1)	3	3
Cinquefoil Knot (5_1)	1	0

Of particular note here is how drastically better the results are than the bound we were assured of. We include now some sample images of these knottable segments:

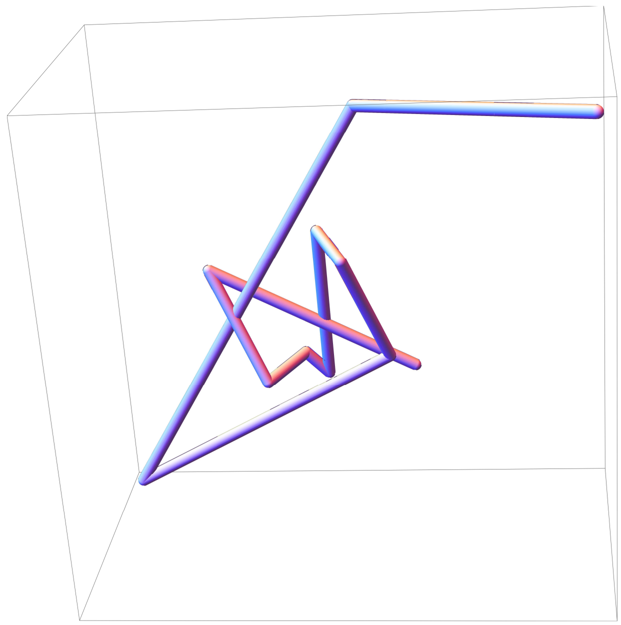

Figure 3.1: Sample 1162 from $Pol_3(1000)$ making knot 4_1

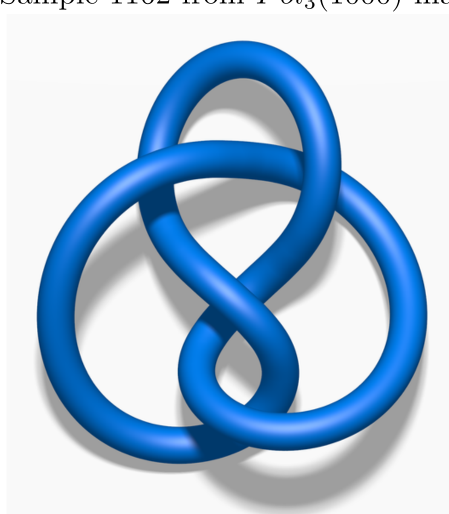

Figure 3.2: Knot 4_1 [2]

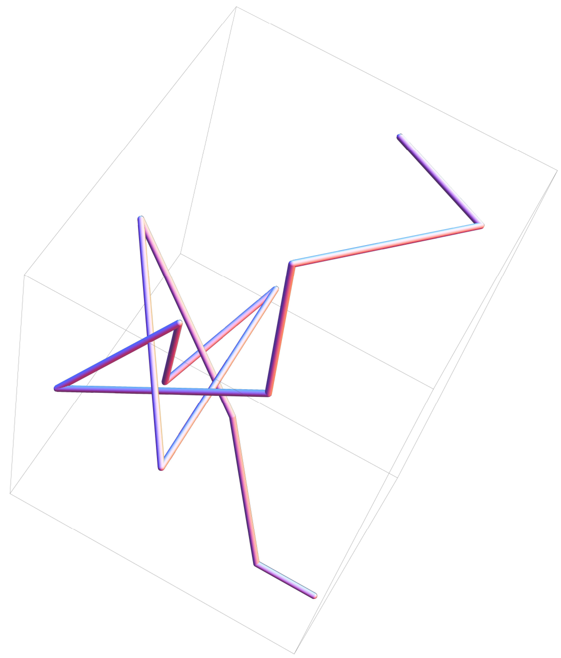

Figure 3.3: Sample 11838 from $Arm_3(1000)$ making knot 5_1

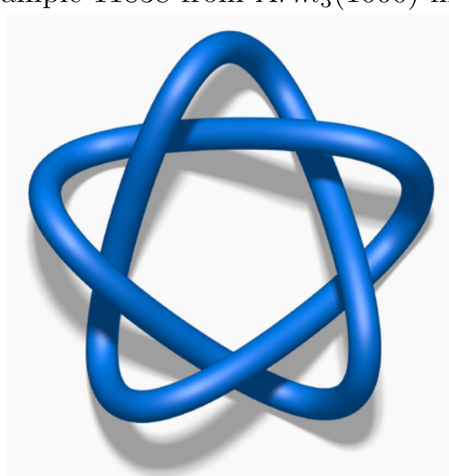

Figure 3.4: Knot 5_1 [2]

