Abstract. What is the longest rope on the unit sphere? Intuition tells us that the answer to this packing problem depends on the rope’s thickness. For a countably infinite number of prescribed thickness values we construct and classify all solution curves. The simplest ones are similar to the seam lines of a tennis ball; others exhibit a striking resemblance to Turing patterns in chemistry, or to ordered phases of long elastic rods stuffed into spherical shells.

1. THE PROBLEM. What is the longest curve on the unit sphere? The most probable answer of any mathematically inclined person to this naïve question is: There is no such thing, since any spherical curve of finite length can be made arbitrarily long by replacing parts of it by more and more “wavy” arcs; see Figure 1. Rephrasing the initial query as “What is the longest rope on the unit sphere?” makes a big difference. A rope, in contrast to a mathematical curve, forms a solid body with positive thickness, so that now this question addresses a packing problem with obvious parallels in everyday life. Is there an optimal way of winding electrical cable onto a reel? Similarly, and economically quite relevantly, can one maximize the volume of yarn wound onto a given bobbin [18], or how should one store long textile fiber band most efficiently to save storage space [17]?

Figure 1. Without a lower bound on the thickness there is no longest curve on $S^2$. By inserting more and more oscillations into a given curve one can make its length arbitrarily large.

Common to all these packing problems, in contrast to the classic Kepler problem of optimal sphere packing [5, 13], is that long and slender deformable objects are to be placed into a fixed volume or onto a given surface. Nature displays fascinating packing strategies on various scales. Extremely long strands of viral DNA are packed very efficiently into the tiny phage heads of bacteriophages [6], and chromatin fibers are folded and organized in various aggregates within the chromatid [19].

To model a rope as a mathematical curve with positive thickness in Euclidean three-space we follow the approach of Gonzalez and Maddocks [11] who considered all triples of distinct curve points $x, y, z$ on a closed curve $\gamma \subset \mathbb{R}^3$, and their respective circumcircle radii $R(x, y, z)$. The smallest of these radii determines the curve’s...
Figure 2. A positive thickness imposes a lower bound both on radius of curvature (left) as well as on global self-distance (right). The tubular neighborhood around the curve does not self-intersect.

\[
\Delta[\gamma] := \inf_{x \neq y \neq z, x, y, z \in \gamma} R(x, y, z). \tag{1.1}
\]

A positive lower bound on this quantity controls local curvature but also prevents the curve from self-intersections; see Figure 2. In fact, it equips the closed curve with a tubular neighborhood of uniform radius \(\Delta[\gamma]\) without self-penetration. This means that there exists a tubular neighborhood which consists of the disjoint union of open disks normal to and centered at the curve, and each with radius \(\Delta[\gamma]\). Conversely, a continuously differentiable closed curve \(\gamma\) that possesses such a tubular neighborhood of radius \(\Theta\) without self-penetration can be shown to have thickness at least \(\Theta\) [12, Lemma 3]. Furthermore, positive thickness characterizes the set of embedded loops with bounded curvature (see [12, Lemmata 2 & 3] and [21, Theorem 1]), and we therefore tacitly assume from now on that our curves are simple, have positive length, and are continuously differentiable.

With this mathematical concept of thickness at our disposal we can reformulate the original question of finding the longest ropes on the unit sphere as a variational problem, where we first focus on closed loops.

**Problem (P).** For a given constant \(\Theta > 0\) find the longest closed curve \(\gamma : S^1 \cong \mathbb{R}/(2\pi \mathbb{Z}) \to S^2 := \{x \in \mathbb{R}^3 : |x| = 1\}\) with prescribed minimal thickness \(\Theta\), i.e., with \(\Delta[\gamma] \geq \Theta\).

Before discussing the solvability of this maximization problem for various thickness values we would like to point out that every loop \(\gamma \subset \mathbb{R}^3\) of positive thickness enjoys a strong geometric property, the presence of forbidden balls: Any open ball \(B_{\Theta} \subset \mathbb{R}^3\) of radius \(\Theta \leq \Delta[\gamma]\) whose boundary \(\partial B_{\Theta}\) touches\(^1\) the curve \(\gamma\) at a point \(p \in \gamma\) is not penetrated by the curve, that is, \(B_{\Theta} \cap \gamma = \emptyset\). In fact, otherwise there would be a point \(q \in B_{\Theta} \cap \gamma\), and the plane spanned by the segment \(q - p\) and the tangent vector of \(\gamma\) at \(p\) would intersect \(B_{\Theta}\) in a planar disk of radius at most \(\Theta\). This disk would contain the strictly smaller circle through \(q\) and \(p\) that is tangent to the disk’s boundary at \(p\). Approximating this circle by the circumcircles of the point triples \(q, p, p_i\) for some sequence \(\{p_i\}\) of curve points converging to \(p\) as \(i\) tends to infinity yields a contradiction via \(\Delta[\gamma] \leq R(p, q, p_i) \leq \Theta\) for sufficiently large \(i\).

A direct consequence of the presence of forbidden balls is that Problem (P) is not solvable at all if the prescribed thickness is strictly greater than 1; there are simply no spherical curves whose thickness exceeds the value 1. Indeed, for any point \(p\) on

\(^1\)Throughout the paper we say that a submanifold \(M \subset \mathbb{R}^3\) touches another submanifold \(N \subset \mathbb{R}^3\) if their topological boundaries intersect non-transversally. In the present case it means that \(\gamma\) is tangent to \(\partial B_{\Theta}\) at \(p\).
a spherical curve $\gamma$ with thickness $\Delta[\gamma] > 1$ there exists an open ball of radius $\Delta[\gamma]$ touching the unit sphere (and therefore $\gamma$ as well) at $p$ and containing all of the unit sphere but $p$. However, this ball is forbidden and hence contains no curve point, so that $p$ is the only curve point on $S^2$. This settles Problem (P) for $\Theta > 1$.

