SHORTEST CLOSED CURVE TO INSPECT A SPHERE

A Dissertation
Presented to
The Academic Faculty

By

James Wenk

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics
Department of Mathematics

Georgia Institute of Technology

August 2022

SHORTEST CLOSED CURVE TO INSPECT A SPHERE

Thesis committee:

Dr. Mohammad Ghomi, Advisor
School of Mathematics
Georgia Institute of Technology

Dr. Rob Kusner
Department of Mathematics
University of Massachusetts

Dr. Igor Belegradek
School of Mathematics
Georgia Institute of Technology

Dr. Galyna Livshyts
School of Mathematics
Georgia Institute of Technology

Dr. Jason Cantarella
Department of Mathematics
University of Georgia

Dr. Michael Loss
School of Mathematics
Georgia Institute of Technology

Date approved: July 5, 2022

I would like to dedicate this thesis to my family. To my Mother and Father, for their unwavering support and encouragement throughout my life, without which I would not be who I am today. To my sister Laura, for the fun and laughter we had always brightened a dreary week. To Clara, whose patience and understanding of my mathematical obsessions were that of a Saint. And to my cats, C.C. and Tiger, who passed during my studies, you will always be remembered for the fuzzy balls of love you were.

ACKNOWLEDGMENTS

    I would like to thank the members of my thesis committee for their help and feedback in preparing my thesis. Special thanks to my advisor, who's guidance and assistance on formalizing mathematics was extremely helpful, and one day I hope to emulate. I would also like to thank all the friends I have made along the way. The shared nerdiness, jokes and complaining we had made this journey bearable.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

We show that in Euclidean 3-space any closed curve γ which lies outside the unit sphere and contains the sphere within its convex hull has length $\geq 4\pi$. Equality holds only when γ is composed of 4 semicircles of length π, arranged in the shape of a baseball seam, as conjectured by V. A. Zalgaller in 1996.

CHAPTER 1

INTRODUCTION

What is the shortest closed orbit a sattelite may take to inspect the entire surface of a round asteroid? This is a well-known optimization problem [1, 2, 3, 4, 5, 6] in classical differential geometry and convexity theory, which was resolved in [7]. In this work we give a more fleshed out argument, which proceeds the same as in the previous work. Now the question which we are interested in may be precisely formulated as follows. A curve γ in Euclidean space $\mathbf{R}^3$ *inspects* a sphere S provided that it lies outside S and each point p of S can be "seen" by some point q of γ, i.e., the line segment pq intersects S only at p. It is easily shown that the latter condition holds if and only if S lies in the convex hull of γ. The supremum of the radii of the spheres which are contained in the convex hull of γ and are disjoint from γ is called the *inradius* of γ. Thus we seek the shortest closed curve with a given inradius. The answer is as follows:

Theorem 1.1. *Let $\gamma\colon [a,b] \to \mathbf{R}^3$ be a closed rectifiable curve of length L and inradius r. Then*

$$L \geq 4\pi r. \tag{1.1}$$

Equality holds only if, up to a reparameterization, γ is simple, $\mathcal{C}^{1,1}$, lies on a sphere of radius $\sqrt{2}\,r$, and traces consecutively 4 semicircles of length πr.

It follows that the image of the minimal curve is unique up to a rigid motion, and resembles the shape of a baseball seam as shown in Figure 1.1, which settles a conjecture of Viktor Zalgaller made in 1996 [1]. The previous best estimate was $L \geq 6\sqrt{3}\,r$ obtained in 2018 [4]. Here we use some notions from [4] together with other techniques from integral geometry (Crofton type formulas), geometric knot theory (unfoldings of space curves), and geometric measure theory (tangent cones, sets of positive reach) to establish the above

1
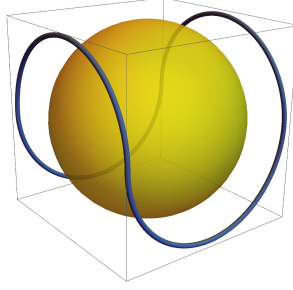
Figure 1.1: The unique minimizer, known as the Baseball Curve

theorem. We also derive a number of formulas (chapter 3, chapter 6, and chapter 8) for the inspection efficiency of curves, which may be verified with the aid of the computer software package that we have provided [8].

Our main approach for proving Theorem 1.1 is as follows. Since (Equation 1.1) is invariant under rescaling and rigid motions, we may assume that $r = 1$ and γ inspects the unit sphere $\mathbf{S}^2$, in which case we say simply that γ is an *inspection curve*. Then we define the *horizon* of γ (chapter 3) as the measure in $\mathbf{S}^2$ counted with multiplicity of the set of points $p \in \mathbf{S}^2$ where the tangent plane $T_p\mathbf{S}^2$ intersects γ:

$$H(\gamma) := \int_{p \in \mathbf{S}^2} \#\gamma^{-1}(T_p\mathbf{S}^2) \, dp.$$

Since γ is closed, $T_p\mathbf{S}^2$ intersects γ at least twice for almost every $p \in \mathbf{S}^2$. Thus $H(\gamma) \geq 8\pi$. Next we define the *(inspection) efficiency* of γ as

$$E(\gamma) := \frac{H(\gamma)}{L(\gamma)}. \tag{1.2}$$

So to establish (Equation 1.1) it suffices to show that $E(\gamma) \leq 2$. Now note that, since H is additive, for any partition of γ into subsets γ_i, $i \in I$,

$$E(\gamma) = \sum_i \frac{H(\gamma_i)}{L(\gamma)} = \sum_i \frac{L(\gamma_i)}{L(\gamma)} E(\gamma_i) \leq \sup_i E(\gamma_i). \tag{1.3}$$

So the desired upper bound for $E(\gamma)$ may be established through a partitioning of γ into

subsets γ_i with $E(\gamma_i) \le 2$.