Chapter 4

Bibliography

[1] Ted Ashton and Jason Cantarella. Vecttools.

http://www.jasoncantarella.com/vecttools, October 29, 2013.

[2] Dror Bar-Natan and Scott Morrison. The Knot Atlas.

http://katlas.math.toronto.edu/, April 01, 2014.

[3] C. Bradford Barber, David P. Dobkin, and Hannu Huhdanpaa. The Quickhull algorithm for convex hulls. *Association for Computing Machinery Transactions on Mathematical Software*, 22(4):469-483, 1996. http://www.qhull.org

[4] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press. 2004.

[5] Jason Cantarella, Tetsuo Deguchi, and Clayton Shonkwiler. Probability theory of random polygons from the quaternionic viewpoint. *Communications on Pure and Applied Mathematics*, 2013. To appear. arXiv:1206.3161 [math.DG].

[6] Jason Cantarella, Alexander Y. Grosberg, Robert Kusner, and Clayton Shonkwiler. The expected total curvature of random polygons. Preprint, 2012. arXiv:1210.6537 [math.DG].

[7] Henri Cartan. Sur la mesure de Haar. *Comptes Rendus de l'Académie des Sciences de Paris*, 221:759-762, 1940.

[8] Persi Diaconis, Morris Eaton, and Steffen Lauritzan. Finite de Finetti theorems in linear modules and multivariate analysis. *Scandinavian Journal of Statistics*, (4), 289. doi:10.2307/4616249

[9] Persi Diaconis and David Freedman. A dozen de Finetti-style results in search of a theory. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques* 23 no. S2:297-423, 1987.

[10] Yuanan Diao, John C. Nardo, and Yanqing Sun, Global knotting in equilateral random polygons. *Journal of Knot Theory Ramifications*, 10(04):597-607, 2001.

[11] Jose A. Diaz-Garcia and Ramon Gutierrez-Jaimez. Complex bimatrix variate generalised beta distributions. 2009. arXiv:0906.1131 [math.ST].

[12] Rick Durrett. *Probability Theory and Examples*. Fourth Edition. Cambridge University Press. 2010.

[13] Morris L. Eaton. On some multivariate distributions and random orthogonal matrices. *In Proceedings of the Berkely conference in honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R.A. Olshen, eds.) 2681-691. Wadsworth, Monterey, California. 1985.

[14] John B. Fraleigh and Raymond A. Beauregard. *Linear Algebra*. Second Edition. Addison-Wesley Publishing Company. 1991.

[15] Bert Fristedt and Lawrence F. Gray. *A Modern Approach To Probability Theory.* Boston, Mass.: Birkhuser, 1997.

[16] Jean Gallier. Basics of manifolds and classical Lie groups: the exponential map, Lie groups, and Lie algebras. *Geometric Methods and Applications*, Volume 38. 459-528. 2011.

[17] N. R. Goodman, *Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). The Annals of Mathematical Statistics*, Vol 34, Issue 1, 152-177, 1936.

[18] Piotr Graczyk, Letac Grard , and Massam Hlne. The complex Wishart distribution and the symmetric group. *The Annals Of Statistics*, Vol. 31, Issue 1, 287-309. 2003.

[19] Izrail S Gradshteyn and Iosif M Ryzhik. *Table of Integrals, Series, and Products*. Elsevier/Academic Press, Amsterdam, 7th edition, 2007.

[20] Jean-Claude Hausmann and Allen Knutson. Polygon spaces and Grassmannians, *L'Enseignement Mathmatique. Revue Internationale. 2e Srie*, 43(1-2):173-198, 1997.