If we intersect the union of all forbidden touching balls $B_\Theta$, $\Theta \leq \Delta[\gamma]$, for a loop $\gamma \subset S^2$ with the unit sphere, we easily deduce (see Figure 3) that every curve point of a spherical curve carries a pair of geodesic balls $B_\vartheta(\xi) := \{ \eta \in S^2 : \text{dist}_{S^2}(\eta, \xi) < \vartheta \}$ of radius $\vartheta := \arcsin \Theta$ on $S^2$ that do not intersect $\gamma$. Here $\text{dist}_{S^2}(\xi, \eta) := \angle(\xi, \eta)$ for $\xi, \eta \in S^2$ denotes the intrinsic distance on $S^2$. Thus we have shown:

**Forbidden Geodesic Balls (FGB).** A closed spherical curve $\gamma : S^1 \rightarrow S^2$ with thickness $\Delta[\gamma] \geq \Theta > 0$ does not intersect any open geodesic ball $B_\vartheta(\xi)$ on $S^2$ whose boundary $\partial B_\vartheta(\xi)$ is tangent to $\gamma$ in at least one curve point.

One can imagine a bow tie consisting of two open geodesic balls of spherical radius $\vartheta$ attached to the curve at their common boundary point. This bow tie can be moved freely along the curve without ever hitting any part of the curve.

![Figure 3. Left: A great circle is the thickest curve on $S^2$. Right: The unit sphere cut along a normal plane that is orthogonal to $\gamma$ at the point $p \in \gamma$. The grey spatial forbidden ball rotated along the dashed circle generates a forbidden geodesic ball of radius $\vartheta = \arcsin \Theta$ on $S^2$.](image)

The full strength of Property (FGB) is frequently used later on to completely classify infinitely many explicit solutions of Problem (P). For the moment it helps us to quickly solve that problem for $\Theta = 1$. Take any point $p$ on an arbitrary spherical curve $\gamma$ with thickness 1. The two forbidden open geodesic balls of spherical radius $\vartheta = \arcsin 1 = \frac{\pi}{2}$ touching $\gamma$ at $p$ are two complementary hemispheres $S^+$ and $S^-$ that—according to (FGB)—do not intersect $\gamma$. Hence $\gamma$ must be the equator as the only closed curve contained in the complement $S^2 \setminus (S^+ \cup S^-)$. Thus the equator is the only spherical curve with thickness 1 and hence—up to congruence—the unique solution to Problem (P) for $\Theta = 1$.

But what about other thickness values $\Theta \in (0, 1)$? Is the variational problem (P) solvable at all? The answer is yes, and once one has analyzed the continuity properties of the constraint $\Delta[\gamma] \geq \Theta$, this can be proven with the direct method in the calculus of variations. The necessary arguments for this (and for the constructions and classification results in Sections 2 and 3) are carried out in full detail in [9].

**Theorem 1.1 (Existence [9, Theorem 1.1]).** For each prescribed minimal thickness $\Theta \in (0, 1]$ Problem (P) possesses (at least) one solution $\gamma_\Theta$. In addition, every such solution attains the minimal thickness, i.e., $\Delta[\gamma_\Theta] = \Theta$. 
2. INFINITELY MANY EXPLICIT SOLUTIONS. Knowing that solutions exist does not necessarily mean that we know their actual shapes, unless $\Theta = 1$ where we have identified the equator as the only solution. For general variational problems it is mostly impossible to extract explicit information about the shapes of solutions; even uniqueness is usually a challenging issue. Here, however, the situation is different, and this has to do with the fact that every spherical curve $\gamma$ with positive thickness $\Delta[\gamma] \geq \Theta > 0$ carries an open spherical tubular neighborhood of radius $\vartheta := \arcsin \Theta$.

$$\mathcal{T}_\vartheta(\gamma) := \{\xi \in S^2 : \text{dist}_{S^2}(\gamma, \xi) < \vartheta\},$$

which equals the disjoint union of open subarcs of great circles of uniform length $2\vartheta$ on the sphere. Each of these great arcs is centered at a curve point $p \in \gamma$, and is orthogonal to the tangent vector of $\gamma$ at $p$. In accordance with the spatial case, we say that such a spherical tubular neighborhood has no self-penetration.

The converse is also true: the existence of such a spherical tubular neighborhood $\mathcal{T}_\vartheta(\gamma)$ around a curve $\gamma \subset S^2$ implies that this curve has thickness $\Delta[\gamma] \geq \sin \vartheta$. To see this correspondence, it is enough to show that the normal disks constituting the spatial tubular neighborhood $B_{\vartheta}(\gamma)$ are disjoint if and only if the great arcs of length $2\vartheta$ whose union equals $\mathcal{T}_\vartheta(\gamma) \subset S^2$ are disjoint. This can easily be verified using the central projection from the origin (see Figure 4). We define the spherical thickness of a curve $\gamma \subset S^2$ as the supremum of all radii $\tau$ such that the spherical tubular neighborhood $\mathcal{T}_\tau(\gamma)$ has no self-penetration. Thus we conclude that a spherical curve $\gamma$ has (spatial) thickness $\Theta := \Delta[\gamma]$ if and only if it has spherical thickness $\vartheta = \arcsin \Theta$.