To find the desired partition, we may start by assuming that γ in Theorem 1.1 has minimal length among all (closed) inspection curves, and is parameterized with constant speed (chapter 2). Then we apply an "unfolding" procedure, studied by Cantarella, Kusner and Sullivan [9], to transform γ into a planar curve $\widetilde{\gamma}$ with the same arclength and *height*, i.e., radial distance function from the origin o of $\mathbf{R}^3$ (chapter 4). It follows that $E(\gamma) = E(\widetilde{\gamma})$. Furthermore, the minimality of γ will ensure that $\widetilde{\gamma}$ is "locally convex with respect to o" [10]. Consequently $\widetilde{\gamma}$ may be partitioned into a collection of curves $\widetilde{\gamma}_i$ we call spirals (chapter 5). A *spiral* is a planar curve which lies outside the unit circle $\mathbf{S}^1$, is locally convex with respect to o, has monotone height, and is orthogonal to the position vector of its closest boundary point to o. We will show that the efficiency of any spiral is at most 2 by polygonal approximations and a variational argument (chapter 7), which establish (Equation 1.1).

The rest of the paper will be devoted to characterizing minimal inspection curves. First we note that equality holds in (Equation 1.1) only when $E(\gamma) = 2$, which forces all the spirals $\widetilde{\gamma}_i$ to have efficiency 2 as well. Then we show that a spiral has efficiency 2 only when it has constant height $\sqrt{2}$. The argument will depend on whether the minimum height of the spiral is above or below $\sqrt{2}$. In the former case we will compute that the *instantaneous efficiency* of the spiral is less than 2 (chapter 8), and in the latter case we devise a variational procedure called *lifting and splitting* to estimate the efficiency of the portion of the spiral below the height $\sqrt{2}$ (chapter 9). Once we know that any minimal inspection curve γ has constant height $\sqrt{2}$, we proceed to the final stages of the characterization (chapter 10). The simplicity of γ follows from a Crofton type formula of Blaschke-Santalo. Then a characterization of $\mathcal{C}^{1,1}$ submanifolds in terms of their tangent cones [11] and reach in the sense of Federer, ensures the regularity of γ. Finally we show that γ is composed of 4 semicircles by constructing a "nested partition" of γ [12], which completes the proof of Theorem 1.1. The last technique goes back to the proofs of the classical 4-vertex theorem due to Kneser and Bose [13, 14], which has been developed further by Umehara

and Thorbergsson [15, 16].

The question we study in this work belongs to a circle of long standing optimization problems for the length of a curve in Euclidean space subject to various constraints on its convex hull, including bounds on volume, surface area, width, and inradius [4, 17, 1, 18, 19, 20] [5, A28, A30]. With the exception of results in dimension 2, and the above result in dimension 3, most of these problems remain open; see [4, 21] for more background and references. We should also note that these problems may be posed both for closed and open curves. In the latter case, there are connections to the "lost in a forest problem" of Bellman [22], or its dual version, Moser's "worm problem" [23, 24, 25], which are well-known in computational geometry.

CHAPTER 2

EXISTENCE OF MINIMAL INSPECTION CURVES

The central objects of study in this work are rectifiable curves, which become Lipschitz mappings after reparameterization with constant speed. We begin by recording some basic facts in this regard; more extensive background may be found in [26], [27, chap. 2], or [28, chap. 4]. These notions will be used in this to establish the existence of minimal inspection curves, and study their tangent lines.

Here $\mathbf{R}^n$ denotes the n-dimensional Euclidean space with origin o, standard inner product $\langle \cdot, \cdot \rangle$, and induced norm $|\cdot| := \langle \cdot, \cdot \rangle^{\frac{1}{2}}$. Furthermore $\mathbf{S}^{n-1}$, B^n denote respectively the unit sphere and closed unit ball in $\mathbf{R}^n$. The *interior*, *closure*, and *boundary* of any set $X \subset \mathbf{R}^n$ will be denoted by $\operatorname{int}(X)$, $\overline{X}$, and ∂X respectively. By a *curve* we shall mean a continuous mapping $\gamma \colon [a, b] \to \mathbf{R}^n$, where $[a, b] \subset \mathbf{R}$ denotes a closed interval with $a < b$. We will also use γ to refer to its image, $\gamma([a, b])$. We say that γ is *closed* if $\gamma(a) = \gamma(b)$. A closed curve γ is *simple* if it is one-to-one on $[a, b)$, and is $\mathcal{C}^1$ provided that it is continuously differentiable with $\gamma'_+(a) = \gamma'_-(b)$. Furthermore, γ is $\mathcal{C}^{1,1}$ if γ' is Lipschitz. The *length* of γ is defined as

$$L(\gamma) := \sup \sum_{i=1}^{n} \big| \gamma(t_i) - \gamma(t_{i-1}) \big|,$$

where the supremum is taken over all partitions $a := t_0 \leq t \leq \cdots \leq t_n := b$ of $[a, b]$. We say that γ is *rectifiable* provided that $L(\gamma)$ is finite. Furthermore, γ has *constant speed* C provided $L(\gamma|_{[t,s]}) = C\,|t - s|$ for all $t < s \in [a, b]$. So constant speed curves are rectifiable. We say that $\widehat{\gamma} \colon [c, d] \to \mathbf{R}^n$ is a *reparameterization* of γ provided that there exists a nondecreasing continuous map $\phi \colon [a, b] \to [c, d]$ such that $\gamma = \widehat{\gamma} \circ \phi$. If $[a, b] = [c, d]$, then we say that $\widehat{\gamma}$ is *domain preserving*. Note that $L(\widehat{\gamma}) = L(\gamma)$. It is also well-

known [27, Prop. 2.5.9] that:

