Uniqueness of Clifford Torus with Prescribed Isoperimetric Ratio

Thomas Yu∗ Jingmin Chen†

November 18, 2020

Abstract:
The Marques-Neves theorem asserts that among all the toroidal (i.e. genus 1) closed surfaces, the Clifford torus has the minimal Willmore energy $\int H^2 dA$. Since the Willmore energy is invariant under Möbius transformations, it can be shown that there is a one-parameter family, up to homotheties, of genus 1 Willmore minimizers. It is then a natural conjecture that such a minimizer is unique if one prescribes its isoperimetric ratio. In this article, we show that this conjecture can be reduced to the positivity question of a polynomial recurrence. A proof of the positivity can be found in the companion article [17].

Acknowledgments. This work is partially supported by NSF grants DMS 0915068 and DMS 1115915. We thank Robert Kusner for bringing to our attention the uniqueness problem and his earlier involvement of this project. We are grateful to Manuel Kauers, Stephen Melczer and Pierre Lairez for sharing their expertise in P-recurrences.

Keywords: Canham-Evans-Helfrich model, Willmore energy, Clifford torus, Möbius geometry, Marques-Neves theorem, Uniqueness, P-recurrence, Special functions, Positivity

1 Introduction

In this paper, we provide the proof, modulo a final step to be presented in the article [17], of the following result:

Theorem 1.1. The 3-D Euclidean shape1 of the Clifford torus $\{[\cos u, \sin u, \cos v, \sin v]^T/\sqrt{2} : u, v \in [0, 2\pi]\}$ in $S^3$ stereographically projected to $\mathbb{R}^3$ is uniquely determined by its isoperimetric ratio.

For an account of the interest of this result in connection to the Marques-Neves theorem [14] and the uniqueness of solution of the Canham-Evans-Helfrich model for biomembranes, see [22, Page 1-4] and the article [13] in the Spring 2016 MSRI newsletter. See also the remarks in Section 5. This result also furnishes a rigorous test case for the study of numerical methods; see [4, 6] and the references therein.

Let $T_R := \{(R + \cos v)\cos u, (R + \cos v)\sin u, \sin v : u, v \in [0, 2\pi]\}$, a torus embedded in 3-space with major radius $R \in (1, \infty)$, minor radius 1. Let $i_{(x,y,z)}$ be the inversion map about the unit sphere centered at $(x,y,z)$ of $\mathbb{R}^3$. Recall that the set of all stereographic images of the Clifford torus in $S^3$ to $\mathbb{R}^3$ is tantamount to the set of all images of $T_{\sqrt{2}}$ under Möb(3), the 3-D Möbius group consisting of all rigid motions, scaling, and sphere inversions.

The proof of Theorem 1.1 connects ideas in geometry and combinatorics and is divided into four steps:

∗Department of Mathematics, Drexel University. Email: yut@drexel.edu. He is supported in part by the National Science Foundation grants DMS 0512673 and DMS 0915068.
†Citigroup Global Markets Inc., 390 Greenwich Street, New York, NY 10013, U.S.A.. Email: jingmchen@gmail.com.
1Two subsets of $\mathbb{R}^3$ have the same Euclidean shape if they can be transformed from one to another by a homothety.
I. Prove that the set of all non-homothetic images of $T_{\sqrt{2}}$ under Möb(3) corresponds exactly to the one-parameter family

$$\{i(a,0,0)(T_{\sqrt{2}}) : a \in [0,\sqrt{2} - 1)\}. \quad (1.1)$$

In other words, the cyclides depicted in Figure 1 are exactly the set of all non-homothetic Clifford tori.

This is established in Theorem 2.4 of Section 2.

II. With this result, the conjecture follows if we can show:

$$\text{Iso} : [0,\sqrt{2} - 1) \to \{(3/2)(2\pi^2)^{-1/4}, 1\}, \quad \text{Iso}(a) := v(i(a,0,0)(T_{\sqrt{2}})) \quad (1.2)$$

is a bijection. If so, then each $v_0 \in [(3/2)(2\pi^2)^{-1/4}, 1)$ corresponds to one and only one Clifford torus, namely $i_{\text{Iso}^{-1}(v_0),0,0}(T_{\sqrt{2}})$, with isoperimetric ratio $v_0$, which must be the unique solution of the genus 1 Canham problem with $v_0$ as the constrained isoperimetric ratio.

To prove that Iso is a bijection, it suffices to show that Iso is monotonic increasing and

$$\lim_{a \to \sqrt{2} - 1} \text{Iso}(a) = 1. \quad (1.3)$$

In Section 3, we establish Theorem 3.1, which is a more general version of (1.3).

III. To prove that Iso is monotonic increasing, we venture into the realm of special functions. We make the observation that the area and enclosing volume of the cyclides in (1.1), denoted by $A(a)$ and $V(a)$, can be extended analytically to the disc $\{z : |z| < \sqrt{2} - 1\}$ on the complex plane. Moreover, the coefficients $(a_n)_{n \geq 0}$ and $(v_n)_{n \geq 0}$ of their power series at $z = 0$ are holonomic, or P-recursive, sequences, i.e. they satisfy linear recurrences with polynomial coefficients. We work out explicitly these P-recurrences in Section 4.

Since $\text{Iso}^2(a)/(36\pi) = V^2(a)/A^3(a)$, Iso is monotonic increasing iff the logarithm of the right-hand side is. But then we have

$$\frac{d}{da} \ln \frac{V^2(a)}{A^3(a)} = \frac{2V'(a)A(a) - 3V(a)A'(a)}{V(a)A(a)},$$

(While $\dim \text{Möb}(3) = 10$, the fact that there is only one degree of freedom in the set of all non-homothetic images of $T_{\sqrt{2}}$ under Möb(3) is quite trivial to those versed in the subject: the 7-dimensional subgroup of Möb(3) consisting of the homotheties (a.k.a. similarity transformations) in $\mathbb{R}^3$, by definition, contributes nothing to the Euclidean shapes. Next, a subgroup of $SO(4)$ isomorphic to $SO(1) \times SO(1)$ clearly leaves the Clifford torus in $S^3$ invariant; this takes away two more degrees of freedom. (The 2-parameter family of congruent transformations in Möbius geometry are neither congruences nor similarities in Euclidean geometry.) The explicit one-to-one correspondence between the inversion parameter ‘$a$’ and the ‘shape space’ requires a more prudent effort.)
so Iso is monotonic increasing iff $2V'(a)A(a) - 3V(a)A'(a) > 0$ on $(0, \sqrt{2} - 1)$. The fact that $A(z)$ and $V(z)$ are holonomic implies that $D(z) := 2V'(z)A(z) - 3V(z)A'(z)$ is also holonomic; the coefficients $(a_n)_{n \geq 0}$ of the power series of $D(z)$ at $z = 0$ follows the P-recurrence (4.7) derived in Section 4.

The monotonicity of Iso follows if all the terms defined by the P-recurrence (4.7) are positive.