[21] Jean-Claude Hausmann and Allen Knutson. The cohomology ring of polygon spaces, *Annales Institut Fourier*, 48(1):281-321, 1998.

[22] Nicholas J. Highham. Computing the polar decomposition–with applications. *SIAM Journal on Scientific Computing*, 7(4), 1160. 1986.

[23] Benjamin Howard, Christopher Manon, and John J. Millson. The toric geometry of triangulated polygons in Euclidean space. *Canadian Journal of Mathematics*, 63(4):878-937, 2011.

[24] Barry D Hughes. *Random Walks and Random Environments: Volume 1: Random Walks*. Clarendon Press, 1995.

[25] Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous Univariate Distributions. Vol. 1.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York, 2nd edition, 1994.

[26] Douglas Jungreis. Gaussian random polygons are globally knotted. *Journal of Knot Theory and its Ramifications*, 3(4):455-464, 1994.

[27] Michael Kapovich and John J. Millson. The symplectic geometry of polygons in Euclidean space. *Journal of Differential Geometry*, 44(3):479-513, 1996.

[28] Steven G. Krantz, Steven and Harold R. Parks. *Geometric Integration Theory.* Boston, Mass.: Birkhuser, 2008.

[29] Reginald Douglas Lord. The use of the Hankel transform in statistics I. General theory and examples. *Biometrika*, 41(1/2):4455, 1954.

[30] Nathan Moore and Alexander Y. Grosberg. Limits of analogy between self-avoidance and topology-driven swelling of polymer loops. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 72(6 Pt 1):061803, 2005.

[31] Mervin. E. Muller. A note on a method for generating points uniformly on N-dimensional spheres. *Communications of the Association for Computing Machinery*, 2(4):19-20, 1959.

[32] National Institute of Standards and Technology. Digital Library of Mathematical Functions. http://dlmf.nist.gov/, March 23, 2012.

[33] Panagiotou, E., K. C. Millett, and S. Lambropoulou. The linking number and the writhe of uniform random walks and polygons in confined spaces. *Journal of Physics A: Mathematical and Theoretical*, 43(4):045208, 2010.

[34] Kaare B. Petersen and Michael S. Pedersen. *The Matrix Cookbook*. Technical University of Denmark. 2008.

[35] Dale Rolfsen. *Knots And Links*. Berkeley, Calif. : Publish or Perish, 1976.

[36] Michael Rubinstein and Ralph H. Colby. *Polymer Physics*. Oxford: Oxford University Press. 2006.

[37] Muni S. Srivastava. On the complex Wishart distribution. *The Annals Of Mathematical Statistics*, 36(1):313-315, 1965.

[38] E.M. Stein and R. Shakarchi. *Real Analysis, Measure Theory, Integration & Hilbert Spaces*, vol. 3, 1st ed., 2005.

[39] Michael Spivak. *Calculus*. Fourth Edition. Publish or Perish. 2008.

[40] De Witt Sumners and Stuart G. Whittington. Knots in self-avoiding walks. *Journal of Physics A: Mathematical and General*, 21(7):1689-1694, 1999.

[41] Rocco Varela, Kenneth Hinson, Javier Arsuaga, and Yuanan Diao. A fast ergodic algorithm for generating ensembles of equilateral random polygons. *Journal of Physics A:* 42(9):095204, 2009.

[42] Stuart G. Whittington. Topology of polymers. *New Scientific Applications of Geometry and Topology* (ed D W Sumners). 7395. Providence, Rhode Island: American Mathematical Society. 1992.