![Figure 4.](image.png)

Figure 4. Orthogonal cross section of the spatial tubular neighborhood $B_{\vartheta}(\gamma)$ (grey region). Illuminated by a point-like light source at the origin, each such cross section casts a shadow on the surface of the sphere in the form of a great arc of length $2\vartheta$ centered at $p$ and normal to $\gamma$ (black arc). If two different grey disks intersect then the corresponding black great arcs must intersect as well. If, on the other hand, two such black great arcs intersect, then consider the two dashed arcs, each of which is in 1-1 correspondence to the respective great arc under the central projection from the origin. Those two dashed arcs must intersect as well, and therefore also the respective grey disks that contain them.

It was shown more than 70 years ago by Hotelling [14] and in more generality by Weyl [24] that the volume of such a uniform tubular neighborhood is proportional to the length $\mathcal{L}$ of its centerline. Adapted to the present situation of thick curves on the unit sphere, this classic theorem reads

$$\text{area}(\mathcal{T}_\vartheta(\gamma)) = 2 \sin \vartheta \cdot \mathcal{L}(\gamma)$$
for any curve $\gamma \subset S^2$ with (spatial) thickness $\triangle[\gamma] \geq \Theta = \sin \vartheta$. Consequently, any curve $\gamma \subset S^2$ with thickness $\triangle[\gamma] \geq \Theta$ whose spherical tubular neighborhood $\mathcal{T}_\Theta(\gamma)$ covers all of $S^2$, i.e., with
\[
\text{area}(\mathcal{T}_\Theta(\gamma)) = 4\pi = \text{area}(S^2),
\]
has maximal length among all spherical curves with prescribed minimal thickness $\Theta$. In other words, sphere-filling thick curves provide solutions to Problem (P).

Are there any thickness values $\Theta \in (0,1)$ such that we find sphere-filling curves of that minimal thickness, i.e., curves $\gamma \subset S^2$ with $\triangle[\gamma] \geq \Theta$ such that for $\vartheta = \arcsin \Theta$ we have relation (2.2)?

If we relax for a moment our assumption that we are searching for one connected closed curve then we easily find sphere-filling ensembles of curves. For $\Theta_n := \sin(\pi/2n) =: \vartheta_n$, $n \in \mathbb{N}$, the stack of latitudinal circles $C_i$ with $\text{dist}_{S^2}(C_0, \text{north pole}) = \vartheta_n$ and mutual distance $\text{dist}_{S^2}(C_i, C_{i-1}) = 2\vartheta_n$ for $i = 1, \ldots, n-1$ forms a set of $n$ spherical curves, and this ensemble has spherical thickness $\vartheta_n = \pi/2n$. Their mutually disjoint spherical tubular neighborhoods completely cover the sphere:
\[
\text{area}\left[ \bigcup_{i=0}^{n-1} \mathcal{T}_{\vartheta_n}(C_i) \right] = 4\pi.
\]

This collection of latitudinal circles can now be used to construct one closed sphere-filling curve. Let us explain in detail how, for the case $n = 4$. We cut the sphere with the 4 latitudinal circles along a longitudinal circle into an eastern hemisphere $S^e$ and a western hemisphere $S^w$. Each hemisphere now contains a stack of 4 latitudinal semicircles. Keeping the western hemisphere $S^w$ fixed, we rotate the eastern hemisphere $S^e$ by an angle of $2\vartheta_4 = \pi/4$ so that all the endpoints of the now-turned semicircles on $S^e$ meet endpoints of the semicircles on $S^w$; see Figure 5.

![Figure 5](image)

This modified collection of semicircles still has spherical thickness $\vartheta_4 = \pi/8$ and is sphere filling, since in the construction the sphere-filling stack of the original latitudinal circles was only cut orthogonally and reunited along one longitudinal circle, which does not change the thickness and sphere-filling property of the ensemble. We also observe that this new ensemble, which resembles to some extent the seam lines on a tennis ball, forms one closed curve, and hence solves our problem—at least for this particular given spatial thickness $\Theta_4 = \sin \vartheta_4 = \sin(\pi/8)$. Are there other solutions for $n = 4$? Why not continue rotating the eastern hemisphere $S^e$ against the fixed hemisphere $S^w$ to obtain more solutions? It turns out that a total rotation by $4\vartheta_4 = \pi/2$ yields two connected components, which is not what we are looking for. But turning $S^e$ by an angle of $6\vartheta_4 = 3\pi/4$ leads to another solution: a new single closed loop which is the mirror image of the first one; see Figure 5.
One can show that this procedure works well for arbitrary \( n \in \mathbb{N} \), and with a little elementary algebra\(^2\) we can determine the exact number of solutions:

**Theorem 2.1 (Explicit solutions).** For each \( n \in \mathbb{N} \) and each \( k \in \{0, \ldots, n - 1\} \) whose greatest common divisor with \( n \) equals 1, the construction described above starting with \( n \) latitudinal circles \( C_0, \ldots, C_{n-1} \) with spherical distance

\[
\text{dist}_{S^2}(C_0, \text{north pole}) = \vartheta_n = \frac{\pi}{2n}, \quad \text{dist}_{S^2}(C_i, C_{i-1}) = 2\vartheta_n, \quad i = 1, \ldots, n - 1,
\]

and rotating the eastern hemisphere against the fixed western hemisphere by an angle of \( 2k\vartheta_n \) leads to \( \varphi(n) \) explicit piecewise circular solutions of the variational problem \((P)\) for prescribed minimal thickness \( \Theta_n = \sin \vartheta_n \).

Here, \( \varphi \) denotes the Eulerian totient function from number theory: \( \varphi(n) \) gives the number of positive integers \( k \leq n \) so that the greatest common divisor of \( k \) and \( n \) equals 1. In our example above, \( n = 4 \), we indeed found \( \varphi(4) = 2 \) explicit solutions by rotating the eastern hemisphere by the amount of \( 2k\vartheta_4 \) for \( k = 1 \) and for \( k = 3 \).