*IV. Prove that all terms defined by the P-recurrence (4.7) are positive.

This last step is carried out in [17].

Steps I-III are carried out in the next three sections.

2 Step I: Non-homothetic Clifford tori

Our first goal is to characterize all the Euclidean shapes of the Clifford tori, i.e. we would like to find a parametrization of the ‘shape space’

$$\{i(x,y,z)(T_{\sqrt{3}}) : (x, y, z) \in \mathbb{R}^3 \setminus T_{\sqrt{3}}\} / \text{Hom}(3).$$

(2.1)

Here ‘$/\text{Hom}(3)$’ means we identify two point sets if they can be transformed from one to another by a homothety in $\mathbb{R}^3$. Since we are primarily interested in Euclidean shapes here, we avoid sphere inversions centered at points on $T_R$ itself. To help us gain a better insight of the underlying structure, we also study the more general shape space

$$\{i(x,y,z)(T_R) : R > 1, (x, y, z) \in \mathbb{R}^3 \setminus T_R\} / \text{Hom}(3).$$

(2.2)

Maxwell’s characterization of a cyclide. It is well-known that any (torodial) cyclide $C$ has two orthogonal planes of mirror symmetry; see, for example, [15, 2, 3]. We make the observation that the Euclidean shape of a toroidal cyclide $C$ is uniquely determined by certain measurements of the cross section of $C$ with either one of the two symmetry planes.

We use Maxwell’s characterization of cyclides [15, 2, 3]: any cyclide $C$ is the envelope of all the spheres centered at the points $P$ on a given ellipse $E$ with radii $r(P)$, $P \in E$, satisfying $r(P) + FP = L$, where $F$ is one of the foci of $E$ and $L$ is a constant in a suitable range. We can think of $L$ as the length of a taut string attached in one end to $F$; the string slides smoothly on $E$ and traces out spheres with the other end. See Figure 2. Under this characterization, $C$ is a torodial cyclide if and only if

$$a > L - a > f,$$

where $a$, $f$ and $L$ are the major radius of $E$, the focal length of $E$, and the length of the string, respectively. Moreover, the Euclidean shape of $C$ can be characterized by the ratio $a : f : L$.\(^{3}\)

The major axis of $E$ lies on the intersecting line of the two symmetry planes of $C$. In the following, $P_1$ refers to the symmetry plane where $E$ lies, whereas $P_2$ ($\perp P_1$) refers to the other symmetry plane. The cross section $C \cap P_1$ consists of two circles exterior to each other, whereas the cross section $C \cap P_2$ consists of two circles with one lying inside the other (see Figure 2).

Denote the radii of the two circles in $C \cap P_1$ by $r_1$ and $r_2$ and the distance between the two centers by $d$ (see Figure 2). Similarly, let $\tilde{r}_1$ and $\tilde{r}_2$ be the radii of the two circles in $C \cap P_2$ and $\tilde{d}$ be the distance between the two centers. By convention, $r_1 \geq r_2$, $\tilde{r}_1 \geq \tilde{r}_2$. The three sets of measurements $(r_1, r_2, d)$, $(\tilde{r}_1, \tilde{r}_2, \tilde{d})$ and $(a, f, L)$ of a cyclide $C$ are related by the following equations:

$$a = \frac{d}{2}, \quad f = \frac{r_1 - r_2}{2}, \quad L = \frac{d + r_1 + r_2}{2},$$

(2.3)

$$\tilde{r}_1 = \frac{d + (r_1 + r_2)}{2}, \quad \tilde{r}_2 = \frac{d - (r_1 + r_2)}{2}, \quad \tilde{d} = r_1 - r_2.$$  

(2.4)

Since the maps $(a, f, L) \mapsto (r_1, r_2, d)$ and $(r_1, r_2, d) \mapsto (\tilde{r}_1, \tilde{r}_2, \tilde{d})$ are linear isomorphisms, we conclude that:

\(^{3}\)This already explains why the shape space (2.2) is two-dimensional.
Lemma 2.1. Each of the three ratios
\[ a : f : L, \quad r_1 : r_2 : d \quad \text{and} \quad \tilde{r}_1 : \tilde{r}_2 : \tilde{d} \]
determines the Euclidean shape of the cyclide \( \mathcal{C} \).

For any \( \varrho > 0 \), let \( C(\varrho) = C(\varrho; R) \) be the circle in the \( \varrho \)-z plane with a diameter connecting \((\varrho, 0)\) and \(((R^2 - 1)/\varrho, 0)\); see Figure 3. By convention, \( C(0) = C(\infty) \) is the z-axis. In general, we have
\[ C(\varrho) = C((R^2 - 1)/\varrho). \]

These circles on the plane can be extended to the following tori in 3-D:
\[ T(\varrho) := T(\varrho; R) := \{(\rho \cos(\theta), \rho \sin(\theta), z) : (\rho, z) \in C(\varrho), \theta \in [0, 2\pi]\}. \quad (2.5) \]

For any fixed \( R \), the torus \( T(\varrho) \) lies completely outside, on, or inside the torus \( T \) when \( \varrho \in [0, R - 1) \cup (R + 1, \infty] \), \( \varrho = R \pm 1 \), or \( \varrho \in (R - 1, R + 1) \), respectively. In particular, \( T(R \pm 1; R) = T_R \). On the \( \rho \)-z plane, these correspond to the red, green and blue circles in Figure 3. While the one-parameter family of circles
\[ \left\{ C(\varrho) : \varrho \in [0, \sqrt{R^2 - 1}] \right\} \]
partitions the \( \rho \)-z plane,\(^4\) the corresponding one-parameter family of tori
\[ \left\{ T(\varrho) : \varrho \in [0, \sqrt{R^2 - 1}] \right\} \]
partitions \( \mathbb{R}^3 \). We shall see that how these circles and tori characterize the shape spaces (2.1) and (2.2).

Theorem 2.2. For any fixed \( R \in (1, \infty) \) and \( \varrho \in [0, \infty] \setminus \{R - 1, R + 1\} \), all the cyclides in
\[ \left\{ i_{(x,y,z)}(T_R) : (x, y, z) \in T(\varrho; R) \right\}, \quad (2.6) \]
are homothetic in \( \mathbb{R}^3 \).

\(^4\) Any \((\rho, z), \rho > 0\), lies on the circle \( C(\varrho^+) = C(\varrho^-) \), where
\[ \varrho^\pm = \frac{(\rho^2 + z^2 + R^2 - 1) \pm \sqrt{(\rho^2 + z^2 + R^2 - 1)^2 - 4\rho^2(R^2 - 1)}}{2\rho}. \]
We divide the proof into 3 steps.

Proof: We divide the proof into 3 steps.

1° By rotational symmetry, the shape of $i_{(\rho, \cos(\theta), \rho \sin(\theta), z)}(T_R)$ is independent of $\theta$. So it suffices to prove that all cyclides of the form

$$i_{(\rho, 0, z)}(T_R), \quad (\rho, z) \in C(\varrho),$$

are homothetic.

