ON DISTORTION AND THICKNESS OF KNOTS

ROBERT B. KUSNER AND JOHN M. SULLIVAN

Introduction

What length of rope (of given diameter) is required to tie a particular knot? Or, to turn the problem around, given an embedded curve, how thick a regular neighborhood of the curve also is embedded? Intuitively, the diameter of the possible rope is bounded by the distance between strands at the closest crossing in the knot. But of course the distance between two points along a curve goes to zero as the points approach each other, so to make the notion precise, we need to exclude some neighborhood of the diagonal.

Various notions of thickness have been proposed recently. For example, [LSDR] defines a thickness by considering the distance function between points on the curve only where it has critical points. But the definition there also involves the minimum radius of curvature of the curve, and thus is unbounded for polygonal curves. Here we introduce and compare two new families of thickness measures. One makes use of Gromov's concept of distortion (see [Gro1] and [Gro2, p. 114] and [GLP, pp. 6–9]); it applies to all rectifiable curves (including polygons). The other generalizes the notion from [LSDR]. Our main result is a basic inequality (Theorem 5.2) between these measures.

The distortion thickness should permit us to prove the existence of thickest curves of prescribed length (or dually, shortest curves of prescribed thickness) in each knot class; such curves are of interest to chemists and biologists modeling polymers and DNA (see, for example, [KBM$^+$]). Moreover, curvature bounds should follow from the optimality, and we conjecture that, for an optimal knot, all reasonable measures of thickness should be equal.

1. Notation and Definitions

We will deal throughout with an embedded rectifiable closed curve γ in $\mathbb{R}^n$, of finite length $L(\gamma)$. Thus we may assume that $\gamma\colon \mathbb{R}/L\mathbb{Z} \to \mathbb{R}^n$

Date: November 14, 1996; revised January 17, 1997.

The first author was funded in part by NSF grant DMS 94-04278, and both authors enjoyed the hospitality of the IMA during the Spring 1996 term.

is a Lipschitz parameterization by arclength. (Often we will normalize by rescaling so that $L = 2\pi$.) If $p, q \in \gamma$ are two points on the curve, $|p - q|$ will denote the straight-line (chord) distance between them in $\mathbb{R}^n$, and $d(p, q)$ the arclength distance between them along γ. This arclength distance is always measured the shorter way around γ, so that it is never more than $L/2$. Given any point p on γ, there is a unique *opposite point* p^* such that $d(p, p^*) = L/2$.

Definition. The *distortion* between distinct points p and q on the curve γ is

$$\delta(p, q) := \frac{d(p, q)}{|p - q|} \geq 1.$$

This is a Lipschitz constant for the inverse of the map γ (which itself has Lipschitz constant 1). The distortion of γ is the maximum distortion between any points:

$$\delta(\gamma) := \sup_{p,q} \delta(p, q),$$

where the supremum is over $(\gamma \times \gamma) \smallsetminus \Delta$, the set of all pairs of distinct points on γ. We can also consider a restricted distortion, only considering opposite pairs:

$$\delta_\circ(\gamma) := \sup_{p \in \gamma} \delta(p, p^*).$$

Clearly $\delta_\circ(\gamma) \leq \delta(\gamma)$ since the supremum is over a subset; also, the supremum in $\delta_\circ$ is achieved, though that for δ may not be.

Definition. Given any real number $b \geq 1$, the *distortion thickness* of a curve $\gamma \subset \mathbb{R}^n$ is

$$\tau_b(\gamma) := \inf_{\delta(p,q) > b} |p - q|,$$

the minimum distance between points of distortion more than b.

The idea here is that on a smooth curve, nearby points have distortion hardly greater than one, so by taking $b > 1$ we eliminate these nearby pairs before taking the minimum distance.

We will define the remaining notions only for $C^{1,1}$ curves, which have a well-defined, Lipschitz continuous tangent vector field, and a curvature κ almost everywhere. A pair of points (p, q) on the curve is called *symmetric* if the unit tangent vectors T_p and T_q make equal angles with the chord $p - q$, that is, $T_p \cdot (p - q) = T_q \cdot (q - p)$. For instance, any two points on a round circle are symmetric. An ordered pair (p, q) of distinct points is said to be *self-critical* [LSDR] if q is a critical point for the distance from p, that is $T_q \cdot (p - q) = 0$. A pair which is symmetric and self-critical is called *doubly self-critical.*

For any $C^{1,1}$ curve γ, define $c_1(\gamma)$ to be the minimum straight-line distance $|p-q|$ among self-critical pairs (p, q), and $c_2(\gamma)$ to be the minimum among doubly self-critical pairs. Also, let $r(\gamma)$ be the minimum radius of curvature along the curve, $r(\gamma) := 1/\sup_\gamma \kappa$; this is positive because a $C^{1,1}$ curve has bounded curvature.