Appendix A

Proof of Proposition 81

Recall that we have shown that the expected turning angle is given by the integral:

$$\theta(n) = \int_0^\pi \int_0^\infty \int_0^\infty \frac{n-2}{2^{n-2}\pi\Gamma\left(\frac{n}{2}-1\right)} e^{-\frac{1}{2}(r_1+r_2)} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-2}\psi \, \mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}\psi.$$

To save space, let us denote by $\gamma(n) := \dfrac{n-2}{2^{n-2}\pi\Gamma\left(\frac{n}{2}-1\right)}$. Since we are already using $z = \sqrt{r_1^2 + r_2^2 + 2r_2r_2\cos(\psi)}$, we will introduce a change of integration variables of $(r_1, r_2, \theta) \mapsto (x, y, z)$ with z as defined, $x = \dfrac{r_1 + r_2}{2}$, and $y = \dfrac{r_1 - r_2}{2}$. In doing so, it will be helpful to note that $r_1 = x + y$, $r_2 = x - y$ and $\theta = \arccos\left(\dfrac{z^2 - 2(x^2+y^2)}{2(x+y)(x-y)}\right)$. Further, this change of variables will involve an inverse determinant of the Jacobian with value $\dfrac{2\sqrt{r_1^2 + r_2^2 + 2r_2r_2\cos(\psi)}}{r_1r_2\sin(\theta)} = \dfrac{2z}{x^2-y^2}\left(1 - \left(\dfrac{z^2-2(x^2+y^2)}{2(x+y)(x-y)}\right)^2\right)^{-\frac{1}{2}}$, where we have simplified $\csc(\arccos(\alpha)) = \dfrac{1}{\sqrt{1-\alpha^2}}$. We then further simplify

$\left(1 - \left(\dfrac{z^2-2(x^2+y^2)}{2(x+y)(x-y)}\right)^2\right)^{-\frac{1}{2}} = \dfrac{2(x-y)(x+y)}{\sqrt{(4x^2-z^2)(z^2-4y^2)}}$. Before introducing this change, note that we can simplify $\dfrac{2z}{x^2-y^2}\left(1 - \left(\dfrac{z^2-2(x^2+y^2)}{2(x+y)(x-y)}\right)^2\right)^{-\frac{1}{2}} = \dfrac{4z}{\sqrt{(4x^2-z^2)(z^2-4y^2)}}$

This change produces the integral:

$$\theta(n) = 4\gamma(n)\int_0^\infty\int_{z/2}^\infty\int_{-z/2}^{z/2} e^{-x}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\arccos\left(\frac{z^2-2(x^2+y^2)}{2(x^2-y^2)}\right)\frac{\mathrm{d}y\mathrm{d}x\mathrm{d}z}{\sqrt{(4x^2-z^2)(z^2-4y^2)}}$$

Next, we introduce another change of variables, $(x,y)\mapsto\left(\frac{z}{2}s,\frac{z}{2}t\right)$:

$$\begin{aligned}\theta(n) &= 4\gamma(n)\int_0^\infty\int_{z/2}^\infty\int_{-z/2}^{z/2} e^{-x}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\arccos\left(\frac{z^2-2(x^2+y^2)}{2(x^2-y^2)}\right)\frac{\mathrm{d}y\mathrm{d}x\mathrm{d}z}{\sqrt{(4x^2-z^2)(z^2-4y^2)}}\\ &= \gamma(n)\int_0^\infty\int_1^\infty\int_{-1}^{1} e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\operatorname{arcsec}\left(\frac{s^2-t^2}{2-s^2-t^2}\right)\frac{\mathrm{d}t\mathrm{d}s\mathrm{d}z}{\sqrt{(s^2-1)(1-t^2)}}\end{aligned}$$

We will now apply integration by parts to the integral with respect to t:

$$\begin{aligned}\int_{-1}^{1}\operatorname{arcsec}\left(\frac{t^2-s^2}{t^2+s^2-2}\right)\frac{1}{\sqrt{(1-t^2)}}\,\mathrm{d}t &= \arcsin(t)\operatorname{arcsec}\left(\frac{t^2-s^2}{t^2+s^2-2}\right)\Big|_{-1}^{1}\\ &\quad-\int_{-1}^{1}\arcsin(t)\frac{2t}{s^2-t^2}\sqrt{\frac{s^2-1}{1-t^2}}\mathrm{d}t\end{aligned}$$