By Lemma 2.1, the Euclidean shape of $i_{(\rho, 0, z)}(T_R)$ is determined by the measurements of its cross section at the $x$-$z$ plane. Denote by $P$ the $x$-$z$ plane and $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ be the ortho-projection map onto $P$. Note that

$$\pi \left( i_{(\rho, 0, z)}(T_R) \cap P \right) = i_{(\rho, z)}(\pi(T_R \cap P)).$$

(2.7)

Here $i_{(\rho, z)}$ stands for the circle inversion map in 2-D with respect to the unit circle centered at $(\rho, z)$. Note that $P$ is a symmetry plane of the cyclides (2.7) and that the cross section (2.7) consists of a circle pair. Therefore, by (the implication of) Lemma 2.1, it suffices to check that these circle pairs corresponding to different $(\rho, z) \in C(\varrho)$ are all homothetic. We have reduced the problem into one of plane geometry.

2° We recall a well-known fact about circle inversion. If we invert two circles centered at $(x_1, 0)$ and $(x_2, 0)$ with radii $r_1$ and $r_2$ about a circle centered anywhere on the line

$$\left\{ (x_{ra}, y) \mid x_{ra} = \frac{(x_2^2 - x_1^2) + (r_1^2 - r_2^2)}{2(x_2 - x_1)} \right\},$$

(2.8)

the resulting circle pair is homothetic to the original circle pair. This line is called the radical axis of the circle pair; see Figure 4.

We first determine the image of the circle pair $\pi(T_R \cap P)$ under the circle inversion $i_{(\varrho, 0)}$. The circle pairs in $\pi(T_R \cap P)$ consist of two unit circles with diameters $A_1B_1$ and $A_2B_2$, both on the $x$-axis, with $A_1 = (R - 1, 0)$, $B_1 = (R + 1, 0)$, $A_2 = (-R + 1, 0)$ and $B_2 = (-R - 1, 0)$. The images of $A_1, B_1, A_2, B_2$ under $i_{(\varrho, 0)}$, denoted by $A'_1, B'_1, A'_2, B'_2$, again lie on the $x$-axis and form the diameters $A'_1B'_1, A'_2B'_2$ of circle pair in $i_{(\varrho, 0)}(\pi(T_R \cap P))$.

- When $\varrho \in (0, R - 1)$, $B'_2 < A'_2 < B'_1 < A'_1$.
Figure 4: Radical Axis of a circle pair: (left) two circles exterior to each other; (right) one circle lying inside the other. The blue circles meet the circle pair orthogonally.

- When \( \varrho \in (R + 1, \infty) \), \( B_1' < A_1' < B_2' < A_2' \).
- When \( \varrho \in (R - 1, R + 1) \), \( A_1' < B_2' < A_2' < B_1' \).

In the first two cases, the circle pair are exterior of each other, as in Figure 4(a); in the last case, one circle lies inside the other, as in Figure 4(b). In any case, the resulting circle pair has the following radii and centers:

\[
\begin{align*}
    r_1 &= \frac{|A_1' - B_1'|}{2} = \frac{1}{(\varrho - R)^2 - 1}, \quad r_2 = \frac{|A_2' - B_2'|}{2} = \frac{1}{(\varrho + R)^2 - 1} \quad \text{(2.9)}
\end{align*}
\]

\[
\begin{align*}
    O_1 &= \frac{A_1' + B_1'}{2} = \left( \varrho - \frac{\varrho - R}{(\varrho - R)^2 - 1} \cdot 0 \right), \quad O_2 = \frac{A_2' + B_2'}{2} = \left( \varrho - \frac{\varrho + R}{(\varrho + R)^2 - 1} \cdot 0 \right).
\end{align*}
\]

By (2.8) and (2.9), the radical axis of the circle pair \( i_{(\varrho,0)}(\pi(T_R \cap P)) \) is given by \( \{ (\rho_{ra}, z) : z \in \mathbb{R} \} \) where

\[
    \rho_{ra} = \varrho - \frac{\varrho}{\varrho^2 + 1 - R^2}.
\]

Now the circle pairs in

\[
\left\{ i_{(\rho_{ra}, z)} \circ i_{(\varrho,0)}(\pi(T_R \cap P)) : z \in \mathbb{R} \right\}
\]

are all homothetic. The theorem is proved if we show that every circle pair in \( \left\{ i_{(\rho,z)}(\pi(T_R \cap P)) : (\rho, z) \in \mathcal{C}(\varrho) \right\} \) is homothetic to some circle pair in (2.10). We do so in the last step of the proof.

3° Since an arbitrary composition of inversions can be written as a composition of an inversion (of radius 1) with a homothety (see [1, Page 92]),

\[
i_{(\rho_{ra}, z)} \circ i_{(\varrho,0)} = \mathcal{H} \circ i_{(\rho_{1}, z_1)}. \quad \text{(2.11)}
\]

We can determine \( (\rho_1, z_1) \) using the following properties of an inversion \( i_O \) to find \( (\varrho_1, z_1) \): \( i_O(O) = \infty \), \( i_O(\infty) = O \), and \( i_O(Q_1) = Q_2 \Leftrightarrow i(Q_2) = Q_1 \). By the first property,

\[
i_{(\rho_{ra}, z)} \circ i_{(\varrho,0)}(\rho_{1}, z_1) = \mathcal{H} \circ i_{(\rho_{1}, z_1)}(\rho_{1}, z_1) = \infty.
\]

By the second property,

\[
i_{(\varrho,0)}(\rho_{1}, z_1) = (\rho_{ra}, z).
\]

\[5\text{Here and below, } A < B \text{ simply means } A \text{ is on the left of } B \text{ for two points } A \text{ and } B \text{ are on the first axis of } \mathbb{R}^2.\]
By the third property,
\[(\rho_1, z_1) = i_{(\epsilon, 0)}(\rho_{ra}, z),\]
This means the set of all \((\rho_1, z_1)\) in (2.11) is the image of the line \(\{(\rho_{ra}, z) | z \in \mathbb{R}\}\) under the inversion \(i_{(\epsilon, 0)}\), which is a circle. By symmetry, this circle has a diameter on the x-axis. One end of the diameter is \(i_{(\epsilon, 0)}((\rho_{ra}, \infty)) = (\varrho, 0)\), and the other end is \(i_{(\epsilon, 0)}((\rho_{ra}, 0)) = \frac{R^2 - 1}{\varrho}\). The circle is \(C(\varrho)\).

In virtue of Theorem 2.2, we use the shorthand notation
\[i_\varrho(T_R)\]
to represent the common Euclidean shape of the cyclides in (2.6). Formally, \(i_\varrho(T_R)\) is an element in the shape space (2.2).

To further analyze the shape \(i_\varrho(T_R)\), by Lemma 2.1 and Theorem 2.2, it suffices to analyze the ratio \(r_1 : r_2 : d\) of the cross-section of \(i_{(\epsilon,0,0)}(T_R)\) at its \(P_i\) symmetry plane.