Definition. Given a real number $k > 0$, the *curvature thickness* of a curve γ is

$$\sigma_k(\gamma) := \min\bigl(kr(\gamma), c_2(\gamma)\bigr).$$

For $k = 2$, this is (twice) the notion of thickness studied in [LSDR].

Note. If γ is not $C^{1,1}$, its curvature in some sense is unbounded, so we set $\sigma_k(\gamma) = r(\gamma) = 0$.

Since we are interested in minimizing the length of curves with (any one of our notions of) thickness prescribed, or dually, in maximizing thickness for curves of prescribed length, it useful to have a scale-invariant measure for this optimization problem.

Definition. The *ropelength* of a curve is its length divided by its thickness. This is the length of a similar curve scaled to have thickness 1 (so that it could be made out of a rope of diameter 1). Given a topological knot type K, the *ropelength of* K is the minimum ropelength of any curve γ of type K.

Intuitively, the ropelength of K is the shortest length of rope (of diameter one) which could be used to tie that knot. Of course, there are several notions of ropelength corresponding to our different notions of thickness, but we conjecture they all agree (for reasonable values of k and b), because a curve of shortest ropelength should have nice structure.

Note. Most of the notions we have defined make sense also for any union of immersed rectifiable curves, which need not be closed curves. For a link of several components, we take $d(p, q) = \infty$ if p and q are on different components, so that such pairs are always contenders in the definition of distortion thickness. If the curves are not embedded, there are points p, q such that $d(p, q) > 0$ while $|p - q| = 0$. Such a pair has infinite distortion, and is doubly self-critical, so all our notions agree that such a nonembedded curve has thickness zero.

2. Distortion of Curves

In order to understand the definition of distortion thickness τ_b, we need to explore the distortion of curves. In particular, we will see that every closed curve has distortion at least $\pi/2$, so that $b = \pi/2$ is a natural choice.

Recall that the distortion of γ is defined as a supremum of $\delta(p,q) = d(p,q)/|p-q| \geq 1$. We noted before that if γ is not rectifiable or embedded then we will have some $\delta(p,q) = \infty$. The finiteness of $\delta(\gamma)$ is sometimes called a *uniform chord-arc condition* in harmonic analysis [Tor]. We now note that this supremum may be infinite even for an embedded rectifiable curve.

Example. Consider first a wedge γ of external angle θ. If p and q are at equal distance s from the corner, then $|p-q| = 2s\cos\frac{\theta}{2}$ so $\delta(\gamma) = \delta(p,q) = \sec\frac{\theta}{2}$. Now take a plane curve γ, such as $x^2 = y^3$, with a sharp cusp at the origin. Then as p and q approach the origin along γ at equal distance s from (and on opposing sides of) the cusp, we may approximate a neighborhood of the cusp by narrower and narrower wedges, and conclude that $\lim \delta(p,q) = \infty$.

However, for a C^1 curve, the distortion $\delta(p,q)$ approaches 1 whenever p and q approach a common point, so we can view δ as a continuous function on the compact space $\gamma \times \gamma$, and it achieves its (finite) supremum somewhere (away from the diagonal).

Note. O'Hara [Oha1] defined a family of energy functionals E_j^p for knots as L^p norms

$$jE_j^p(\gamma) := \left\| |x-y|^{-j} - d(x,y)^{-j} \right\|_p$$

and observed that the distortion is a limiting value $\delta(\gamma) = \exp(E_0^\infty(\gamma))$.

Gromov mentions in [Gro2] that any closed rectifiable curve has distortion at least $\pi/2$. In fact, in [GLP] he proves a somewhat stronger result which includes the following proposition as a consequence. For completeness, we include a proof.