The constant will evaluate to $\frac{\pi^2}{2}-\frac{-\pi^2}{2}=\pi^2$, as we have $\lim_{x\to1^-}\arcsin(\pm x)=\pm\frac{\pi}{2}$ and $\operatorname{arcsec}\left(\frac{1-s^2}{s^2-1}\right)=\operatorname{arcsec}(-1)=\pi$. Our integral is now:

$$\theta(n) = \gamma(n)\int_0^\infty\int_1^\infty\int_{-1}^{1} e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\operatorname{arcsec}\left(\frac{s^2-t^2}{2-s^2-t^2}\right)\frac{\mathrm{d}t\mathrm{d}s\mathrm{d}z}{\sqrt{(s^2-1)(1-t^2)}} \quad \text{(A.1)}$$

$$= \gamma(n)\int_0^\infty\int_1^\infty e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\frac{1}{\sqrt{s^2-1}}\left(\pi^2-\int_{-1}^{1}\frac{2t\arcsin(t)}{s^2-t^2}\sqrt{\frac{s^2-1}{1-t^2}}\mathrm{d}t\right)\mathrm{d}s\mathrm{d}z. \quad \text{(A.2)}$$

Let us first deal with the constant term. We will be showing that

$$\gamma(n)\int_0^\infty\int_1^\infty e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\frac{1}{\sqrt{s^2-1}}\pi^2\,\mathrm{d}s\mathrm{d}z=\pi. \quad \text{(A.3)}$$

First, we recognize that $\int_1^\infty \frac{e^{-s\frac{z}{2}}}{\sqrt{s^2-1}}\,\mathrm{d}s$ is the integral representation of $\frac{\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi}}K_0\left(\frac{z}{2}\right) = K_0\left(\frac{z}{2}\right)$, given by 10.32.8 of [32]. So the integral we are interested in is just:

$$\pi^2\gamma(n)\int_0^\infty z^{\frac{n}{2}-1}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)K_0\left(\frac{z}{2}\right)\,\mathrm{d}z.$$

Next, we apply the product identity 10.32.17 from [32] to write $K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)K_0\left(\frac{z}{2}\right) = 2\int_0^\infty K_{\frac{n}{2}-2-0}\left(2\frac{z}{2}\cosh(t)\right)\cosh\left(\left(\frac{n}{2}-2+0\right)t\right)\,\mathrm{d}t$. Bringing out integral to $2\pi^2\gamma(n)\int_0^\infty\int_0^\infty z^{\frac{n}{2}-1}K_{\frac{n}{2}-2}\left(z\cosh(t)\right)\cosh\left(\left(\frac{n}{2}-2\right)t\right)\,\mathrm{d}t$. Working first on the integral with respect to z, Lemma 16 of [6] shows us that $\int_0^\infty z^{\frac{n}{2}-1}K_{\frac{n}{2}-2}\left(z\cosh(t)\right)\,\mathrm{d}z = 2^{\frac{n}{2}-2}\cosh(t)^{-\frac{n}{2}}\Gamma\left(\frac{n}{2}-1\right)$. We now have the integral

$$\begin{aligned}
&\pi^2 2^{\frac{n}{2}-2}\Gamma\left(\frac{n}{2}-1\right)\gamma(n)2\int_0^\infty\cosh(t)^{-\frac{n}{2}}\cosh\left(\left(\frac{n}{2}-2\right)t\right)\,\mathrm{d}t\\
=&\pi^2 2^{\frac{n}{2}-2}\Gamma\left(\frac{n}{2}-1\right)\frac{n-2}{2^{n-2}\pi\Gamma\left(\frac{n}{2}-1\right)}2\int_0^\infty\cosh(t)^{-\frac{n}{2}}\cosh\left(\left(\frac{n}{2}-2\right)t\right)\,\mathrm{d}t\\
=&\pi(n-2)2^{1-\frac{n}{2}}\int_0^\infty\cosh(t)^{-\frac{n}{2}}\cosh\left(\left(\frac{n}{2}-2\right)t\right)\,\mathrm{d}t
\end{aligned}$$

We can now finish the evaluation using 3.5.17 of [19] to give us $\int_0^\infty\cosh(t)^{-\frac{n}{2}}\cosh\left(\left(\frac{n}{2}-2\right)t\right)\,\mathrm{d}t = \frac{2^{\frac{n}{2}-1}}{n-2}$, wherein we obtain a final result of π.