Lemma 2.1. *Any rectifiable curve* $\gamma\colon [a,b] \to \mathbf{R}^n$ *admits a domain preserving reparameterization with constant speed.*

Thus we may assume that all rectifiable curves in this work have constant speed. Next note that if $\gamma\colon [a,b] \to \mathbf{R}^n$ has constant speed C, then for all $t < s \in [a,b]$,

$$|\gamma(t) - \gamma(s)| \leq L\big(\gamma\big|_{[t,s]}\big) = C|t-s|. \tag{2.1}$$

So γ is C-*Lipschitz*, and therefore differentiable almost everywhere by Rademacher's theorem. Then $L(\gamma) = \int_a^b |\gamma'(t)|\, dt$ [27, Thm. 2.7.6]. Furthermore, (Equation 2.1) implies that $|\gamma'| \leq C$ at all differentiable points of γ. On the other hand, $\int_a^b |\gamma'(t)|\, dt/(b-a) = L(\gamma)/(b-a) = C$. Thus $|\gamma'| = C$ almost everywhere. So we have established:

Lemma 2.2. *Let* $\gamma\colon [a,b] \to \mathbf{R}^n$ *be a rectifiable curve. Then* γ *has constant speed if and only if* $|\gamma'| = L(\gamma)/(b-a)$ *almost everywhere.*

In particular, when $L(\gamma) > 0$, we may assume after a reparameterization that $|\gamma'| \neq 0$ almost everywhere. Let $\mathcal{C}^0([a,b], \mathbf{R}^n)$ denote the space of curves $\gamma\colon [a,b] \to \mathbf{R}^n$ with the supremum norm or *uniform metric* [27, p. 47] given by

$$\mathrm{dist}(\gamma_1, \gamma_2) := \sup_{t \in [a,b]} |\gamma_1(t) - \gamma_2(t)|. \tag{2.2}$$

Unless noted otherwise, *convergence* of curves in this paper will be with respect to the uniform metric. It is well-known that the length functional $L\colon \mathcal{C}^0([a,b], \mathbf{R}^n) \to \mathbf{R}$ is lower semi-continuous [27, Prop. 2.3.4(iv)]. The *convex hull* of a set $X \subset \mathbf{R}^n$, denoted by $\mathrm{conv}(X)$, is the interof all closed half-spaces containing X. We say that $\gamma\colon [a,b] \to \mathbf{R}^3$ is an *inspection curve* if it is closed and $\mathbf{S}^2 \subset \mathrm{conv}(\gamma)$.

Proposition 2.3. *There exists an inspection curve of minimum length.*

Proof. After a reparameterization, any inspection curve $\gamma\colon [a,b] \to \mathbf{R}^3$ may be identified with an element of $\mathcal{C}^0([0,1], \mathbf{R}^3)$. Let $X \subset \mathcal{C}^0([0,1], \mathbf{R}^3)$ denote the collection of inspection curves whose length is bounded above by some constant C, chosen sufficiently large so that $X \neq \emptyset$. Since the length functional L is lower semi-continuous on $C^0([0,1], \mathbf{R}^3)$, it suffices to show that X is compact. To this end we may assume that all elements of X have constant speed by Lemma 2.1. Let $\gamma_i \in X$ be a sequence of curves. Then, by Lemma 2.2, $|\gamma_i'| = L(\gamma_i) \leq C$ almost everywhere, and therefore γ_i are C-Lipschitz. Furthermore, since $o \in \text{conv}(\gamma_i)$, γ_i are confined within a ball of radius C centered at o. So by Arzela-Ascoli theorem [27, Thm. 2.5.14], a subsequence of γ_i converges to a C-Lipschitz and therefore rectifiable curve $\gamma\colon [0,1] \to \mathbf{R}^3$. Since γ is C-Lipschitz, $|\gamma'| \leq C$ almost everywhere, which in turn yields that $L(\gamma) \leq C$. Furthermore since $\mathbf{S}^2 \subset \text{conv}(\gamma_i)$, it follows that $\mathbf{S}^2 \subset \text{conv}(\gamma)$, which means that γ is an inspection curve. So we conclude that $\gamma \in X$ as desired. $\qquad\square$

Any curve given by the above proposition will be called a *minimal* inspection curve. Next we establish an important property concerning tangent lines of these curves. Let $\angle(v, w) := \cos^{-1}(\langle v, w \rangle / (|v||w|))$ denote the *angle* between $v, w \in \mathbf{R}^n \setminus \{o\}$. For any rectifiable curve $\gamma\colon [a,b] \to \mathbf{R}^n \setminus \{o\}$, with $L(\gamma) > 0$, we set

$$\alpha(t) := \angle\big(\gamma(t), \gamma'(t)\big).$$

By Lemma 2.2, if γ has constant speed, then $|\gamma'| \neq 0$ almost everywhere. So α is well-defined for almost every $t \in [a,b]$. The *tangent cone* $T_t\gamma$ of γ at $t \in [a,b]$ is the collection of all rays emanating from $\gamma(t)$ which are limits of a sequence of secant lines emanating from $\gamma(t)$ and passing through points $\gamma(s_i)$ as s_i converge to t. If γ is closed and $t = a$ or b, then we set $T_t\gamma := T_a\gamma \cup T_b\gamma$. See [11, Sec. 2] for basic facts and background on tangent cones. If $T_t\gamma$ is a line, then we call it the *tangent line* of γ at t. In particular when γ is differentiable at t and $|\gamma'(t)| \neq 0$, then $T_t\gamma$ is the line through $\gamma(t)$ spanned by $\gamma'(t)$.