Proposition 2.1. *Given any closed rectifiable curve $\gamma \subset \mathbb{R}^n$, the distortion satisfies $\delta(\gamma) \geq \delta_\circ(\gamma) \geq \pi/2$, with equality if and only if γ is a round circle.*

Proof. Rescale so that $L(\gamma) = 2\pi$, and parameterize the curve by arclength $s \in \mathbb{R}/2\pi\mathbb{Z}$. Opposite points p and p^* have arclength $d(p,p^*) = \pi$, so we want to prove for some p that $|p-p^*| \leq 2$. For then $\delta(p,p^*) \geq \pi/2$ and we are done.

Consider the new curve $f(s) := p - p^* = \gamma(s) - \gamma(s+\pi)$ in $\mathbb{R}^n$. Observe that f is Lipschitz with speed at most two (since by the triangle inequality $|f'(s)| \leq |\gamma'(s)| + |\gamma'(s+\pi)| = 2$ almost everywhere), and that $f(s+\pi) = -f(s)$ for all s.

We want to show that $|f| \leq 2$ somewhere; suppose not, so that f lies *outside* the closed ball of radius two in $\mathbb{R}^n$. Then any arc of f from s to $s+\pi$ is an arc between antipodal points in $\mathbb{R}^n$ which avoids this closed ball, so its length exceeds the distance between antipodal points on a sphere of radius two, which is 2π. But since the parameterization of f has speed at most two, the length of this arc is at most 2π, a contradiction.

Note that the only way to get equality $\delta = \pi/2$ is to have $|f| \equiv 2$, with f tracing out a great circle on this sphere. Then p and p^* must always move in opposite directions, and p also traces out a round circle. □

Even if we had only proved this theorem for smooth curves, the following lemma would extend the result to arbitrary curves; it will also be useful in our discussion of distortion thickness.

Lemma 2.2. *In the C^0 topology on the space of curves, the distortion, considered as a map to $[1, \infty]$, is lower semicontinuous.*

Proof. In the C^0 topology on parameterized curves (not necessarily with constant speed parameterizations), the position of any point $p = \gamma(t)$ is by definition a continuous function of γ. Thus, for any points p and q the function $|p-q|$ is continuous in the C^0 topology, and the arclength $d(p,q)$ along γ is lower semicontinuous. Hence $\delta(p,q)$ is lower semicontinuous for each pair of points.

Remember that a real-valued function f on any topological space is lower semicontinuous if and only if $f^{-1}\{x : x > y\}$ is open for all $y \in \mathbb{R}$. Thus it is clear that if every member of some family $\{f_\alpha\}$ of functions is lower semicontinuous, then $f := \sup_\alpha f_\alpha$ is also lower semicontinuous. But the distortion $\delta(\gamma)$ is the supremum of $\delta(p,q)$, a family of lower semicontinuous functions, hence is itself lower semicontinuous. □

Note. This lemma, together with the fact that distortion blows up for nonembedded curves, leads us to expect that a curve minimizing the distortion should exist within each knot class. Gromov [Gro1] and O'Hara [Oha1] have independently observed that there are infinitely many knot types with distortion less than some constant $M < 100$. Gromov asked [Gro2, p. 114] if perhaps all knots have distortion under 100. In any case, distortion does not seem to be as useful for distinguishing knots as O'Hara's original energy E_2^1 [Oha, FHW, KS] has been, or as thickness promises to be.

Note. On a sufficiently smooth curve γ, we can make some observations about any pair (p,q) realizing the maximum distortion. First if we consider the variation of $\delta(p,q)$ from a perturbation fixing $d(p,q)$, we find that (p, q) must be a symmetric pair, meaning the chord between them makes equal angles θ with the tangents to γ at either end. If p and q are not opposite points on γ, then we can also consider variations increasing or decreasing $d(p, q)$; this shows that $\delta = \sec\theta$. But if $q = p^*$ is opposite p, all we can conclude is that $\delta \leq |\sec\theta|$, which bounds θ near $\pi/2$. Note that a doubly self-critical pair realizing the minimum self-distance $c_2(\gamma)$ cannot also realize the maximum distortion, unless the points are antipodal; some slightly more distant pair will have greater distortion.

3. Distortion and a Conjecture in Integral Geometry

Another approach to proving the proposition from the last section would be to use Crofton's formula and related notions from integral geometry, which relate the average sizes of projected images of a curve to the length of the curve.