**Lemma 2.3.** For any \(R \in (1, \infty)\), the \(P_i\) cross section of \(\mathcal{C} = i_{(\epsilon,0,0)}(T_R)\) has the following measurements:

1. When \(\varrho \in [0, R - 1)\) (corresponding to the red circles in Figure 3(a)), the \(P_i\) symmetry plane of \(\mathcal{C}\) is the \(x\)-\(z\) plane, and
   \[r_1 : r_2 : d = \lambda : 1 : \frac{1}{\sqrt{(\lambda - 1)^2 + 4 \lambda R^2}}, \quad \text{where} \quad \lambda = \frac{r_1}{r_2} = \frac{(\varrho + R)^2 - 1}{(\varrho - R)^2 - 1} \in [1, \infty), \quad (2.12)\]
2. When \(\varrho \in (R - 1, \sqrt{R^2 - 1}]\) (corresponding to the blue circles in Figure 3(a)), the \(P_i\) symmetry plane of \(\mathcal{C}\) is the \(x\)-\(y\) plane, and
   \[r_1 : r_2 : d = \lambda : 1 : \frac{\sqrt{(\lambda - 1)^2 + 4 \lambda \frac{R^2}{R^2 - 1}}}{(\varrho + R)^2 - 1}, \quad \text{where} \quad \lambda = \frac{r_1}{r_2} = \frac{(R - 1)(\varrho + 1)^2 - \varrho^2}{(R + 1)(\varrho^2 - (R - 1)^2)} \in [1, \infty). \quad (2.13)\]

**Proof:** The first two steps of the proof of Theorem 2.2 imply that
\[P,\text{ the x-z plane, is } \begin{array}{l}
\text{the } P_1 \text{ symmetry plane of } \mathcal{C} \text{ when } \varrho \in [0, R - 1) \\
\text{the } P_2 \text{ symmetry plane of } \mathcal{C} \text{ when } \varrho \in (R - 1, \sqrt{R^2 - 1}] \\
\end{array}.
\]
In the first case, \(r_1\) and \(O_i\) in (2.9) are such that \(r_1 > r_2\) and \(O_2 < O_1\), and they give the \(r_1, r_2, d\) measurements of \(\mathcal{C}\):
\[r_1 = \frac{1}{(\varrho - R)^2 - 1}, \quad r_2 = \frac{1}{(\varrho + R)^2 - 1}, \quad d = \frac{\varrho + R}{(\varrho + R)^2 - 1} - \frac{\varrho - R}{(\varrho - R)^2 - 1}. \quad (2.14)\]
In the second case, we also have \(r_1 > r_2\) but now \(O_1 < O_2\), and they give the \(\tilde{r}_1, \tilde{r}_2, \tilde{d}\) measurements of \(\mathcal{C}\):
\[\tilde{r}_1 = \frac{1}{1 - (\varrho - R)^2}, \quad \tilde{r}_2 = \frac{1}{(\varrho + R)^2 - 1}, \quad \tilde{d} = -\frac{\varrho + R}{(\varrho + R)^2 - 1} + \frac{\varrho - R}{(\varrho - R)^2 - 1}.
\]
By (2.4), we can convert the \((\tilde{r}_1, \tilde{r}_2, \tilde{d})\) measurements to the \((r_1, r_2, d)\) measurements via \(r_1 = (\tilde{r}_1 - \tilde{r}_2 + \tilde{d})/2, \quad r_2 = (\tilde{r}_1 - \tilde{r}_2 - \tilde{d})/2, \quad d = \tilde{r}_1 + \tilde{r}_2, \tilde{d}\), so
\[r_1 = \frac{R - 1}{\varrho^2 - (R - 1)^2}, \quad r_2 = \frac{R + 1}{(R + 1)^2 - \varrho^2}, \quad d = \frac{1}{(R + \varrho)^2 - 1} - \frac{1}{(R - \varrho)^2 - 1}. \quad (2.15)\]
By routine computations, (2.12) follows from (2.14) and (2.13) follows from (2.15).

Lemma 2.3 has an almost immediate consequence:
Theorem 2.4. For any $R \in (1, \infty)$, $i_{\varrho}(T_R)$ is distinct for each $\varrho \in [0, R - 1]$.

- If $R \neq \sqrt{2}$, then $i_{\varrho}(T_R)$ is distinct for each $\varrho \in [0, R - 1) \cup (R - 1, \sqrt{R^2 - 1}]$.
- If $R = \sqrt{2}$, then $\varrho > 2 - 1$ adds no new shape and hence the shape space (2.1) is in one-to-one correspondence with

$$\left\{ i_{\varrho}(T_{\sqrt{2}}) : \varrho \in [0, \sqrt{2} - 1] \right\}.$$ 

Proof: Recall the two expressions in Lemma 2.3 for $\lambda = \frac{r_1}{r_2}$ in the two intervals of $\varrho$. It is easy to check that both

$$\lambda_1 : [0, R - 1) \rightarrow [1, \infty), \quad \lambda_1(\varrho) = \frac{(\varrho + R)^2 - 1}{(\varrho - R)^2 - 1}$$

and

$$\lambda_2 : (R - 1, \sqrt{R^2 - 1}] \rightarrow [1, \infty), \quad \lambda_2(\varrho) = \frac{(R - 1)((R + 1)^2 - \varrho^2)}{(R + 1)(\varrho^2 - (R - 1)^2)}$$

are bijections: simply check that $\lambda_1$ is monotonic increasing from 1 to $\infty$, and $\lambda_2$ is monotonic decreasing from $\infty$ to 1. As the $r_1 : r_2$ ratio of $i_{\varrho}(T_R)$ is distinct for different $\varrho \in [0, R - 1)$, the first statement of the theorem is true. Likewise, $i_{\varrho}(T_R)$ is also distinct for each $\varrho \in (R - 1, \sqrt{R^2 - 1}]$, $i_{\varrho_1}(T_R) \neq i_{\varrho_2}(T_R)$. There are two cases:

1. If $\lambda_1(\varrho_1) \neq \lambda_2(\varrho_2)$, then $i_{\varrho_1}(T_R) \neq i_{\varrho_2}(T_R)$.
2. If $\lambda_1(\varrho_1) = \lambda_2(\varrho_2)$, then, by the expressions of the $r_2 : d$ ratio in Lemma 2.3, the $r_2 : d$ ratios of $i_{\varrho_1}(T_R)$ and $i_{\varrho_2}(T_R)$ are different exactly when $R^2 \neq \frac{R^2}{R^2 - 1}$. But

$$R^2 = \frac{R^2}{R^2 - 1} \iff R = \sqrt{2}.$$

So we also have $i_{\varrho_1}(T_R) \neq i_{\varrho_2}(T_R)$ in this case.

This argument proves the statement under the second bullet as well.