Figure 6 depicts such sphere-filling closed curves for various \( n \), and one notices a striking resemblance with certain so-called Turing patterns (see Figure 10(a)) observed and analyzed in chemistry and biology as characteristic concentration distributions of different substances; see, e.g., [23]. In that context, the patterns are caused by diffusion-driven instabilities; here, in contrast, the shapes of solutions are a consequence of a simple variational principle.

\[\text{Figure 6. Various solutions of Problem (P). All curves are visualized as tubes of a fixed radius } \Theta = \pi/24, \text{ which coincides with the actual spatial thickness } \Delta[\gamma] = \Theta_n \text{ only for the last curve with } n = 12. \text{ The remaining values of spatial thickness } \Theta_n, \text{ for } n = 2, 3, 5, 7, 11, \text{ all exceed the tubes’ radii depicted in the image.}\]

\(^2\)Such a construction was used for a bead puzzle called the orb or orb it [25] in the 1980s, and the algebra involved was probably known to its inventors.
Similar constructions for thickness values $\Omega_n := \sin(\pi/n)$, $n \in \mathbb{N}$, starting from an initial ensemble of semicircles together with one or two poles on $S^2$, lead to two disjoint families of sphere-filling open curves\(^3\) distinguished by the relative positions of the two endpoints on the sphere; see Figure 7. For all even $n \in \mathbb{N}$ the resulting sphere-filling open curves have antipodal endpoints, which is not the case if $n$ is odd. Let us point out that these open curves occur in the different context of statistical physics, namely as two of three possible configurations of ordered phases of long elastic rods densely stuffed into spherical shells; see [15], in particular their Figures 4a and 4c. Those studies aimed at explaining the possible nematic order of densely packed long DNA in viral capsids.

Figure 7. Various solutions of a version of Problem (P) for open curves. Only the last ones in each row are depicted with full spatial thickness.

3. CLASSIFICATION OF SPHERE-FILLING ROPES. For each positive integer $n$ we have constructed explicitly longest closed ropes of thickness $\Theta_n = \sin(\pi/2n)$ on the unit sphere. Are there more? We know there are for intermediate values $\Theta \neq \Theta_n$ by Theorem 1.1, but even if we stick to these specific countably many values $\Theta_n$ of given minimal thickness might we find more sphere-filling and thus length-maximizing curves of considerably different shapes? The answer may be surprising, but, no, up to rotations our solutions are the only ones, and this “uniqueness” result is actually a consequence of a complete classification of sphere-filling thick curves:

Theorem 3.1 (Classification of sphere-filling loops). If the spherical tubular neighborhood $T_\vartheta(\gamma)$, $\vartheta \in (0, \pi/2]$, of a closed spherical curve $\gamma \subset S^2$ with thickness

\[ n = 3, k = 1 \quad n = 5, k = 1 \quad n = 25, k = 6 \]

\[ n = 4, k = 0 \quad n = 6, k = 0 \quad n = 26, k = 0 \]

\[ \text{Figure 7. Various solutions of a version of Problem (P) for open curves. Only the last ones in each row are depicted with full spatial thickness.} \]

\[^3\text{The (spatial) thickness of an open curve } \gamma \text{ with two endpoints } p \text{ and } q \text{ is the supremum of all radii } r > 0 \text{ such that the tubular neighborhood } B_r(\gamma) \text{ has no self-penetration, where } B_r(\gamma) \text{ consists of three disjoint parts: an open half-ball cap at } p \text{ and at } q, \text{ and the disjoint union of open normal disks centered on the curve.} \]
\[\Delta[\gamma] \geq \Theta = \sin \vartheta \text{ satisfies} \]
\[
\text{area } (\mathcal{T}_\Theta(\gamma)) = 4\pi = \text{area}(S^2),
\]
then there exists a positive integer \(n\), and an integer \(k \in \{0, \ldots, n - 1\}\), with greatest common divisor equal to 1, such that \(\vartheta = \vartheta_n = \pi/(2n)\), \(\Delta[\gamma] = \Theta_n = \sin \vartheta_n\), and such that \(\gamma\) coincides—up to a rotation—with one of the \(\varphi(n)\) explicit solutions of Problem (\(P\)) exhibited in Theorem 2.1.

An analogous result holds also for open curves: any sphere-filling thick open curve must have spherical thickness \(\omega_n = \arcsin \Omega_n = \pi/n\) for some \(n \in \mathbb{N}\), and coincides with a member of one of the two explicitly constructed families of open spherical curves, depending on whether \(n\) is even or odd. So, if one was given the (somewhat strange) task to produce a soccer ball of a given size by deforming a continuous piece of thick rope of suitable length into an airtight spherical hull, then only specific values of rope thickness are possible, and our theorem tells us how one should proceed. There is simply no other way!