Recall that if γ is a plane curve of length L, and we consider the (S^1-worth of) projected images of γ to the line, Crofton's formula says that the average length (with multiplicity) of these projections is $\frac{2}{\pi}L$. Since the multiplicity of the projected closed curve is at least 2 everywhere, we find that the average diameter of the projections is at most L/π. Thus, a plane curve of length 2π has width at most 2 in some direction.

Our proposition follows immediately for plane curves. If γ has length 2π, orient it so that its width horizontally is at most 2. Then find a pair of opposite points p and p^* at the same height (noting that the height difference changes sign). Clearly $|p - p^*| \leq 2$, while $d(p, p^*) = \pi$, so $\delta_\circ(\gamma) \geq \pi/2$.

Janse van Rensburg has suggested [JvR] that any space curve could be "unfolded" to a convex plane curve of the same length, while never decreasing the chord distances $|p - q|$. For instance, if a curve touches a supporting hyperplane in two disjoint places, we could reflect one intervening segment to lie below this plane. This procedure certainly never increases the distortion. If we could prove that it converged to a planar curve (perhaps considering only the polygonal case) then we would have a new proof of our proposition, except that it would seem difficult to show circles are the only curves of distortion $\delta = \pi/2$.

The argument sketched above for plane curves could be applied directly to curves in higher dimensions if the following conjecture is true. Again we would find a pair of opposite points on the original curve that have the same height in the direction of the projection, and they would then have distortion at least $\pi/2$.

Conjecture. If γ is a curve in $\mathbb{R}^n$ of length L, then there is some orthogonal projection to $\mathbb{R}^{n-1}$ in which the image of γ has diameter at most L/π.

Note. For space curves in $\mathbb{R}^3$ of length L, there are two standard analogues of Crofton's formula. The first deals with projections to a line, and says that the average length (with multiplicities), over the sphere of projections, is $L/2$. Thus some projection to a line has diameter at most $L/4$. The second deals with projections to a plane, and says that the average length of these, over the sphere of possible projections, is $\frac{\pi}{4}L$. Thus some projection to a plane has diameter at most $\frac{\pi}{8}L$, but this is not as good as our conjectured L/π.

4. The Distortion Thickness

Recall that the distortion thickness of a curve measures the least self-distance among pairs of points with large distortion:

$$\tau_b(\gamma) := \inf_{\delta(p,q)>b} |p-q|.$$

Clearly, as we increase b, $\tau_b(\gamma)$ is nondecreasing, since we take the infimum over smaller sets. For $b \le 1$, all pairs (p,q) are in contention, except those connected by a straight arc, so $\tau_1 = 0$ for all closed curves. For $b \ge \frac{\pi}{2}$, the infimum may be over the empty set (as it would be for a circle), so the thickness may be infinite. For $b = \frac{\pi}{2}$, although a circle has infinite thickness, no other closed curve does. Thus we concentrate on $b \in (1, \frac{\pi}{2}]$.

Note. If $b < \delta(\gamma)$, then the distortion ropelength L/τ_b is always at least $2\delta(\gamma) \ge \pi$. To check this, take p and q with $\delta(p,q)$ close to $\delta(\gamma)$. Of course $d(p,q) \le L/2$ and $|p-q| \ge \tau_b$, so dividing gives us $\delta(p,q) \le L/2\tau_b$.

Proposition 4.1. *The distortion thickness τ_b is upper semicontinuous in the C^0 topology on the set of closed rectifiable curves, for any fixed b.*

Proof. The claim is that if a sequence of curves γ^k approaches a limit γ^0, then their thicknesses $\tau^k := \tau_b(\gamma^k)$ satisfy

$$\overline{\lim}\, \tau^k \le \tau^0.$$

If not, write $\overline{\lim}\, \tau^k = \tau^0 + 3\epsilon$, for some $\epsilon > 0$; passing to a subsequence, we may assume that $\tau^k > \tau^0 + 2\epsilon$ for all k.