The next two results characterize the bigger shape space (2.2); they are inspiring for us but technically we do not need them for this article. We omit the detailed proofs, which follow the same line of arguments as in that of Theorem 2.4.

Lemma 2.5. For any $R \in (1, \infty)$, $\varrho \in [0, \sqrt{R^2 - 1}]$,

$$i_{\varrho}(T_R) = i_{\varrho'}(T_{R'}) \quad \text{where} \quad (R', \varrho') = \frac{1}{\sqrt{R^2 - 1}} \left( R, \frac{\sqrt{R^2 - 1} - \varrho}{\sqrt{R^2 - 1} + \varrho} \right).$$

(2.16)

Theorem 2.6. Let

$$C_R := \begin{cases} [0, \sqrt{R^2 - 1}] \setminus \{ R - 1 \} & \text{if } R \in (1, \sqrt{2}), \\ [0, \sqrt{2} - 1] & \text{if } R = \sqrt{2} \end{cases}, \quad C := \bigcup_{R \in (1, \sqrt{2})} \{ (R, \varrho) : \varrho \in C_R \}.$$ 

Distinct elements in $C$ correspond to distinct $i_{\varrho}(T_R)$ and the shape space (2.2) is in one-to-one correspondence with

$$\left\{ i_{\varrho}(T_R) : (R, \varrho) \in C \right\}.$$ 

8
3 Step II: Rounding by sphere inversion

Theorem 3.1. If \( S \) is a compact regular surface (with or without boundary) in \( \mathbb{R}^3 \) and \( p \in S \), then

\[
\text{Area}(i_q(S)) \sim \frac{\pi}{|p - q|^2}, \quad q \to p, \quad \overline{pq} \perp T_p S.
\]  

If \( S \) is also closed and orientable (so \( S \) and \( i_q(S) \) have enclosing volumes), then

\[
\text{Volume}(i_q(S)) \sim \frac{\pi}{6|p - q|^3}, \quad q \to p, \quad \overline{pq} \perp T_p S,
\]  

and (consequently)

\[
v(i_q(S)) = \frac{\text{Volume}(i_q(S))}{(4\pi/3)(\text{Area}(i_q(S))/(4\pi))^{3/2}} \to 1, \quad q \to p, \quad \overline{pq} \perp T_p S.
\]

Proof: Without loss of generality assume \( p = (0,0,0) \) and \( T_p S \) is the \( x \)-\( y \)-plane, and let \( \varepsilon \) be a small scalar representing the point \( q = (0,0,\varepsilon) \) approaching the surface orthogonally at the origin. So the surface near \( p \) can be written as the graph of a smooth function \( h(x,y) \), where \( x^2 + y^2 < R^2 \) for some \( R > 0 \) and \( h \) has a vanishing linear approximation at the origin, i.e. \( h(0,0) = 0 = \frac{\partial h}{\partial x}(0,0) = \frac{\partial h}{\partial y}(0,0) \), and so

\[
h(x, y) = O(x^2 + y^2), \quad |\nabla h(x,y)| = O(\sqrt{x^2 + y^2}), \quad (x, y) \to (0, 0).
\]  

Write \( S_\varepsilon := \{(x,y,h(x,y)) : x^2 + y^2 < R^2\} \). By continuity, the area of \( i_{(0,0,\varepsilon)}(S_\varepsilon) \) approaches that of \( i_{(0,0,0)}(S \setminus S_\varepsilon) \) as \( \varepsilon \to 0 \) and hence stays bounded for small \( \varepsilon \). So it suffices to prove (3.1) with \( S \) replaced by \( S_\varepsilon \).

The conformal factor of \( i_\varepsilon \) is \( \lambda^2(a,x) = 1/\|x-a\|^4 \), i.e. \( \langle di_a|_x v, di_a|_x w \rangle = \lambda^2(a,x) \langle v, w \rangle \). Therefore,

\[
\text{Area}(i_{(0,0,\varepsilon)}(S_\varepsilon)) = \int_{x^2+y^2 < R^2} \sqrt{1 + |\nabla h(x,y)|^2} \left|\frac{h(x,y) - \varepsilon}{\sqrt{x^2 + y^2 + (h(x,y) - \varepsilon)^2}}\right|^2 dxdy
\]

\[
= \int_0^{2\pi} \int_0^R \sqrt{1 + |\nabla h(re^{i\theta})|^2} \left|\frac{rre^{i\theta} - \varepsilon}{\sqrt{r^2 + (r - h(re^{i\theta}))^2}}\right|^2 rdrd\theta. \tag{3.5}
\]

Let \( r_\varepsilon(\varepsilon) = |\varepsilon|^{\alpha} \) for any \( \alpha \in (1/2, 1) \) so that

(i) \( |\varepsilon| = o(r_\varepsilon(\varepsilon)) \) and (ii) \( r_\varepsilon(\varepsilon) = o(|\varepsilon|^{1/2}) \), as \( \varepsilon \to 0 \). \tag{3.6}

We then split the inner integral in (3.5) into \( \int_0^{r_\varepsilon(\varepsilon)} + \int_{r_\varepsilon(\varepsilon)}^R \); define

\[
J(\varepsilon) := \int_0^{2\pi} \int_0^{r_\varepsilon(\varepsilon)} \sqrt{1 + |\nabla h(re^{i\theta})|^2} \left|\frac{rre^{i\theta} - \varepsilon}{\sqrt{r^2 + (r - h(re^{i\theta}))^2}}\right|^2 rdrd\theta, \quad K(\varepsilon) := \int_0^{2\pi} \int_{r_\varepsilon(\varepsilon)}^R \sqrt{1 + |\nabla h(re^{i\theta})|^2} \left|\frac{rre^{i\theta} - \varepsilon}{\sqrt{r^2 + (r - h(re^{i\theta}))^2}}\right|^2 rdrd\theta.
\]

We shall prove (3.1) by showing that the former integral is asymptotically equivalent to \( \varepsilon^{-2}/2 \) and the latter grows slower than \( \varepsilon^{-2} \).

For \( J(\varepsilon) \), we compare it with the special case when \( h \equiv 0 \). By (3.4), there exists a constant \( C > 0 \), independent of \( r \) and \( \theta \), such that

\[
|\nabla h(re^{i\theta})|^2, \quad |h(re^{i\theta})| \leq C r^2.
\]

For \( r \in [0, r_\varepsilon(\varepsilon)] \), \( r^2 \leq r_\varepsilon(\varepsilon)^2 = o(|\varepsilon|) \) by (3.6)(ii), so \( \varepsilon - h(re^{i\theta}) \sim \varepsilon \). Also, \( 1 + |\nabla h(re^{i\theta})|^2 \sim 1 \). From this it is easy to see that

\[
J(\varepsilon) \sim \int_0^{2\pi} \int_0^{r_\varepsilon(\varepsilon)} \frac{r}{r^2 + \varepsilon^2} drd\theta. \tag{3.7}
\]
The right-hand side is \(J(\varepsilon)\) in the case of \(h \equiv 0\), whose asymptotic can be easily determined:
\[
\int_0^{2\pi} \int_0^{r_r(\varepsilon)} \frac{r}{\sqrt{r^2 + \varepsilon^2}} dr d\theta = \frac{2\pi}{\varepsilon^2} \int_0^{r_r(\varepsilon)/\varepsilon} s ds \left(1 + s^2\right)^{1/2} = \frac{2\pi}{\varepsilon^2} \left[\frac{1}{2} - \frac{1}{2(1 + (r_r(\varepsilon)/\varepsilon)^2)}\right] \sim \frac{\pi}{\varepsilon^2}, \quad \varepsilon \to 0.
\]

(3.8)

In the last step above, we used (3.6)(i).