Let us explain the main ideas of the proof of this classification result. The presence of forbidden geodesic balls (FGB) allows us to prove a fundamental touching principle for spherical curves \(\gamma\) with positive thickness \(\Delta[\gamma]\); see Part A below. This principle guarantees then that the number of possible local touching situations between the curve and geodesic balls with radius equal to \(\Delta[\gamma]\) is very limited (Part B). The combination of these pieces of information leads to a geometric rigidity for sphere-filling curves reflected in two sorts of possible global patterns (Part C).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig8.png}
\caption{A curve \(\gamma\) with spherical thickness \(\vartheta\) that touches a geodesic circle of radius \(\vartheta\) in two non-antipodal points \(P\) and \(Q\) joins them with a subarc of that circle. Otherwise one of the dotted geodesic circles would intersect \(\gamma\) three times, leading to a lower thickness.}
\end{figure}

\textbf{A. The touching principle} addresses the situation when a spherical curve \(\gamma\) with \(\Delta[\gamma] \geq \Theta = \sin \vartheta\) touches the boundary \(\partial \mathscr{B}_\Theta(\xi_0)\) of a geodesic ball \(\mathscr{B}_\Theta(\xi_0)\) in \(S^2\) in at least two non-antipodal points \(P, Q \in \partial \mathscr{B}_\Theta(\xi_0)\) with \(\text{dist}_{S^2}(P, Q) =: 2\vartheta_1 < 2\vartheta\). In this situation the boundary of the strictly smaller geodesic ball \(\mathscr{B}_\vartheta_1(\xi_1)\) for which \(P\) and \(Q\) are antipodal is intersected transversally by \(\gamma\) at \(P\) and \(Q\), which means that the open geodesic ball \(\mathscr{B}_\vartheta_1(\xi_1)\) contains curve points; see Figure 8. On the other hand, \(\partial \mathscr{B}_\vartheta_1(\xi_1)\) contains no further curve point \(T\) different from \(P\) and \(Q\), since this would imply for the corresponding (Euclidean) circumcircle that
\[
R(P, Q, T) \leq \sin \vartheta_1 < \sin \vartheta = \Theta,
\]
contradicting our assumption \(\Delta[\gamma] \geq \Theta\); recall formula (1.1). Consequently, there is a whole subarc of \(\gamma\) connecting \(P\) and \(Q\) contained in \(\mathscr{B}_\vartheta_1(\xi_1)\), but not in the original
larger ball $B_\vartheta(\xi_0)$ since this one is a forbidden ball according to (FGB). Where can we locate this arc within the set $B_{\vartheta_1}(\xi_1) \setminus B_{\vartheta}(\xi_0)$? Sweeping out the region $B_{\vartheta_1}(\xi_1) \setminus B_\vartheta(\xi_0)$ with intermediate geodesic circles $\partial B_\vartheta(\xi_s)$ with center $\xi_s$ on the great arc connecting $\xi_0$ and $\xi_1$ and containing $P$ and $Q$ for each $s \in [0, 1]$ (so that $\vartheta_s = |P - \xi_s|$, $\vartheta_0 := \vartheta$), we use the same argument as the one that led to (3.3) to show that there are no curve points in $B_{\vartheta_1}(\xi_1) \setminus B_\vartheta(\xi_0)$. Thus we have proven:

**Touching Principle (TP).** A closed spherical curve $\gamma : S^1 \to S^2$ with (spatial) thickness $\Delta[\gamma] \geq \Theta = \sin \vartheta$ that touches a geodesic circle $\partial B_\vartheta \subset S^2$ of spherical radius $\vartheta$ at two non-antipodal points $P$ and $Q$ contains the shorter circular subarc of $\partial B_\vartheta$ connecting $P$ and $Q$.

We benefit from the touching principle since it allows us to characterize sphere-filling curves of thickness $\Theta = \Delta[\gamma] = \sin \vartheta$ in terms of their local behavior when touching geodesic balls of spherical radius $\vartheta$: for any open geodesic ball $B_\vartheta$ disjoint from $\gamma$ — and there are plenty of those, e.g., all forbidden balls by (FGB) — one of the following three touching situations is guaranteed for the intersection $\mathcal{I} := \partial B_\vartheta \cap \gamma$:

**B. Possible local touching situations.**

(a) $\gamma$ touches $\partial B_\vartheta$ in exactly two antipodal points, i.e., $\mathcal{I} = \{P, Q\}$ with $\text{dist}_{S^2}(P, Q) = 2\vartheta$, or

(b) the intersection $\mathcal{I}$ is a relatively closed semicircle of spherical radius $\vartheta$, or

(c) this intersection $\mathcal{I}$ equals the full geodesic circle $\partial B_\vartheta$.

To see this we notice first that for a sphere-filling curve the relatively closed intersection $\mathcal{I}$ is nonempty, since otherwise a slightly larger ball $B_{\vartheta_+}$ for some small positive $\epsilon$ would not contain any curve point, which leads to $\mathcal{I} \cap T_0(\gamma) = \emptyset$, and hence $4\pi = \text{area}(T_0(\gamma)) \leq \text{area}(S^2 \setminus B_\vartheta) < 4\pi$ contradicting (2.2).

Similarly, one can rule out that the set $\mathcal{I}$ is contained in a relatively open semicircle on $\partial B_\vartheta$, since then two extremal points $\eta, \zeta \in \mathcal{I}$ realizing the diameter of $\mathcal{I}$ would have spherical distance $\text{dist}_{S^2}(\eta, \zeta) < 2\vartheta$. This is one of the frequent occasions that the touching principle (TP) comes into play. It implies that the whole circular subarc of $\partial B_\vartheta$ connecting $\eta$ and $\zeta$ is contained in $\mathcal{I}$ and therefore equals $\mathcal{I}$ in this situation. But the fact that $\gamma$ lifts off the geodesic circle $\partial B_\vartheta$ at $\eta$ and $\zeta$ before completing a full semicircle allows us to move the closed ball $B_\vartheta$ slightly “away” from $\mathcal{I}$, that is, in the direction orthogonal to and away from the geodesic arc connecting $\eta$ and $\zeta$. This way we obtain a slightly shifted closed ball of the same radius without any contact to $\gamma$, a situation that we have ruled out above. Therefore, $\mathcal{I}$ is not contained in any relatively open semicircle on $\partial B_\vartheta$.