Thus almost all tangent cones of a constant speed curve γ, with $L(\gamma) > 0$, are lines. The following lemma generalizes an earlier observation [4, Lem. 7.4] for polygonal curves.

Lemma 2.4. *Let $\gamma\colon [a,b] \to \mathbf{R}^3$ be a constant speed minimal inspection curve. Then tangent lines of γ avoid* $\operatorname{int}(B^3)$. *In particular*

$$\alpha(t) \geq \sin^{-1}\left(\frac{1}{|\gamma(t)|}\right),\tag{2.3}$$

for almost every $t \in [a,b]$.

Proof. Let T be a tangent line of γ at $t \in [a,b]$. Suppose towards a contradiction that T intersects $\operatorname{int}(B^3)$. Set $X := \operatorname{conv}(\{\gamma(t)\} \cup B^3)$. Then T intersects $\operatorname{int}(X)$. Consequently there exists a open interval $U \subset [a,b]$ of the form (t,s) or (s,t) such that $\gamma(U) \subset \operatorname{int}(X)$, and $\gamma(s)$ lies on ∂X. Let $\overline{U}$ denote the closure of U. Replacing $\gamma(\overline{U})$ with a line segment connecting $\gamma(t)$ and $\gamma(s)$ (or if $\gamma(t) = \gamma(s)$ then cutting out $\gamma(U)$) yields a closed curve β with $L(\beta) < L(\gamma)$. On the other hand, $\operatorname{conv}(\beta) = \operatorname{conv}(\gamma)$, since $\gamma(U) \subset \operatorname{int}(\operatorname{conv}(X)) \subset \operatorname{int}(\operatorname{conv}(\gamma))$. In particular $\mathbf{S}^2 \subset \operatorname{conv}(\beta)$. So β is an inspection curve shorter than γ, which is the desired contradiction. Now (Equation 2.3) follows from basic trigonometry. $\square$

CHAPTER 3

THE INTEGRAL FORMULA FOR EFFICIENCY

As we had mentioned in the introduction, the *efficiency* of any rectifiable curve $\gamma\colon [a,b] \to \mathbf{R}^3$ with $|\gamma| \geq 1$ is defined as

$$E(\gamma) := \frac{H(\gamma)}{L(\gamma)}, \qquad \text{where} \qquad H(\gamma) := \int_{p \in \mathbf{S}^2} \#\gamma^{-1}(T_p\mathbf{S}^2)\, dp. \qquad (3.1)$$

Recall also that $H(\gamma)$ is called the *horizon* of γ, and $\#$ indicates cardinality. When γ is an inspection curve it follows from Caratheodory's convex hull theorem that

$$\#\gamma^{-1}(T_p\mathbf{S}^2) \geq 2$$

for almost every $p \in \mathbf{S}^2$, see [4, Lemma 7.1]. Thus $H(\gamma) \geq 8\pi$. So to prove (Equation 1.1) it suffices to show that $E(\gamma) \leq 2$. To this end we use the area formula in geometric measure theory [29, Thm 3.2.3] to compute $E(\gamma)$. This generalizes previous work in [4, Sec. 7.2] where the following proposition had been established under the additional restrictions that γ be piecewise $\mathcal{C}^{1,1}$ and its tangent lines pass through o only finitely many times. Recall that $\alpha := \angle(\gamma, \gamma')$.

Proposition 3.1. *Let $\gamma\colon [a,b] \to \mathbf{R}^3$ be a constant speed curve with $|\gamma| \geq 1$. Then*

$$E(\gamma) = \frac{1}{b-a} \int_a^b \int_0^{2\pi} \frac{1}{|\gamma|^2} \left| \sqrt{|\gamma|^2 - 1}\, \sin\left(\alpha\right) \cos(\theta) + \cos\left(\alpha\right) \right|\, d\theta dt. \qquad (3.2)$$

Proof. Let $\overline{\gamma}(t) := \gamma(t)/|\gamma(t)|$. Since γ is Lipschitz, $\overline{\gamma}$ is Lipschitz as well. So there exists a point $x \in \mathbf{S}^2 \setminus \overline{\gamma}$. Let e_1 be a $\mathcal{C}^1$ unit tangent vector field on $\mathbf{S}^2 \setminus \{x\}$, and set $e_2(p) := p \times u(p)$. Then $(e_1(p), e_2(p))$ is a Lipschitz orthonormal frame on any compact

subset of $\mathbf{S}^2 \setminus \{x\}$. So if we set

$$e_1(t) := e_1\big(\overline{\gamma}(t)\big), \quad e_2(t) := e_2\big(\overline{\gamma}(t)\big),$$

then $t \mapsto (\overline{\gamma}(t), e_1(t), e_2(t))$ is a Lipschitz orthonormal frame. Next set

$$\lambda := \frac{1}{|\gamma|}, \quad \text{and} \quad \rho := \sqrt{1 - \lambda^2}.$$