Now, the infimum in the definition of τ^0 may not be realized, but we can certainly find a pair of points (p,q) which come close to realizing

τ^0. That is, $|p-q|_0$, their straight-line distance on γ^0, is within ϵ of τ^0, while $\delta^0(p,q)$, their distortion on γ^0, is greater than b. So we have

$$\tau^k > |p-q|_0 + \epsilon, \qquad \delta^0(p,q) > b.$$

Since $|p-q|_k \to |p-q|_0$, but τ^k stays greater than this distance, we must have that (p,q) is not in contention for τ^k for all large enough k, that is, $\delta_k(p,q) \le b$. On the other hand, the semicontinuity of δ means that $\liminf \delta_k(p,q) \ge \delta_0(p,q) > b$. This contradiction completes the proof. □

Note. The length of curves is lower semicontinuous (but not continuous) in this C^0 topology. Although we have discussed the existence of an arclength parameterization for each curve, we do not assume when discussing limits that constant-speed parameterizations are used. For instance, consider curves obtained by replacing one side of a square by finer and finer zigzags of twice the length. If each is parameterized by arclength, they approach a limit, which traces out the square at varying speed.

Corollary 4.2. *In the C^0 topology, the distortion ropelength of curves is lower semicontinuous.*

Proof. It is the quotient of a lower semicontinuous function (arclength) by an upper semicontinuous function (thickness). □

Let us consider some example curves. First take γ to be the unit circle. Given any $\phi \in [0, \frac{\pi}{2})$, any pair of points at angle 2ϕ is at distance $|p-q| = 2\sin\phi$, so it has distortion $\phi/\sin\phi$. Thus if $b = \phi/\sin\phi \in [1, \frac{\pi}{2})$ the thickness $\tau_b(\gamma)$ will be $2\sin\phi$. (The infimum here is not realized.)

Next, recall that the distortion of a wedge of (external) angle θ is $\sec\frac{\theta}{2}$. Given b, set $\theta = 2\sec^{-1} b$; if a curve γ has any corner sharper than this, its thickness $\tau_b(\gamma)$ will be zero. However, on a polygonal curve with no corners sharper than this, no pair of points on adjacent segments will be in contention for τ_b.

Suppose γ bounds the convex hull of the unit circle and an exterior point (placed so that the angle at this point is θ). This curve has thickness $\tau_b = 0$ if $b < \sec\frac{\theta}{2}$, but if $b \ge \sec\frac{\theta}{2}$ its thickness is the same as that of the circle. This shows that the thickness is not C^0-continuous.

Proposition 4.3. *Suppose γ is a $C^{1,1}$ curve. If the infimum defining τ_b is attained, this is at a doubly self-critical pair, so $\tau_b \ge c_2$. If not, but $\tau_b < \infty$, there is a pair of points with $\delta(p,q) = b$ and $|p-q| = \tau_b$.*

Proof. If the infimum is attained at (p,q) with $\delta(p,q) > b$, then all nearby pairs on γ also have $\delta > b$ and are in contention for τ_b, so since

$|p-q|$ is a minimum, (p,q) must be a doubly self-critical pair. Otherwise, take a minimizing sequence (p_n,q_n), and pass to a convergent subsequence with some limit (p,q). Because the sequence was minimizing, $|p-q| = \tau_b(\gamma)$, with $\delta(p,q) \geq b$. But if $\delta(p,q) > b$, this pair would achieve the minimum, so instead $\delta(p,q) = b$. (Note also that this implies $d(p,q) = b\tau_b$.) ☐

Note. The minimum doubly self-critical distance c_2 can be achieved at a pair of arbitrarily small distortion $\delta > 1$, as if we put small hooks at the ends of a straight segment before continuing with a huge loop. Thus it is not clear in general that $\tau_b \leq c_2$. Perhaps, however, the arc between the points in a self-critical pair always has large distortion somewhere. Hence we might consider a refined definition of τ_b which allows in the infimum any (p,q) for which a subarc has distortion $\delta > b$.

5. The Curvature Thickness

Recall that $\sigma_k(\gamma) := \min\big(kr(\gamma), c_2(\gamma)\big)$, which is the smaller of the minimum doubly self-critical distance c_2 and the minimum radius of curvature r scaled by the factor k. This definition makes sense for any $k \geq 0$, but we are most interested in $k \in [1,2]$.

The special case of this definition with $k = 2$ was introduced in [LSDR]. This thickness σ_2 is characterized there as the maximum diameter of a tube around a C^2 curve γ for which the normal exponential map is an embedding. We will discuss this and similar notions in Section 6.