For \(K(\varepsilon)\), note that \(\nabla h\) is bounded on a compact set, so
\[
K(\varepsilon) = \int_0^{2\pi} \int_{r_r(\varepsilon)}^R \frac{1}{\sqrt{1 + \left|\nabla h(r\varepsilon)\right|^2}} r^2 dr d\theta 
\leq 2\pi \int_{r_r(\varepsilon)}^R \frac{C}{\varepsilon^2} r^2 dr 
\leq 2\pi C \int_{r_r(\varepsilon)}^\infty r^{-3} dr = \pi C r_{\ast}(\varepsilon)^{-2} = o(\varepsilon^{-2}).
\]

In the last step above, we again used (3.6)(i).

We have completed the proof of (3.1).

Let \(B\) be a ball whose boundary is tangent to \(S\) at \(p\) and lies inside of \(S\), so Volume\((B) \leq\) Volume\((S)\) and also
\[
\text{Volume}(i_q(B)) \leq \text{Volume}(i_q(S)).
\]

As before, write \(|p - q| = \varepsilon\). Since \(i_q(B)\) is a ball with diameter \(\sim 1/\varepsilon\),
\[
\text{Volume}(i_q(B)) \sim \frac{4\pi}{3} \left(\frac{1}{2\varepsilon}\right)^3 = \frac{\pi}{6\varepsilon^3}, \quad \varepsilon \to 0.
\]

So Volume\((i_q(S))\) grows at least as fast as \(\pi/(6\varepsilon^3)\). By the first part of the theorem and the isoperimetric inequality, Volume\((i_q(S))\) cannot grow faster than \(\pi/(6\varepsilon^3)\), and (3.2) is proved.

Remark 3.2. We thank I. Pinelis for the help in analyzing the asymptotic of the area integral (3.5); see https://mathoverflow.net/questions/353648/asymptotic-of-an-area-integral.

4 Step III: Reduction to P-recurrence

In this section we express by P-recurrences the surface area and enclosing volume of \(i_a(T_{\sqrt{\varepsilon}})\), where \(a = [a, 0, 0]^T, a \in [0, \sqrt{2} - 1]\), which are the same as those of SCT\(_a(T_{\sqrt{\varepsilon}})\). (Recall \(i(T_{\sqrt{\varepsilon}}) = T_{\sqrt{\varepsilon}}\).) From these, an associated P-recurrence related to the isoperimetric ratio of \(i_a(T_{\sqrt{\varepsilon}})\) will also be derived.

4.1 Area and volume integrals

The conformal factor of a special conformal transformation SCT\(_a := i \circ t_a \circ i\) is
\[
\lambda^2(a, x) = \frac{1}{(1 + 2(a, x) + \langle a, a, \langle x, x\rangle\rangle^2),}
\]
i.e. \(\langle dS_a|_v, dS_a|_w\rangle = \lambda^2(a, x)\langle v, w\rangle\). So the area and enclosing volume of SCT\(_{[a, 0, 0]}(T_{\sqrt{\varepsilon}})\) are given by
\[
A(a) = \int_0^{2\pi} \int_0^{2\pi} Q(a; x)^{-2} d\text{Area}(u, v), \quad V(a) = \int_0^1 \int_0^{2\pi} Q(a; x)^{-3} d\text{Vol}(u, v, r),
\]
where
\[
Q(a; x) := \frac{1}{\lambda^2([a, 0, 0]|^T, x)} = 1 + 2a_1 a + ||x||^2 a^2,
\]
\[
x(u, v, r) = \begin{bmatrix} \sqrt{2} + r \sin(v) \cos(u), & \sqrt{2} + r \sin(v) \sin(u), & r \cos(v) \end{bmatrix}, \quad u, v \in [0, 2\pi], \quad r \in [0, 1],
\]
\[
d\text{Area}(u, v) = (\sqrt{2} + \sin(v)) du dv, \quad d\text{Vol}(u, v, r) = r(\sqrt{2} + r \sin(v)) du dv dr.
\]

Notice also that
\[
\langle x, x\rangle = ||x||^2 = 2 + r^2 + 2\sqrt{2} r \sin(v).
\]
4.2 Holomorphic extension

The integral definitions of $A$ and $V$ above extend from the interval $[0, \sqrt{2} - 1)$ to a holomorphic function on the open disk

$$D := \{ z \in \mathbb{C} : |z| < \sqrt{2} - 1 \}.$$

To see this, note that the roots of $Q(z; x)$, viewed as a quadratic polynomial in $z$, can be expressed as

$$-x_1 \pm i \sqrt{x_2^2 + x_3^2} \over |x|^2,$$

so their moduli are both $1/|x|$. But $x$ is a point in the boundary or interior of the solid torus $T$, so $|x| \in [\sqrt{2} - 1, \sqrt{2} + 1]$, which is equivalent to $1/|x| \in [\sqrt{2} - 1, \sqrt{2} + 1]$. This means

$$Q(z; x) \neq 0, \quad \forall z \in D, \ x \in T.$$

Therefore $Q(z; x)^{-L}$, $L = 2$ or $3$, is holomorphic in the first argument and continuous in the second. A standard argument in complex analysis shows that $A$ and $V$, defined based on the integrals in (4.1), extend to holomorphic functions on $D$.

So from now on, we write $A(z)$ and $V(z)$ instead of $A(a)$ and $V(a)$.

4.3 Power series at $z = 0$

Since $A$ and $V$ are even functions, the odd power Taylor coefficients at $z = 0$ all vanish. Denote by $a_j$ and $v_j$ the coefficients of $z^{2j}$ in the expansions of $A(z)$ and $V(z)$ at $z = 0$, respectively. An observation here is that

$$\frac{d^n A}{dz^n}(0) = \int_0^\pi \left. \frac{d^n}{dz^n} Q(z; x(u, v, 1))^{-2} \right|_{z=0} d\text{Area}(u, v),$$

$$\frac{d^n V}{dz^n}(0) = \int_0^\pi \left. \frac{d^n}{dz^n} Q(z; x(u, v, r))^{-3} \right|_{z=0} d\text{Vol}(u, v, r),$$

and, thanks to the evaluation at $z = 0$, the integrands above are polynomials in $x_1$ and $|x|^2$, hence are trigonometric polynomials in $(u, v)$.