Define $F \colon [a, b] \times [0, 2\pi] \to \mathbf{S}^2$ by

$$F(t, \theta) = \lambda(t)\overline{\gamma}(t) + \rho(t)\big(\cos(\theta)e_1(t) + \sin(\theta)e_2(t)\big).$$

Then $\theta \mapsto F(t, \theta)$ parameterizes the *horizon circle* $H(\gamma(t))$, i.e., the set of points in $\mathbf{S}^2$ generated by all the rays which emanate from $\gamma(t)$ and are tangent to $\mathbf{S}^2$ (if $\gamma(t) \in \mathbf{S}^2$, then $H(\gamma(t))$ degenerates into a single point). So, for all $p \in \mathbf{S}^2$,

$$F^{-1}(p) = \gamma^{-1}(T_p\mathbf{S}^2).$$

Thus, since F is Lipschitz, the area formula [29, Thm 3.2.3] yields that

$$H(\gamma) = \int_{p \in \mathbf{S}^2} \#F^{-1}(p)\, dp = \int_a^b \int_0^{2\pi} JF(t, \theta)\, d\theta dt,$$

where $JF := |\partial F/\partial t \times \partial F/\partial \theta|$ is the Jacobian of F. Next, for every differentiable point $t \in [a, b]$ of γ let

$$E_\gamma(t) := \int_0^{2\pi} JF(t, \theta)\, d\theta.$$

By the Lebesgue differentiation theorem, for almost every $t \in [a, b]$,

$$E_\gamma(t) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} E_\gamma(s)\, ds = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} H\big(\gamma|_{[t-\varepsilon, t+\varepsilon]}\big).$$

10

So $E_\gamma(t)$ does not depend on the choice of the frame (e_1, e_2). We claim that for almost every point $t_0 \in [a, b]$ of γ we may choose the frame (e_1, e_2) so that

$$JF(t_0, \theta) = \frac{1}{|\gamma(t_0)|^2} \left| \sqrt{|\gamma(t_0)|^2 - 1} \sin\left(\alpha(t_0)\right) \cos(\theta) + \cos\left(\alpha(t_0)\right) \right| |\gamma'(t_0)|. \quad (3.3)$$

This would complete the proof because $E(\gamma) = (\int_a^b E_\gamma(t)\, dt)/L(\gamma)$, and since the speed is constant $|\gamma'| = L(\gamma)/(b-a)$. To establish (Equation 3.3) note that if t_0 is a differentiable point of γ, then it is a differentiable point of $\overline\gamma$ as well. There are two cases to consider: either $\overline\gamma'(t_0) \neq 0$ or $\overline\gamma'(t_0) = 0$.

First suppose that $\overline\gamma'(t_0) \neq 0$. Let C be the great circle in $\mathbf{S}^2$ which is tangent to $\overline\gamma$ at $\overline\gamma(t_0)$. Set $e_1(\overline\gamma(t_0)) := \overline\gamma'(t_0)/|\overline\gamma'(t_0)|$. We may extend e_1 smoothly to a unit tangent vector field in a neighborhood of $\overline\gamma(t_0)$ on $\mathbf{S}^2$ so that $e_1(p)$ is tangent to C when $p \in C$. Recall that $e_2(p) := p \times e_1(p)$, and $e_1(t) := e_1(\overline\gamma(t))$, $e_2(t) := e_2(\overline\gamma(t))$. Let $v := |\overline\gamma'(t_0)|$. Then

$$\begin{pmatrix} \overline\gamma \\ e_1 \\ e_2 \end{pmatrix}'(t_0) = \begin{pmatrix} 0 & v & 0 \\ -v & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \overline\gamma \\ e_1 \\ e_2 \end{pmatrix}(t_0).$$

Now one may compute [8] that at $t = t_0$,

$$\begin{aligned} \frac{\partial F}{\partial t} &= \left(\lambda' + \rho v \cos(\theta)\right)\overline\gamma + \left(\lambda v + \rho' \cos(\theta)\right)e_1 + \rho' \sin(\theta)e_2, \\ \frac{\partial F}{\partial \theta} &= -\rho \sin(\theta)e_1 + \rho \cos(\theta)e_2. \end{aligned}$$

It follows that

$$\left| \frac{\partial F}{\partial t} \times \frac{\partial F}{\partial \theta} \right|^2 = \rho^2 \left(v^2 \cos^2(\theta) - 2v \cos(\theta)(\lambda'\rho - \lambda\rho') + \left((\rho')^2 + (\lambda')^2\right) \right).$$

11

It is easy to check that $\lambda'\rho - \lambda\rho' = \lambda'/\rho$ and $(\rho')^2 + (\lambda')^2 = (\lambda'/\rho)^2$. So

$$JF = |\rho v \cos(\theta) - \lambda'|,$$

which is equivalent to (Equation 3.3) since

$$v = \frac{\sqrt{|\gamma'|^2|\gamma|^2 - \langle\gamma,\gamma'\rangle^2}}{|\gamma|^2} = \frac{\sin(\alpha)|\gamma'|}{|\gamma|} \qquad \text{and} \qquad \lambda' = -\frac{\langle\gamma,\gamma'\rangle}{|\gamma|^3} = -\frac{\cos(\alpha)|\gamma'|}{|\gamma|^2}.$$