If $\mathcal{I}$ is contained in a relatively closed semicircle we may assume that it contains apart from the antipodal endpoints of that semicircle also at least one third point, since otherwise we would be in situation (a) and could stop here. Therefore, by virtue of the touching principle (TP) $\mathcal{I}$ coincides completely with that closed semicircle, which is option (b). If, however, $\mathcal{I}$ is not contained in any semicircle we can simply look at one point $q \in \mathcal{I}$ and its antipodal point $q' \in \partial B_\vartheta$. If $q'$ happens to be also in $\gamma$, then both semicircles connecting $q$ and $q'$ would contain further curve points and therefore $\mathcal{I} = \partial B_\vartheta$ again by the touching principle, and we end up with option (c).

If $q' \notin \mathcal{I}$, then the largest open subarc $\alpha$ of $\partial B_\vartheta$ containing $q'$ but no point of $\gamma$ must be shorter than $\pi$ unless $\mathcal{I}$ is contained in a semicircle, a situation we brought to a close before. Applying the touching principle to the two endpoints of $\alpha$ we find in
fact that $q'$ lies on the subarc of $\gamma$ connecting these endpoints on $\partial \mathcal{B}_\vartheta$, which exhausts the last possible situation to verify that our list of situations (a)--(c) is complete.

We are going to use the local structure established in Parts A and B to prove geometric rigidity of sphere-filling curves $\gamma$ with positive spatial thickness $\Theta = \sin \vartheta$.

C. Global patterns of sphere-filling curves.

(C1). If $\gamma \subset S^2$ intersects a normal plane $E$ orthogonally at $k$ distinct points whose mutual spherical distance equals $2\vartheta$, then $k$ is even and $\gamma$ contains a semicircle of spherical radius $\vartheta$ in each of the two hemispheres bounded by $E \cap S^2$.

(C2). If $\gamma$ contains one latitudinal semicircle $S \subset \partial \mathcal{B}_\vartheta$, then $\vartheta = \pi/2n$ for some $n \in \mathbb{N}$ and the portion of $\gamma$ in the corresponding hemisphere consists of the whole stack of $n$ latitudinal semicircles (including $S$) with mutual spherical distance $\vartheta$.

Before providing the proofs for these rigidity results let us explain how we can combine these to prove Theorem 3.1.

Proof of Theorem 3.1. The goal is to show the existence of a latitudinal semicircle of spherical radius $\vartheta$ contained in $\gamma$ in order to apply the global pattern (C2), which then assures that $\gamma$ consists of a stack of latitudinal semicircles with mutual spherical distance $\vartheta$ in one hemisphere, say in $S^w$. This behavior in $S^w$ leads to the characteristic intersection point pattern in the longitudinal circle $\partial S^w$ needed in order to apply (C1), which in turn guarantees the existence of a semicircle of radius $\vartheta$ on $S^c$ whose endpoints a priori do not need to lie on $\partial S^c$. Again property (C2) applied to this semicircle on $S^c$ leads to a whole stack of $n$ equidistant semicircles now at least partially contained in $S^c$. The only way both stacks of semicircles fit together to form one closed spherical curve of prescribed spherical thickness $\vartheta$ is that the second stack is completely contained in $S^c$. Our construction described in Section 2 finally reveals the only possible loops made of two such stacks of equidistant latitudinal semicircles meeting $\partial S^c = \partial S^w$ orthogonally, which completes the proof of the classification theorem. The logic of the proof resembles a ride on a merry-go-round: (C1) produces the semicircle on $S^w$ necessary to use (C2) to obtain the stack of semicircles on $S^w$, which itself generates the point pattern needed to apply (C1) on $S^c$ and finish the task via (C2) on $S^c$. The only problem is: how do we enter the merry-go-round? We have to show that one portion of $\gamma$ is a semicircle of spherical radius $\vartheta$ without assuming the intersection point pattern needed in (C1).

Let $k$ be the integer such that $(k - 1)\vartheta < \pi \leq k\vartheta$. For a fixed point $p \in \gamma$ we walk along a unit-speed geodesic ray $\eta_p$ emanating from $p$ in a direction orthogonal to $\gamma$ at $p$ in search of such a semicircle. The geodesic ball $\mathcal{B}_\vartheta(\eta_p(\vartheta))$ is a forbidden ball by means of (FGB), i.e., $\gamma \cap \mathcal{B}_\vartheta(\eta_p(\vartheta)) = \emptyset$, where $\eta_p(\vartheta)$ denotes the point reached on the geodesic ray after a spherical distance $\vartheta$. According to the possible local touching situations we find the desired semicircle on $\partial \mathcal{B}_\vartheta(\eta_p(\vartheta)) \cap \gamma$ (option (b) or (c) in B), unless the antipodal point $\eta_p(2\vartheta)$ is contained in $\gamma$. In that case we continue along the same geodesic ray passing through $\eta_p(2\vartheta)$ orthogonally to $\gamma$, until we either find a closed semicircle on one of the geodesic circles $\partial \mathcal{B}_\vartheta(\eta_p(2i(1 - 1)\vartheta))$, $i = 1, \ldots, k$, or $\vartheta = \pi/k$, and all “antipodal points” $\eta_p(2i\vartheta)$ are contained in $\gamma$, so that $\eta_p(2k\vartheta) = p$. In other words, either we have found the desired semicircle during the walk along $\eta_p$, or we have walked once around the whole longitudinal circle traced out by $\eta_p$ generating $k$ equidistant points where $\gamma$ intersects $\eta_p$ orthogonally. But this is exactly what is needed to apply (C1) to finally establish the existence of the semicircle we are looking for.
One final comment on why this exact quantization takes place, i.e., why we find \( \vartheta = \pi / k \) so that the walk along \( \eta_p \), pinpointing the centers \( \eta_p((2i-1)\vartheta) \) of geodesic balls on the way, actually leads exactly back to the starting point \( p = \eta_p(2\pi) \). The successive localization of forbidden balls according to (FGB) and the possible local touching situations yield the fact that all open geodesic balls \( \mathcal{B}_0(\eta_p((2i-1)\vartheta)) \) for \( i = 1, \ldots, k \) are disjoint from \( \gamma \). If, for instance, the walk had stopped too late since the step size \( \vartheta \) was too large, \( k\vartheta > \pi \), then \( \text{dist}_{\mathbb{S}^2}(p, \eta_p((2k-1)\vartheta)) = 2\pi - (2k-1)\vartheta < \vartheta \), that is, \( p \in \mathcal{B}_0(\eta_p((2k-1)\vartheta)) \cap \gamma \), a contradiction. \( \square \)