We now have

$$\theta(n) = \pi - \gamma(n)\int_0^\infty\int_1^\infty e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\frac{1}{\sqrt{s^2-1}}\left(\int_{-1}^1\frac{2t\arcsin(t)}{s^2-t^2}\sqrt{\frac{s^2-1}{1-t^2}}\mathrm{d}t\right)\mathrm{d}s\mathrm{d}z \tag{A.4}$$

$$= \pi - \gamma(n)\int_0^\infty\int_1^\infty e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\left(\int_{-1}^1\frac{2t\arcsin(t)}{s^2-t^2}\frac{1}{\sqrt{1-t^2}}\mathrm{d}t\right)\mathrm{d}s\mathrm{d}z. \tag{A.5}$$

From here, making the substitution $t \mapsto \sin(\phi)$, will produce:

$$\begin{aligned}
\theta(n) &= \pi - \gamma(n) \int_0^\infty \int_1^\infty e^{-\frac{sz}{2}} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-1} \left(\int_{-1}^{1} \frac{2t \arcsin(t)}{s^2 - t^2} \frac{1}{\sqrt{1-t^2}} \mathrm{d}t \right) \mathrm{d}s \mathrm{d}z \\
&= \pi - \gamma(n) \int_0^\infty \int_1^\infty e^{-\frac{sz}{2}} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-1} \left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{2\sin(\phi)\phi}{s^2 - \sin^2(\phi)} \frac{1}{\sqrt{1-\sin^2(\phi)}} \cos(\phi) \mathrm{d}\phi \right) \mathrm{d}s \mathrm{d}z \\
&= \pi - \gamma(n) \int_0^\infty \int_1^\infty e^{-\frac{sz}{2}} K_{\frac{n}{2}-2}\left(\frac{z}{2}\right) z^{\frac{n}{2}-1} \left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{2\sin(\phi)\phi}{s^2 - \sin^2(\phi)} \mathrm{d}\phi \right) \mathrm{d}s \mathrm{d}z
\end{aligned}$$

Notice that we may decompose $\dfrac{2\sin(\phi)}{s^2 - \sin^2(\phi)} = \left(\dfrac{1}{s - \sin(\phi)} - \dfrac{1}{s + \sin(\phi)} \right)$. 3.794 of [19] tells us that $\displaystyle\int_0^\pi \frac{x \, \mathrm{d}x}{1 + a^2 + 2a\cos(x)} = \frac{\pi^2}{2(1-a^2)} + \frac{4}{1-a^2} \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k+1)^2}$. This leads us to make the change of variables $\zeta = \dfrac{\pi}{2} - \phi$, so that we can turn the integral with respect to ϕ into:

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} \frac{2\phi \sin(\phi)}{s^2 - \sin^2(\phi)} \mathrm{d}\phi &= \int_{-\pi/2}^{\pi/2} \frac{\phi}{s - \sin(\phi)} - \frac{\phi}{s + \sin(\phi)} \mathrm{d}\phi \\
&= \int_0^\pi \left(\frac{\pi}{2} - \zeta \right) \left(\frac{1}{s - \cos(\zeta)} - \frac{1}{s + \cos(\zeta)} \right) \mathrm{d}\zeta \\
&= -\int_0^\pi \zeta \left(\frac{1}{s - \cos(\zeta)} - \frac{1}{s + \cos(\zeta)} \right) \mathrm{d}\zeta
\end{aligned}$$

In this last line, we have used 2.553(3) of [19] to see that

$\displaystyle\int_0^\pi \frac{1}{s - \cos(\zeta)} - \frac{1}{s + \cos(\zeta)} \mathrm{d}\zeta = 0$. Next, in order to use 3.794, we must express

$$\frac{\zeta}{s + \cos(\zeta)} = b \left(\frac{\zeta}{1 + a^2 + 2a\cos(\zeta)} \right) \text{ and } \frac{\zeta}{s - \cos(\zeta)} = d \left(\frac{\zeta}{1 + c^2 + 2c\cos(\zeta)} \right).$$