Using either (4.1) or the generalized binomial theorem to expand $Q(z; x)^{-L}$ into a power series of $z$, i.e.

$$Q(z; x)^{-L} = \sum_{n=0}^\infty \left( \frac{n + L - 1}{n} \right) (-1)^n (2x_1 z + |x|^2 z^2)^n,$$

together with the identity (of Wallis’ integrals):

$$\int_0^{2\pi} \cos^n(v) \, dv = \int_0^{2\pi} \sin^n(v) \, dv = \begin{cases} \frac{2\pi}{\sin\left(\frac{n\pi}{2}\right)}, & n \text{ even } \\ 0, & n \text{ odd } \end{cases},$$

we have

$$a_j = \sum_{\ell=0}^{j} (-1)^{j-\ell} (j + \ell + 1) \left( \frac{j + \ell}{j - \ell} \right) \int_0^{2\pi} \int_0^{2\pi} (2x_1(u, v, 1))^2 \|x(u, v, 1)\|^{2(j-\ell)} d\text{Area}(u, v),$$

$$= \sum_{\ell=0}^{j} (-1)^{j-\ell} (j + \ell + 1) \left( \frac{j + \ell}{j - \ell} \right) 4^\ell \int_0^{2\pi} \cos^{2\ell}(u) \, du \int_0^{2\pi} (\sqrt{2} + \sin(v))^{2\ell+1} (3 + 2\sqrt{2} \sin(v))^{j-\ell} dv.$$
\[ a_j = \sqrt{2\pi}^2 \sum_{\ell=0}^{j} (-1)^{j-\ell} (j+\ell+1) \binom{j+\ell}{j-\ell,\ell,\ell} \alpha_{\ell,j}, \]

\[ \alpha_{\ell,j} = 2^{\ell+2} 3^{-\ell} \sum_{p=0}^{2\ell+1} \sum_{q=0}^{j-\ell} \binom{2\ell+1}{p} \binom{j-\ell}{q} \left( \binom{p+q}{p+q}/2 \right) 2^{(q-3p)/2} 3^{-q}. \]  

(4.2)

Similarly,

\[ v_j = \sum_{\ell=0}^{j} (-1)^{j-\ell} \binom{j+\ell+1}{j-\ell,\ell,\ell} 2^{\ell+1} \int_0^1 \int_0^{2\pi} \binom{2\ell+1}{p} \binom{j-\ell}{q} \left( \binom{p+q}{p+q}/2 \right) 2^{(q-3p)/2} \eta_{p,q,\ell,j} \]

\[ \eta_{p,q,\ell,j} = \int_0^{2\pi} r(\sqrt{2 + r \sin(v)})^{2\ell+1} (2 + r^2 + 2\sqrt{2r} \sin(v))^{j-\ell} dr. \]

So,

\[ v_j = \sqrt{2\pi}^2 \sum_{\ell=0}^{j} (-1)^{j-\ell} (j+\ell+1) (j+\ell+2) \binom{j+\ell}{j-\ell,\ell,\ell} v_{\ell,j}, \]

\[ \nu_{\ell,j} = 2^{\ell+1} \sum_{p=0}^{2\ell+1} \sum_{q=0}^{j-\ell} \binom{2\ell+1}{p} \binom{j-\ell}{q} \left( \binom{p+q}{p+q}/2 \right) 2^{(q-3p)/2} \eta_{p,q,\ell,j}, \]  

(4.3)

And we have the following power series:

\[ \frac{1}{\sqrt{2\pi}^2} A(z) = 4 + 52 z^2 + 477 z^4 + 3809 z^6 + \frac{451625}{16} z^8 + \cdots \]

\[ \frac{1}{\sqrt{2\pi}^2} V(z) = 2 + 48 z^2 + \frac{1269}{2} z^4 + 6600 z^6 + \frac{1928025}{32} z^8 + \cdots \]

By the expressions (4.2)-(4.3), \( \frac{1}{\sqrt{2\pi}^2} a_n, \frac{1}{\sqrt{2\pi}^2} v_n \) are rational.

### 4.4 Isoperimetric Ratio

To show that the isoperimetric ratio of \( \text{SCT}_{[a,0,0]}(T\sqrt{2}) \) is monotonic increasing in \( a \in [0, \sqrt{2} - 1] \), it suffices to show

\[ \Delta(a) := \frac{d}{da} \ln \frac{V(a)}{A(a)} = 2 \frac{V'(a)}{V(a)} - 3 \frac{A'(a)}{A(a)} > 0, \quad \text{or} \quad 2V'(a)A(a) - 3V(a)A'(a) > 0. \]

It happens that \( \Delta(a) \) is proportional to the distance between the area and volume centers of the cyclide \( \text{SCT}_{[a,0,0]}(T) \). Precisely, \( \Delta(a) = 12 [x^A(a) - x^V(a)] \) where \( x^A(a) \) and \( x^V(a) \) are the first coordinates of the area and volume centers of \( \text{SCT}_{[a,0,0]}(T\sqrt{2}) \), respectively.\(^6\) This follows from the observation that \( (\text{SCT}_{[a,0,0]} \circ x)_1 = \frac{1}{2} Q'(a;x)/Q(a;x) \).

\(^6\)According to this formula, Theorem 1.1 is proved if one establishes that \( x^A(a) \neq x^V(a) \) for all \( a \in (0, \sqrt{2} - 1) \). At one point, the first author presented an incomplete argument to R. Kusner based on this approach. He thanks Kusner for helping him to see the flaw in his argument.
By the Taylor expansions of $A(a)$ and $V(a)$, we have

$$
\frac{1}{2\pi i} \left( 2V'(a)A(a) - 3V(a)A'(a) \right) = \sum_k \left[ 2(v_1 a_k + 2v_2 a_{k-1} + \cdots + (k+1)v_{k+1}v_0) - 3(a_1 v_k + 2a_2 v_{k-1} + \cdots + (k+1)a_{k+1}v_0) \right] a^{2k+1} = 72 a + 1932 a^3 + 31248 a^5 + \frac{790101}{2} a^7 + \frac{17208645}{4} a^9 + \cdots
$$

(4.4)

4.5 P-recurrences

The combinatorial expressions (4.2)-(4.3), together with the closure properties of holonomic sequences [23, 20, 8], show that $(a_n)_{n \geq 0}$ and $(v_n)_{n \geq 0}$ are P-recursive, i.e. they satisfy linear recurrences with polynomial coefficients. Equivalently, their generating functions, namely

$$
\bar{A}(z) = \sum_{n \geq 0} a_n z^n, \quad \bar{V}(z) = \sum_{n \geq 0} v_n z^n,
$$

are holonomic or $D$-finite, i.e. they satisfy linear differential equations with polynomial coefficients. The generating functions of $(a_n)_{n \geq 0}$ and $(v_n)_{n \geq 0}$ are related to the original area and volume functions $A(z)$ and $V(z)$ simply by $A(z) = \bar{A}(z^2)$ and $V(z) = \bar{V}(z^2)$. The generating function of the sequence $(d_n)_{k \geq 0}$, defined by (4.4), is given by

$$
\bar{D}(z) := \sum_{n=0}^{\infty} d_n z^n = 2\bar{V}'(z)\bar{A}(z) - 3\bar{V}(z)\bar{A}'(z).
$$

Since holonomic functions are closed under Hadamard product (hence differentiation), product, and linear combination, $(d_n)_{n \geq 0}$ is also holonomic.