The further main results of [LSDR] deal also with σ_2 for C^2 curves and, translated into our notation, say that

1. If γ has ropelength $L/\sigma_2 < \frac{n\pi}{2}$ for some integer n, then there is some n-gon isotopic to γ (in the same knot class).
2. Thus, any nontrivial knot has ropelength at least $\frac{5\pi}{2}$.
3. If γ is a C^2 curve then its ropelength is at least twice its distortion: $L(\gamma)/\sigma_2(\gamma) \geq 2\delta(\gamma)$.
4. If γ is a C^2 curve, then $c_1(\gamma) \geq \sigma_2(\gamma)$. In other words, $c_1(\gamma) = c_2(\gamma)$ when either is less than $2r(\gamma)$

Note. This last result shows that $\sigma_k = \min(kr, c_1)$ for any $k \leq 2$, so that we could have defined σ_k in terms of c_1 instead of c_2.

Note. O'Hara [Oha1] attributes the definition of $c_1(\gamma)$ to Kuiper, and shows that if γ has $L/c_1 < n$ then it is isotopic to an n-gon. This is a weaker conclusion than the similar result of [LSDR], but requires no curvature bound.

Note. One of the authors has pointed out [Sul] that the discretization for thickness suggested by Stasiak [Sta] (and presumably used in numerical work reported in [KBM+]) actually discretizes σ_1. Here we approximate a smooth curve γ by a polygon with short edges of nearly equal length. Now look at solid cylinders of equal diameter around each edge. Increase the diameter until some nonadjacent cylinders first intersect; this gives an approximation of $\sigma_1(\gamma)$.

It is clear that $\sigma_k = 0$ for any curve with a sharp corner. Thus it is not possible to have any useful bound of the form $\sigma_k \geq C\tau_b$. However, we can bound τ_b below in terms of σ_k, because a curvature bound implies that short segments of a curve have small distortion. We use a lemma, which improves one from [Oha, p. 242] by a factor of two, and is in fact just a special case of Schur's Theorem (see [Che, p. 36]).

Lemma 5.1. *If $\kappa(\gamma) \leq 1$, then for any points p and q on γ with $d(p,q) = d \leq 2\pi$, we have $|p-q| \geq 2\sin d/2$, or equivalently,*

$$\delta(p,q) \leq \frac{d/2}{\sin d/2}.$$

Proof. Let x be the point halfway along the arc from p to q. Orient the curve so that the tangent vector is vertical at x. At arclength s from x, the curve has turned less than angle s, so the vertical component of the unit tangent vector is still at least $\cos s$. Thus out to arclength $d/2$, the vertical component of the difference vector $x-p$ or $q-x$ is at least $\int_0^{d/2} \cos s\, ds = \sin d/2$. Thus $|p-q| \geq 2\sin d/2$, as desired. □

Definition. For any $b \in [1, \frac{\pi}{2}]$, define k_b to be the unique solution of the equivalent equations

$$b = \frac{\arcsin k_b/2}{k_b/2}, \qquad \frac{bk_b/2}{\sin bk_b/2} = b.$$

Note that k_b increases monotonically from 0 to 2 as b increases from 1 to $\frac{\pi}{2}$, and thus $bk_b \leq \pi$.

Theorem 5.2. *For any $1 \leq b \leq \frac{\pi}{2}$ and any closed curve γ, we have*

$$\tau_b(\gamma) \geq \sigma_{k_b}(\gamma).$$

Proof. Write $k = k_b$. If $\sigma_k = 0$ (as when γ is not $C^{1,1}$) there is nothing to prove. Otherwise, rescale so that $\sigma_k(\gamma) = k$. Then we have $c_2(\gamma) \geq k$ and $\kappa \leq 1$ everywhere on γ. Now apply Proposition 4.3; if the infimum defining τ_b is attained, $\tau_b \geq c_2 \geq k$, and we are done.

Otherwise, the proposition gives us points p and q with $|p-q| = \tau_b$ and $d(p,q) = b\tau_b$. If $\tau_b < k$ we want to use the lemma to derive a contradiction. Certainly then $d(p,q) < kb \le \pi$ so the lemma applies, giving

$$b = \delta(p,q) \le \frac{b\tau_b/2}{\sin b\tau_b/2}.$$

But $\frac{x}{\sin x}$ is an increasing function, so since $b\tau_b < bk$, we get $b < \frac{bk/2}{\sin bk/2}$ contradicting the definition of k_b. □

In particular, we have $\tau_{\frac{\pi}{3}} \ge \sigma_1$ and $\tau_{\frac{\pi}{2}} \ge \sigma_2$. Since $\tau_b(\gamma)$ and $\sigma_k(\gamma)$ are nondecreasing in b and k respectively, it follows that $\tau_b \ge \sigma_k$ whenever $b > 2\arcsin(k/2)/k$. We can also get bounds for other b and k.