It remains to consider the case where $\overline{\gamma}'(t_0) = 0$. Then $e_1'(t_0) = e_2'(t_0) = 0$ as well. So, at $t = t_0$,

$$\frac{\partial F}{\partial t} = \lambda'\overline{\gamma} + \rho'\cos(\theta)e_1 + \rho'\sin(\theta)e_2,$$

and $\partial F/\partial\theta$ is as computed above. Since $\overline{\gamma}' = 0$, $\alpha = 0$ or π. Thus $|\lambda'| = |\gamma'|/|\gamma|^2$. So

$$\left|\frac{\partial F}{\partial t} \times \frac{\partial F}{\partial\theta}\right|^2 = \rho^2\left((\lambda')^2 + (\rho')^2\right) = (\lambda')^2 = \left(\frac{|\gamma'|}{|\gamma|^2}\right)^2.$$

Hence $JF = |\gamma'|/|\gamma|^2$, which establishes (Equation 3.3) since $\alpha = 0$ or π. $\qquad\square$

Note that if $\gamma\colon [a,b] \to \mathbf{R}^3$ has constant speed C, then

$$\big||\gamma(s)| - |\gamma(t)|\big| \leq |\gamma(t) - \gamma(s)| \leq C|t - s|.$$

So the function $|\gamma|\colon [a,b] \to \mathbf{R}$, which we call the *height* of γ, is Lipschitz. In particular, $|\gamma|$ is differentiable almost everywhere. Furthermore note that if t is a differentiable point of both γ and $|\gamma|$, then $|\gamma|' = \langle\gamma,\gamma'\rangle/|\gamma|$ at t. Thus for almost every $t \in [a,b]$

$$\alpha(t) = \cos^{-1}\left(\frac{\langle\gamma(t),\gamma'(t)\rangle}{|\gamma(t)|C}\right) = \cos^{-1}\left(\frac{|\gamma|'(t)}{C}\right). \tag{3.4}$$

This shows, via Proposition 3.1, that $E(\gamma)$ depends only on $|\gamma|$. Hence we conclude

Corollary 3.2. *Let γ_1, $\gamma_2 \colon [a, b] \to \mathbf{R}^3$ be constant speed curves with $L(\gamma_1) = L(\gamma_2)$. Furthermore suppose that $|\gamma_1(t)| = |\gamma_2(t)| \geq 1$ for all $t \in [a, b]$. Then $E(\gamma_1) = E(\gamma_2)$.*

CHAPTER 4

UNFOLDING OF MINIMAL INSPECTION CURVES

Here we describe a natural operation, called unfolding [9], which transforms a rectifiable space curve into a planar one. This operation preserves the arclength and height of the curve, and thus will preserve its efficiency due to the results of the last section. Furthermore we will show that the unfolding of any minimal inspection curve satisfies a certain convexity condition. Let $\gamma \colon [a, b] \to \mathbf{R}^3 \setminus \{o\}$ be a rectifiable curve. Then we set $\overline{\gamma} := \gamma/|\gamma|$, and let

$$\theta_\gamma(t) := L\big(\overline{\gamma}\big|_{[a,t]}\big) = \int_a^t |\overline{\gamma}'(t)| dt,$$

denote the arclength function of $\overline{\gamma}$ (θ_γ measures the "cone angle" [9] or "vision angle" [30, 31] of γ from the point of view of o). The *(cone) unfolding* of γ is defined as the planar curve $\widetilde{\gamma} \colon [a, b] \to \mathbf{R}^2$ given by

$$\widetilde{\gamma}(t) := |\gamma(t)| e^{i\theta_\gamma(t)},$$

where $e^{i\theta_\gamma} = (\cos(\theta_\gamma), \sin(\theta_\gamma))$. In other words, $\widetilde{\gamma}$ is generated by the isometric immersion (or unrolling) into $\mathbf{R}^2$ of the conical surface generated by the line segments $o\gamma(t)$. Note that $|\widetilde{\gamma}(t)| = |\gamma(t)|$. Furthermore, assuming γ is reparameterized with constant speed,

$$\widetilde{\gamma}' = (|\gamma|' + i|\gamma|\theta_\gamma') e^{i\theta_\gamma}, \qquad \text{and} \qquad \theta_\gamma' = |\overline{\gamma}'| = \frac{1}{|\gamma|^2} \sqrt{|\gamma|^2 |\gamma'|^2 - \langle \gamma, \gamma' \rangle^2}, \qquad (4.1)$$

almost everywhere. Thus it follows that, for almost all $t \in [a, b]$,

$$|\widetilde{\gamma}'|^2 = (|\gamma|')^2 + |\gamma|^2 (\theta_\gamma')^2 = \frac{\langle \gamma, \gamma' \rangle^2}{|\gamma|^2} + \frac{1}{|\gamma|^2}\Big(|\gamma|^2|\gamma'|^2 - \langle \gamma, \gamma' \rangle^2\Big) = |\gamma'|^2.$$

So γ and $\widetilde{\gamma}$ have equal height and length. Hence, by Corollary 3.2,

Proposition 4.1. *Let* $\gamma\colon [a,b] \to \mathbf{R}^3$ *be a rectifiable curve with* $|\gamma| \neq 0$, *and* $\widetilde{\gamma}$ *be the unfolding of* γ. *Then* $E(\gamma) = E(\widetilde{\gamma})$.