Let us establish the global patterns (C1) and (C2) in more detail since they served as the core tools in the proof of our classification theorem.

We start with the proof of (C1). Here it suffices to focus on one of the two hemispheres \( \mathbb{S}^w \) and \( \mathbb{S}^c \) determined by \( E \), say on \( \mathbb{S}^w \). Since \( \gamma \) is simple and closed the curve can leave \( \mathbb{S}^w \) merely as often as it enters \( \mathbb{S}^w \), which immediately gives \( k = 2n \) for some \( n \in \mathbb{N} \). Moreover, \( \mathbb{S}^w \) is homeomorphic to a flat disk so that we can find nearest neighboring exit and entrance points \( p, q \in E \cap \gamma \) with minimal spherical distance \( \text{dist}_{\mathbb{S}^2}(p, q) = 2\vartheta \) such that the closed subarc \( \beta \subset \mathbb{S}^w \cap \gamma \) connecting \( p \) and \( q \) satisfies \( E \cap \beta = \{p, q\} \). We will show that \( \beta \) contains the desired semicircle of spherical radius \( \vartheta \). Since \( \gamma \) intersects \( E \) orthogonally we infer from (FGB) that the open geodesic ball \( \mathcal{B} = \mathcal{B}_\vartheta \) with \( p, q \in \partial \mathcal{B} \) as antipodal boundary points is disjoint from \( \gamma \). If there were a third point \( b \in \beta \cap \partial \mathcal{B} \) distinct from \( p \) and \( q \), then—according to the touching principle (TP)—the whole semicircle on \( \partial \mathcal{B} \) with endpoints \( p \) and \( q \) would be contained in \( \gamma \), and we would be done. Otherwise we trace the open spherical region \( \mathcal{B} \) bounded by \( \beta \cup (\mathbb{S}^c \cap \partial \mathcal{B}) \) with geodesic rays \( \eta_b \) emanating from arbitrary points \( b \in \beta \) orthogonally into the region \( \mathcal{B} \). Notice that \( \mathcal{B} \) is disjoint from \( \gamma \), and that \( \eta_p(2\vartheta) = q \) and \( \eta_q(2\vartheta) = p \), where the argument of \( \eta \) indicates how long one has to travel along the geodesic ray to reach the destination point. In addition, the forbidden ball property (FGB) implies \( \gamma \cap \mathcal{B}_\vartheta(\eta_b(\vartheta)) = \emptyset \) and therefore \( \mathcal{B}_\vartheta(\eta_b(\vartheta)) \subset \mathcal{B} \) for all points \( b \in \beta \); see Figure 9.

![Figure 9](image_url)

**Figure 9.** Towards the proof of (C1). Geodesic rays \( \eta_b \) emanating from points \( b \) on the subarc \( \beta \subset \gamma \cap \mathbb{S}^w \) to trace the enclosed spherical region \( \mathcal{B} \). The depicted antipodal touching (3.4) cannot hold throughout \( \beta \) by virtue of Brouwer’s fixed point theorem.

According to Part B either

\[
\partial \mathcal{B}_\vartheta(\eta_b(\vartheta)) \cap \gamma = \{b, \eta_b(2\vartheta)\} \quad \text{for all } b \in \beta
\]

(3.4)

(the antipodal situation (a)), or \( \gamma \) contains a semicircle \( S_b = \partial \mathcal{B}_\vartheta(\eta_b(\vartheta)) \cap \gamma \) containing itself the point \( b \) for some \( b \in \beta \). This semicircle lies completely in the western
hemisphere $S^w$, which concludes the proof of (C1). To see the latter assume contrari-
wise that there exists a point $z$ on $S^w \cap S_b \setminus E$. Then by connectivity $p$ or $q$ would lie on
the semicircle $S_b$, too, which immediately implies that $S_b$ and hence also $b \in \partial B \cap \beta$
lies on the original geodesic circle $\partial B$, a situation that we had excluded already.

It remains to exclude the antipodal touching (3.4) throughout the subarc $B \subset \gamma \cap
S^w$. We use Brouwer’s fixed point theorem for the continuous map $f : \beta \to S^w$ defined by
$f(b) := \eta_b(2\vartheta)$, which we claim actually maps $\beta$ into $\beta$. Relation (3.4) in fact
 guarantees that $f(b) \in \gamma \setminus B$. Hence $f(b)$ is either contained in $S^w \setminus B \cap \gamma = \beta$
in which case we are done, or $f(b)$ lies in $\partial B \cap S^w \cap \gamma$. But then the antipodal partner $b$
of $f(b)$ would also lie on $\partial B$, which was excluded earlier. Consequently, Brouwer’s
theorem is applicable and leads to a fixed point $b^* = f(b^*) = \eta_{b^*}(2\vartheta)$, which implies
$2\vartheta = 2\pi$ because $\eta_{b^*}$ parametrizes a unit speed great circle on $S^2$. But this contradicts
our assumption that $\vartheta \in (0, \pi/2)$.