In the left-hand side of the equations, we have a coefficient of ± 1 for the $\cos(\zeta)$ term, so we must have $b = 2a$ and $d = -2c$ with $a > 0$ and $c < 0$. This tells us that $s = \frac{1+a^2}{b} = \frac{1+a^2}{2a}$ and $s = \frac{1+c^2}{d} = -\frac{1+c^2}{2c}$. Applying the quadratic formula, we see that $a = s \pm \sqrt{s^2-1}$ and $c = -s \pm \sqrt{s^2-1}$. In [19], we see that 3.794 only applies for $a^2 < 1$, so we will choose $a = s - \sqrt{s^2-1}$ and $c = -s + \sqrt{s^2-1}$, where we see that $c = -a$. We may then evaluate our integral as:

$$
\begin{aligned}
I &= \int_{-\pi/2}^{\pi/2} \frac{2\phi \sin(\phi)}{s^2 - \sin^2(\phi)} \, \mathrm{d}\phi \\
&= -\int_0^{\pi} \zeta \left(\frac{1}{s - \cos(\zeta)} - \frac{1}{s + \cos(\zeta)} \right) \mathrm{d}\zeta \\
&= -\int_0^{\pi} d \frac{\zeta}{1 + c^2 + 2c\cos(\zeta)} - b \frac{\zeta}{1 + a^2 + 2a\cos(\zeta))} \, \mathrm{d}\zeta \\
&= -\int_0^{\pi} -2c \frac{\zeta}{1 + c^2 + 2c\cos(\zeta)} - 2a \frac{\zeta}{1 + a^2 + 2a\cos(\zeta))} \, \mathrm{d}\zeta \\
&= 2\int_0^{\pi} c \frac{\zeta}{1 + c^2 + 2c\cos(\zeta)} + a \frac{\zeta}{1 + a^2 + 2a\cos(\zeta))} \, \mathrm{d}\zeta \\
&= 2\left(c \left(\frac{\pi^2}{2(1-c^2)} + \frac{4}{1-c^2} \sum_{k=0}^{\infty} \frac{c^{2k+1}}{(2k+1)^2} \right) + a \left(\frac{\pi^2}{2(1-a^2)} + \frac{4}{1-a^2} \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k+1)^2} \right) \right) \\
&= 2\left((-a) \left(\frac{\pi^2}{2(1-a^2)} + \frac{4}{1-a^2} \sum_{k=0}^{\infty} \frac{(-a)^{2k+1}}{(2k+1)^2} \right) + a \left(\frac{\pi^2}{2(1-a^2)} + \frac{4}{1-a^2} \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k+1)^2} \right) \right) \\
&= 4a \frac{4}{1-a^2} \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k+1)^2} \\
&= 8 \frac{2a}{1-a^2} \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k+1)^2} \\
&= 8 \frac{1}{\sqrt{s^2-1}} \sum_{k=0}^{\infty} \frac{a^{2k+1}}{(2k+1)^2}
\end{aligned}
$$

We then recognize that

$$4\sum_{k=0}^{\infty}\frac{a^{2k+1}}{(2k+1)^2}=a\sum_{k=0}^{\infty}4\frac{a^{2k}}{(2k+1)^2}=a\sum_{k=0}^{\infty}\frac{(a^2)^k}{\left(k+\frac{1}{2}\right)^2}=a\Phi\left(a^2,2,\frac{1}{2}\right).$$

Substituting this in will bring us to our final form:

$$\begin{aligned}
\theta(n) &= \pi-\gamma(n)\int_0^{\infty}\int_1^{\infty}e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{2\sin(\phi)\phi}{s^2-\sin^2(\phi)}\mathrm{d}\phi\right)\mathrm{d}s\mathrm{d}z \\
&= \pi-\gamma(n)\int_0^{\infty}\int_1^{\infty}e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\left(8\frac{1}{\sqrt{s^2-1}}\sum_{k=0}^{\infty}\frac{a^{2k+1}}{(2k+1)^2}\right)\mathrm{d}s\mathrm{d}z \\
&= \pi-\gamma(n)\int_0^{\infty}\int_1^{\infty}e^{-\frac{sz}{2}}K_{\frac{n}{2}-2}\left(\frac{z}{2}\right)z^{\frac{n}{2}-1}\left(\frac{2a}{\sqrt{s^2-1}}\Phi\left(a^2,2,\frac{1}{2}\right)\right)\mathrm{d}s\mathrm{d}z
\end{aligned}$$