**Proposition 4.1.** The P-recurrences of $(a_n)_{n \geq 0}$, $(v_n)_{n \geq 0}$ and $(d_n)_{n \geq 0}$ are given by

$$
\sum_{i=0}^{3} p_i(n) a_{n+i} = 0, \text{ where } \begin{bmatrix} p_0(n) \\ p_1(n) \\ p_2(n) \\ p_3(n) \end{bmatrix} = \begin{bmatrix} -84 & -136 & -81 & -21 & -2 \\ 399 & 730 & 484 & 137 & 14 \\ -474 & -835 & -529 & -143 & -14 \\ 54 & 99 & 66 & 19 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ n \\ n^2 \\ n^3 \\ n^4 \end{bmatrix}
$$

(4.5)

$$
\sum_{i=0}^{3} q_i(n) v_{n+i} = 0, \text{ where } \begin{bmatrix} q_0(n) \\ q_1(n) \\ q_2(n) \\ q_3(n) \end{bmatrix} = \begin{bmatrix} -252 & -303 & -136 & -27 & -2 \\ 960 & 1384 & 730 & 167 & 14 \\ -1008 & -1436 & -748 & -169 & -14 \\ 90 & 141 & 82 & 21 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ n \\ n^2 \\ n^3 \\ n^4 \end{bmatrix}
$$

(4.6)

$$
\sum_{i=0}^{7} r_i(n) d_{n+i} = 0, \text{ where } \begin{bmatrix} r_0(n) & r_1(n) & \ldots & r_7(n) \end{bmatrix}^T = M[1, n, n^2, \ldots, n^7]^T,
$$

where

$$
M = \begin{bmatrix}
1630207484 & -3179176725 & -560587685 & -1216898711 & -1675359251 & -626799 & -71441 & -115 \\
12520515306 & -2703017807 & -5096689337 & -1191234237 & -1676596733 & -572961 & -6767 & -11777 \\
9722946318 & -20622045951 & -39842434525 & -9281378227 & -14746920065 & -505661 & -5903 & -11215 \\
3650349795 & -7614819665 & -15607466733 & -3487178227 & -5663849267 & -1889656 & -2089 & -3838 \\
14098358975 & -30355428855 & -6455795533 & -14522451277 & -23184921987 & -7950599 & -90372 & -15832 \\
49580268527 & -10852868633 & -24206298455 & -54584142237 & -93239713222 & -32846953 & -3635 & -6483 \\
31661835197 & -67982352837 & -15058059693 & -33741412337 & -56713045699 & -2127735 & -2399 & -429 \\
290565640325 & -62775118251 & -15218545533 & -34012038487 & -5985443365 & -23547541 & -2562 & -463 \\
3065 & 61331596 & 64627498 & 183738067 & 36248415 & 12252 & 356 & 139 & 1
\end{bmatrix}
$$

(4.7)

Moreover, these are the only P-recurrences with the corresponding order $(r)$ and degree $(d)$ for the three sequences. (E.g., (4.5) is the only P-recurrence with $(r, d) = (3, 4)$ satisfied by the sequence $(a_n)$.)
As one may expect from the appearance of the result, the proof is computer-assisted. The first part of the proposition, namely, the sequences defined by (4.2)-(4.4) satisfy the P-recurrences (4.5)-(4.7), can be established by a refinement of Zeilberger’s creative telescoping method \cite{23} due to Koutschan \cite{11} (implemented in his Mathematica package HolonomicFunctions.) We can check the second part of the claim in an elementary fashion. Assume that we have established that \((a_n)\) follows a P-recurrence of order \(r = 3\) and degree \(d = 4\), then the \((d + 1)(r + 1) = 20\) coefficients in the polynomials satisfy, for every index \(n\), a homogeneous linear equation with rational coefficients determined by the terms \(a_n, a_{n+1}, a_{n+2}, a_{n+3}\). Using the first \(N + 4\) terms of the sequence \(a_n\) with any \(N \geq 20\), easily computable by (4.2), we can set up a homogeneous linear system that must be satisfied by the 20 coefficients. Using a symbolic linear solver to explicitly work out a basis of the null space of the rational \(N \times 20\) coefficient matrix, and seeing that the basis consists of one vector in \(\mathbb{R}^{20}\) with a certain \(N \geq 20\), would not only prove the claimed uniqueness (up to a scaling factor), but also reproduce the P-recurrence in (4.5). This method is called ‘guessing’ in [8], as it can be used to guess (with high confidence) what the P-recurrence might be when used with a big enough \(N\).

Using asymptotic techniques [16, 7, 21, 5] of holonomic functions, aided by rigorous interval arithmetic computation, it can be shown that

\[
d_n \sim c \cdot (\sqrt{2} + 1)^{2n} n^3 \ln(n), \quad c > 8.
\]  

This is more than enough for proving that \(d_n\) is eventually positive, but is insufficient for verifying full positivity. A proof of positivity, based on streamlining existing complex analytic techniques [16, 5] in order to bound the error in (4.8), can be found in the preprint [17] by Melczer and Mezzarobba.

5 Final Remarks

This paper connects a special case of the theory of Willmore surfaces to the theory of special functions, with the hope that it may mobilize some interests in (i) the more ambitious uniqueness question discussed in [22, Page 1-4] and (ii) the positivity problem of P-recurrence, which is a well-known open problem in combinatorics [9, 18]. The main uniqueness conjecture in [22, Page 1-4] speculates the uniqueness of genus 0 and 1 isoperimetric ratio constrained Willmore surfaces (the so-called Canham problem.) The uniqueness conjecture consists of three cases: (i) the genus 0 case for any isoperimetric ratio \(v \in (0, 1]\) (existence is proved in [19]), (ii) the genus 1 case for \(v \in (0, (3/2)(2\pi^2)^{-1/4}]\) (existence is unknown at the time of writing; see [10]), and (iii) the genus 1 case for \(v \in [(3/2)(2\pi^2)^{-1/4}, 1]\), which correspond to the Clifford torus and the Marques-Neves theorem. This article, together with [17], solves part (iii). Part (i) includes the most representative case of the curious bi-concave shape of red-blood cells. With all likelihood, the solution of part (i) and (ii) will require entirely different techniques. However, it remains to see if the special function approach here applies to the understanding of the higher genus Lawson surface \(\xi_{g,1}\), which is conjectured to be the genus \(g\) Willmore minimizer [12, 6]. (Recall that the Clifford torus corresponds to \(\xi_{1,1}\).)

References