The proof of (C2) can be sketched as follows. Let $n$ be the integer such that
$\pi/(2n) \leq \vartheta < \pi/(2n - 2)$. According to (FGB) one has a forbidden ball: $\gamma \cap B_\vartheta = \emptyset$, and the idea is to start from the initial semicircle $S_1 := S \subset \partial B_\vartheta$ and “scan” the
remaining part $S^w \setminus B_\vartheta$ with unit speed geodesic rays $\eta_{p,\vartheta}$ emanating from every point
$p \in S_1$ orthogonally to $S_1$, into the open region $S^w \setminus B_\vartheta$. Again by (FGB) we find
$B_\vartheta(\eta_p(\vartheta))) \cap \gamma = \emptyset$ for each such starting point $p \in S_1$. All possible touching situations
documented in Part B guarantee the existence of at least one second curve point $q$ on
$\partial B_\vartheta(\eta_p(\vartheta )))$, and we claim that $q$ must be antipodal to $p$, i.e., $q = \eta_p(2\vartheta)$. If
not, then according to option (b) or (c) in Part B, the points $p$ and $q$ are contained in
a semicircle on $\partial B_\vartheta(\eta_p(\vartheta ))) \cap \gamma$. But this semicircle is hit by neighboring geodesic
“scanning” rays $\eta_p, \eta_p$ emanating from $r \in S_1$ for $r$ close to $p$, which would lead to a
nonempty intersection of $\gamma$ with the neighboring forbidden ball $B_\vartheta(\eta_p(\vartheta )))$ contradicting
(FGB). Hence we have shown that only antipodal curve points can be generated by
this procedure: $\partial B_\vartheta(\eta_p(\vartheta ))) \cap \gamma = \{ p, \eta_p(2\vartheta) \}$ for all $p \in S_1$, which produces a
second semicircle $S_2 := \{ \eta_p(2\vartheta) : p \in S_1 \} \subset \partial B_{3\vartheta}$ contained in $\gamma$, with spherical
distance $2\vartheta$ to the first semicircle $S_1$.

It is obvious how to continue this procedure—now starting the “scanning” rays
from $S_2$—to obtain a whole stack of semicircles $S_i = \partial B_{(2i-1)\vartheta}$ for $i = 1, \ldots, n$. If
this stack is too high because the spherical thickness is too large with respect to $n$,
i.e., if $\pi/(2n - 1) \leq \vartheta < \pi/(2n - 2)$, then the stack would “spill over” onto the next
hemisphere $S'$ producing a final semicircle $S_n$ contained in $S' \cap \gamma$ with spherical radius
$(2n - 1)\vartheta - \pi \in [0, \vartheta)$ that is too small: it contradicts the spherical thickness
$\vartheta = \arcsin \Delta[\gamma]$ of $\gamma$ since its curvature is too large. If the stack is not high enough
$(\pi/2n < \vartheta < \pi/(2n - 1))$ then the last semicircle $S_n$ is still on the correct hemisphere
$S^w$ but has spherical radius $\pi - (2n - 1)\vartheta \in (0, \vartheta)$, which is again too small for the
thick curve $\gamma$.

4. FINAL REMARKS AND OPEN PROBLEMS. For a countably infinite number of
thickness values we have established a complete picture of the solution set for Problem
(P) using the sphere-filling property to a large extent. The general existence theorem,
Theorem 1.1, however, also guarantees the existence of longest ropes on the unit
sphere for all intermediate thickness values $\Theta \neq \Theta_n$. What are their actual shapes?

Theorem 3.1 ascertains that those solutions cannot be sphere-filling. In [9] we
constructed a comparison curve that could serve as a promising candidate for prescribed
minimal thickness $\Theta \in (\Theta_2, \Theta_1)$, but this question remains to be investigated, as well
as the interesting connections to Turing patterns and the statistical behavior of long
elastic rods under spherical confinement mentioned in Sections 1 and 2. In addition,
if one replaces the unit sphere with other supporting manifolds such as the standard torus or higher-dimensional spheres or even $\mathbb{R}^3$, then the issue of analyzing the shapes of optimally packed ropes is wide open; see for example the interesting conjectures in [16].

Thickness as defined in (1.1) can be attributed to curves of low regularity, which turned out to be useful to prove general existence results for thick elastic rods and ideal knots and links [12, 3, 10]. An ideal knot or link, in particular, is a minimizer of ropelength, the quotient of length and thickness, within its isotopy class. The only analytically known ideal knot, however, is the circle. Cantarella, Kusner, and Sullivan [3] have identified several families of ideal links (see Figure 10(c)), whereas the shape of the ideal trefoil depicted in Figure 10(b) is only a numerical approximation. In general questions regarding shape, uniqueness, and even regularity of ideal configurations are open.

REFERENCES


HENRYK GERLACH received his Ph.D. in 2010 from the École Polytechnique Fédérale de Lausanne under the guidance of John H. Maddocks and Peter Buser and was supported by the Swiss National Science Foundation.

Institut de Mathematiques B, Laboratory for Computation and Visualization in Mathematics and Mechanics, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
henryk.gerlach@gmail.com

HEIKO VON DER MOSEL completed his Ph.D. under the supervision of Stefan Hildebrandt at the University of Bonn in 1996. He is a professor of Mathematics at RWTH Aachen University, and his research is devoted to the calculus of variations and geometric analysis.

Institut für Mathematik, RWTH Aachen University, D-52062 Aachen, Germany
heiko@instmath.rwth-aachen.de

876 © THE MATHEMATICAL ASSOCIATION OF AMERICA [Monthly 118