Next we develop some geometric properties of $\widetilde{\gamma}$. First we record that

Lemma 4.2. *Let* $\gamma\colon [a,b] \to \mathbf{R}^3$ *be a minimal inspection curve with constant speed. Then* $\widetilde{\gamma}$ *is locally one to one.*

Proof. It suffices to show that θ_γ is increasing, or $\theta'_\gamma > 0$ almost everywhere. The formula for θ'_γ in (Equation 4.1), via the Cauchy-Schwartz inequality, shows that $\theta'_\gamma \geq 0$, and $\theta'_\gamma = 0$ only when γ' vanishes, or else γ and γ' are parallel. But γ' can vanish only on a set of measure zero, since γ has constant speed. Furthermore if γ and γ' are parallel, then $\alpha = 0$. But by Lemma 2.4, $\alpha \neq 0$ almost everywhere, which completes the proof. $\qquad\square$

A *convex body* $K \subset \mathbf{R}^2$ is a compact convex set with interior points. We say that a planar curve $\gamma\colon [a,b] \to \mathbf{R}^2$ is *locally convex* provided that it is locally one-to-one and each point $t \in [a,b]$ has an open neighborhood $U \subset [a,b]$ such that $\gamma(U)$ lies on the boundary of a convex body $K \subset \mathbf{R}^2$. A *local supporting line* ℓ for γ at t is a line passing through $\gamma(t)$ with respect to which $\gamma(U)$ lies on one side. If ℓ does not pass through o and $\gamma(U)$ lies on the side of ℓ which contains o, then we say that ℓ lies *above* γ. Finally, if γ is locally convex and through each point of it there passes a local support line which lies above γ, then we say that γ is locally convex *with respect to* o.

Proposition 4.3. *Let* $\gamma\colon [a,b] \to \mathbf{R}^3$ *be a minimal inspection curve with constant speed. Then* $\widetilde{\gamma}$ *is locally convex with respect to* o.

Proof. By Lemma 4.2, every $t \in [a,b]$ has a neighborhood $U \subset [a,b]$ such that $\widetilde{\gamma}$ is one-to-one on U. Furthermore, assuming that U is small, $\widetilde{\gamma}(U)$ will be star-shaped with respect to o, i.e., for every $s \in U$ the line passing through o and $\widetilde{\gamma}(s)$ intersects $\widetilde{\gamma}(U)$ only at $\widetilde{\gamma}(s)$. Thus connecting the end points of $\widetilde{\gamma}(\overline{U})$ to o yields a simple closed curve, say Γ. We call

the segments which run between o and end points of $\widetilde{\gamma}(\overline{U})$ the sides of Γ, and let θ denote the interior angle of Γ at o. We may assume that U is so small that $\theta \leq \pi$. Then we claim that the region K bounded by Γ is convex, which will complete the proof. To this end let p_0, $p_1 \in \text{int}(K)$. There exist a curve $p\colon [0, 1] \to \text{int}(K)$ with $p(0) = p_0$, and $p_1 = p(1)$, since $\text{int}(K)$ is path connected by Jordan curve theorem. Let $\overline{t} \in [0, 1]$ be the supremum of all points $t \in [0, 1]$ such that the line segment $p(0)p(t) \subset \text{int}(K)$. If $\overline{t} = 1$, for all pairs of points p_0, $p_1 \in \text{int}(K)$, then the line segment $p(0)p(1) \subset \text{int}(K)$. So $\text{int}(K)$ is convex, which implies that K is convex, and we are done.

Suppose then, towards a contradiction, that $\text{int}(K)$ is not convex. Then $\overline{t} < 1$ for some pair of points p_0, $p_1 \in \text{int}(K)$. Note also that $\overline{t} > 0$ since $p_0 \in \text{int}(K)$. Consequently an interior point x of $p(0)p(\overline{t})$ intersects $\partial K = \Gamma$, while $p(0)p(\overline{t}) \subset K$. Since $\theta \leq \pi$, x cannot lie on a side of Γ, for then either $p(0)$ or $p(1)$ will be forced to lie on a side of Γ as well, which is not possible as they are interior points of K. So x must lie on $\widetilde{\gamma}(U)$. Now we may slightly perturb the segment $p(0)p(\overline{t})$ so that a point of it leaves K while its end points remain in $\text{int}(K)$. Then we obtain a line segment σ whose end points lie on $\widetilde{\gamma}(U)$ while its interior lies outside K. Thus if we replace the segment of $\widetilde{\gamma}(U)$ which lies between the end points of σ with the line segment σ, we obtain a star-shaped curve $\widetilde{\beta}$ with $L(\widetilde{\beta}) < L(\widetilde{\gamma})$.