Corollary 5.3. *For any $b \in [1, \frac{\pi}{2}]$, $k \ge k_b$ and any curve γ,*

$$\tau_b(\gamma) \ge \frac{k_b}{k}\sigma_k(\gamma).$$

In particular, $\tau_{\frac{\pi}{3}} \ge \sigma_k/k$ for $k \ge 1$.

Proof. Apply the theorem to b and k_b, and use the fact that by definition of σ_k, we have $\sigma_{k_b} \le \sigma_k \le \frac{k}{k_b}\sigma_{k_b}$. □

Note. Combining the result $\tau_{\frac{\pi}{2}} \ge \sigma_2$ with our previous note about distortion ropelength, we recover the result from [LSDR] that $L/\sigma_2 \ge 2\delta(\gamma)$. In fact, their proof is similar in spirit to our proof of Theorem 5.2.

Note. If γ is the unit circle, then for any b we have $\tau_b(\gamma) = \sigma_{k_b}(\gamma) = k_b$, so the theorem is sharp.

6. Thickness Notions Related to Normal Tubes

Other authors have suggested various notions of thickness defined as the largest diameter of a normal tube before some property fails. From [LSDR], we have seen that σ_2 is the infimum of diameters for which the normal exponential map fails to be an embedding.

Diao *et al.* [DEJvR] suggest that we require the $\frac{d}{2}$-neighborhood to be a solid torus and then look at the intersection of this neighborhood with each normal plane (assuming γ is C^1). If further we ask that the zero-component of each such intersection be a meridian disk in the solid torus, intersecting γ only at the original point, then the first d for which this fails is their C^1-continuous measure of thickness, which we call $t_D(\gamma)$.

O'Hara [Oha2, p. 60] suggests defining thickness simply as the infimum of all d such that the $\frac{d}{2}$-neighborhood of γ is not topologically a solid torus; we will call this $t_O(\gamma)$. However, the notion he actually uses is different. Let t_B be the infimal diameter of balls in space whose intersection with γ is not a single arc (and thus is disconnected or all of γ). O'Hara's implicit claim is that $t_B \leq t_O$.[1] Certainly we can have $t_B < t_O$, as for a polygon, where $t_B = 0$.

It is clear from these definitions that $\sigma_2 \leq t_D \leq t_O$, and that $t_D \leq c_1$. Combining these with the final result of [LSDR] shows that $t_D = c_1 = c_2$ whenever these are less than $2r(\gamma)$.

For the ellipse $(x/a)^2 + (y/b)^2 = 1$ with $a < b$, we have $c_2 = 2a = t_O$, $r = a^2/b < a$, and $c_1 = t_D$ with $\sigma_2 = 2r < t_D < t_O$. (Explicit values of τ_b here would involve elliptic integrals.) Suppose we now take the upper half of this ellipse, and attach it in a $C^{1,1}$ fashion to a large loop in the lower halfplane; this new curve has $c_2 = 2a << t_O$.

Now consider a helix of radius ρ and pitch p, the intersection of a cylinder of radius ρ with the helicoid $z = pr\cos\theta$. (To get a closed curve, we could splice a long piece of this helix into a huge loop.) For steep pitch, this has no doubly self-critical points, but if $p^2 \leq -\min\left(\frac{\sin x}{x}\right)$, we have $c_2 < 3\rho$, approaching 0 as the pitch decreases. If the pitch is shallow enough, $c_2 \leq 2\rho$ and then $t_O = c_2$, but for an intermediate range of pitches (when $c_2 > 2\rho$ but is still finite), we again find $t_O >> c_2$, because by the time the normal tube is thick enough to see the self-critical points, it already extends across the axis of the cylinder, so it is still homotopy equivalent to the helix.