Appendix B

Proof of Proposition 96

Re-writing this in terms of r_i and θ_i gives us:

$$\begin{aligned}
&\frac{1}{4}(1 - \sigma(r_2 \sin(\theta_2))\sigma(r_2 \sin(\theta_2) + r_3 \sin(\theta_3))) \\
&* (1 - \sigma(\cos(\theta_3)(r_2 \sin(\theta_2)) - \sin(\theta_3)(r_2 \cos(\theta_2))) \\
&* \sigma(\cos(\theta_3)(r_2 \sin(\theta_2)) - \sin(\theta_3)(r_2 \cos(\theta_2) + r1)))
\end{aligned}$$

This is quite a lot to work with, so we will simplify it first. After distributing the terms and using the fact that $\sigma(a)\sigma(b) = \sigma(ab)$, we get:

$$\begin{aligned}
&\frac{1}{4}(1 - \sigma(r_2^2 \sin^2(\theta_2) + r_2 r_3 \sin(\theta_2)\sin(\theta_3)) \\
&- \sigma((r_2 \cos(\theta_3)\sin(\theta_2) - r_2 \cos(\theta_2)\sin(\theta_3)) \\
&\quad * (r_2 \cos(\theta_3)\sin(\theta_2) - r_2 \cos(\theta_2)\sin(\theta_3) - r_1 \sin(\theta_3))) \\
&+ \sigma((r_2^2 \sin^2(\theta_2) + r_2 r_3 \sin(\theta_2)\sin(\theta_3)) \\
&\quad * (r_2 \cos(\theta_3)\sin(\theta_2) - r_2 \cos(\theta_2)\sin(\theta_3)) \\
&\quad * (r_2 \cos(\theta_3)\sin(\theta_2) - r_2 \cos(\theta_2)\sin(\theta_3) - r_1 \sin(\theta_3))))
\end{aligned}$$

Factoring out r_2, and using that $\sigma(r_2) = 1$, we can simplify this to:

$$\frac{1}{4}(1 - \sigma(r_2 \sin^2(\theta_2) + r_3 \sin(\theta_2)\sin(\theta_3))$$
$$- \sigma((\cos(\theta_3)\sin(\theta_2) - \cos(\theta_2)\sin(\theta_3))(r_2\cos(\theta_3)\sin(\theta_2) - r_2\cos(\theta_2)\sin(\theta_3) - r_1\sin(\theta_3)))$$
$$+ \sigma((r_2 \sin^2(\theta_2) + r_3\sin(\theta_2)\sin(\theta_3))(\cos(\theta_3)\sin(\theta_2) - \cos(\theta_2)\sin(\theta_3))$$
$$* (r_2\cos(\theta_3)\sin(\theta_2) - r_2\cos(\theta_2)\sin(\theta_3) - r_1\sin(\theta_3))))$$

Finally, using the fact that $\sin(\alpha - \beta) = \cos(\beta)\sin(\alpha) - \cos(\alpha)\sin(\beta)$ with the fact that $0 < \theta_2 < \pi$, we arrive at the proposed simplified form:

$$\frac{1}{4}(1 - \sigma(r_2\sin(\theta_2) + r_3\sin(\theta_3))$$
$$- \sigma(\sin(\theta_2 - \theta_3)(r_2\sin(\theta_2 - \theta_3) - r_1\sin(\theta_3)))$$
$$+ \sigma((r_2\sin(\theta_2) + r_3\sin(\theta_3))(r_2\sin(\theta_2 - \theta_3) - r_1\sin(\theta_3))\sin(\theta_2 - \theta_3)))$$