7. Shortest Curves in a Knot Class

We have made some progress in understanding the structure of curves minimizing the distortion ropelength, and believe the following:

Conjecture. For each nontrivial knot type K there exists a shortest curve $\gamma \subset \mathbb{R}^3$ of distortion thickness 1 in that knot type; that is, each knot type has a curve γ of shortest distortion ropelength. Any such γ has bounded curvature, and thus is of smoothness class $C^{1,1}$. Away from doubly self-critical points realizing the thickness, γ must be straight. If we scale to make the distortion thickness $\tau_{\frac{\pi}{2}}(\gamma) = 1$, then the curvature of γ is bounded by 1, so that also the curvature thickness $\sigma_k(\gamma) = 1$ for all $k \geq 1$. Finally, for any $b \geq \pi/3$, we have $\tau_b(\gamma) = 1$, so all our notions of thickness agree for a ropelength minimizer γ.

Note. It seems that the ropelength minimizer for the Hopf link must be the obvious candidate: two unit circles in perpendicular planes passing through each others' centers. For this link, $r = c_1 = c_2 = 1$, so $\sigma_k = 1$ for any $k \geq 1$. Also, the distortion thickness for any $b \geq \frac{\pi}{3}$ is $\tau_b = 1$. If

[1] In fact, for any diameter greater than t_O, he claimed to get a ball with disconnected intersection with γ. This can never be true if γ is a circle, so we have tried to correct for this in our definition of t_B.

we extend the definitions of t_D, t_O and t_B to links in the obvious ways, these also all equal 1. Presumably, in the ropelength minimizer for the connected sum of two Hopf links, the middle component (linking both others) will be a stadium curve, built of two semicircles and two straight segments; this shows we cannot expect minimizers to be C^2.

References

[Che] S. S. Chern. Curves and surfaces in euclidean space. In S. S. Chern, editor, *Studies in Global Geometry and Analysis*, pages 16–56. Math. Assoc. Amer., 1967.

[DEJvR] Yuanan Diao, Claus Ernst, and E. J. Janse van Rensburg. Energies of knots. Preprint, 1996.

[FHW] Michael H. Freedman, Zheng-Xu He, and Zhenghan Wang. On the Möbius energy of knots and unknots. *Annals of Math.* **139**(1994), 1–50.

[Gro1] Mikhael Gromov. Homotopical effects of dilatation. *J. Differential Geometry* **13**(1978), 303–310.

[Gro2] Mikhael Gromov. Filling Riemannian manifolds. *J. Differential Geometry* **18**(1983), 1–147.

[GLP] Mikhael Gromov, J. Lafontaine, and P. Pansu. *Structures métriques pour les variétés riemanniennes.* Cedic/Fernand Nathan, Paris, 1981.

[JvR] E. J. Janse van Rensburg. Personal communication, June 1996.

[KBM+] Vsevolod Katritch, Jan Bednar, Didier Michoud, Robert G. Scharein, Jacques Dubochet, and Andrzej Stasiak. Geometry and physics of knots. *Nature* **384**(November 1996), 142–145.

[KS] Robert B. Kusner and John M. Sullivan. Möbius energies for knots and links, surfaces and submanifolds. In Willam H. Kazez, editor, *Geometric Topology*, pages 570–604. Amer. Math. Soc./International Press, 1997. Proceedings of the Georgia Int'l Topology Conference, August 1993.

[LSDR] Richard A. Litherland, Jon Simon, Oguz Durumeric, and Eric Rawdon. Thickness of knots. Preprint, 1996.

[Oha] Jun O'Hara. Energy of a knot. *Topology* **30**(1991), 241–247.

[Oha1] Jun O'Hara. Family of energy functionals of knots. *Topology Appl.* **48**(1992), 147–161.

[Oha2] Jun O'Hara. Energy functionals of knots II. *Topology Appl.* **56**(1994), 45–61.

[Sta] Andrzej Stasiak. Lecture at the AMS special session on *Physical Knot Theory*, Iowa City, March 1996.

[Sul] John M. Sullivan. Lecture at the IMA workshop on *Topology and Geometry in Polymer Science*, June 1996.

[Tor] Tatiana Toro. Lecture at the IMA workshop on *Topics Related to Nonlinear PDE*, March 1996.

Institute for Advanced Study
E-mail address: `kusner@math.umass.edu`

University of Illinois
E-mail address: `sullivan@geom.umn.edu`