Parameterize $\widetilde{\beta}$ by letting $\widetilde{\beta}(t)$ be the point where the ray generated by $\widetilde{\gamma}(t)$ intersects $\widetilde{\beta}$. Then $|\widetilde{\beta}(t)| \geq |\widetilde{\gamma}(t)|$. Now set

$$\beta(t) := \frac{|\widetilde{\beta}(t)|}{|\widetilde{\gamma}(t)|}\gamma(t).$$

Then $\widetilde{\beta}$ is the unfolding of β. So $L(\beta) = L(\widetilde{\beta}) < L(\widetilde{\gamma}) = L(\gamma)$. On the other hand, note that $o \in \text{conv}(\beta)$; otherwise there exists $u \in \mathbf{S}^2$ such that $\langle \beta(t), u \rangle > 0$ for all $t \in [a, b]$, which in turn yields that $\langle \gamma(t), u \rangle > 0$, which is not possible since $o \in \text{conv}(\gamma)$. So $o \in \text{conv}(\beta)$ which yields that $\lambda\beta(t) \in \text{conv}(\beta)$ for all $0 \leq \lambda \leq 1$. In particular $\gamma \subset \text{conv}(\beta)$. So it follows that $\text{conv}(\beta) \supset \text{conv}(\gamma) \supset \mathbf{S}^2$. Thus β is an inspection curve

shorter than γ, which is the desired contradiction. $\square$

CHAPTER 5

SPIRAL DECOMPOSITION OF THE UNFOLDING

Using the local convexity property established in the last section, we will show here that the unfolding of a minimal inspection curve admits a partition into certain segments we call spirals. First we note that a locally convex curve $\gamma\colon [a,b] \to \mathbf{R}^2$ is rectifiable, and therefore may be reparameterized with constant speed C. Then $|\gamma'| = C$ at almost all differentiable points of γ by Lemma 2.1; however, local convexity ensures more:

Lemma 5.1. *Let* $\gamma\colon [a,b] \to \mathbf{R}^2$ *be a locally convex curve with constant speed* C. *Then one sided derivatives of* γ *exist at all points. Furthermore,* $|\gamma'_+(a)| = |\gamma'_-(b)| = |\gamma'_\pm(t)| = C$ *for all* $t \in (a,b)$.

Proof. Existence of one sided derivatives of γ follows from existence of one-sided derivatives for convex functions [32, Thm. 1.5.4]. Since γ has constant speed C, for all $t \in (a,b)$,

$$|\gamma'_\pm(t)| = \lim_{s \to t^\pm} \frac{|\gamma(s) - \gamma(t)|}{|s - t|} = \lim_{s \to t^\pm} \frac{|\gamma(s) - \gamma(t)|}{L(\gamma|_{[t,s]})} \frac{L(\gamma|_{[t,s]})}{|t - s|} = C \lim_{s \to t^\pm} \frac{|\gamma(s) - \gamma(t)|}{L(\gamma|_{[t,s]})}.$$

It is not hard to show that, due to local convexity,

$$\lim_{s \to t^\pm} \frac{|\gamma(s) - \gamma(t)|}{L(\gamma|_{[t,s]})} = 1,$$

which establishes the desired result for $t \in (a,b)$. The cases where $t = a$ or $t = b$ follow similarly, where we will only have $s \to t^+$ or $s \to t^-$ respectively. $\qquad\square$

Let $\gamma\colon [a,b] \to \mathbf{R}^2 \setminus \{o\}$ be a locally convex curve with constant speed. If $t \in (a,b)$ is a differentiable point of γ then $\gamma'_+(t) = \gamma'_-(t) = \gamma'(t)$. Thus the above lemma shows that $|\gamma'(t)| \neq 0$ at all differentiable points of γ, since $C = L(\gamma)/(b-a) > 0$. In particular $\alpha := \angle(\gamma, \gamma')$ will be well defined at all differentiable points of γ. We also count a, b

18

among differentiable points of γ, and set $\gamma'(a) := \gamma'_+(a)$, $\gamma'(b) := \gamma'_-(b)$. We say that γ is a *spiral* provided that (i) γ is locally convex with respect to o, (ii) $|\gamma|$ is nondecreasing, (iii) $\alpha(a) = \pi/2$, and (iv) $|\gamma(a)| \geq 1$. Note in particular that condition (ii), via (Equation 3.4), implies that

$$\alpha(t) \leq \frac{\pi}{2} \tag{5.1}$$

at all differentiable points $t \in [a, b]$ of γ. We say that γ is a *strict* spiral if $|\gamma|$ is increasing. By a *spiral decomposition* of a constant speed curve $\gamma\colon [a, b] \to \mathbf{R}^2$ we mean a collection U_i of mutually disjoint open subsets of $[a, b]$ such that (i) $\gamma|_{\overline{U}_i}$ is a strict spiral, after switching the direction of γ if necessary, and (ii) $|\gamma|' = 0$ almost everywhere on $[a, b] \setminus \cup_i \overline{U}_i$. By a *parameter shift* we mean replacing t with $(t + x) \bmod (b - a)$ for some $x \in [a, b]$. The main result of this is:

Proposition 5.2. *Let* $\gamma\colon [a, b] \to \mathbf{R}^3$ *be a minimal inspection curve. Then the unfolding of* γ *admits a spiral decomposition, after a parameter shift.*

We may assume, after a reparameterization, that γ has constant speed. Let $\widetilde{\gamma}$ be the unfolding of γ and $x \in (a, b)$ be a local minimum point of the height function $|\gamma| = |\widetilde{\gamma}|$. Then $\widetilde{\gamma}$ is locally supported from below by a circle of radius $|\widetilde{\gamma}(x)|$ centered at o. Thus, since $\widetilde{\gamma}$ is locally convex with respect to o, there can pass only one local support line of $\widetilde{\gamma}$ through $\widetilde{\gamma}(x)$. Consequently $\widetilde{\gamma}$ is differentiable at x [32, Thm. 1.5.15]. Furthermore, note that the local support line at $\widetilde{\gamma}(x)$ must be orthogonal to $\widetilde{\gamma}(x)$, since x is a local minimum of $|\widetilde{\gamma}|$. So $\langle \widetilde{\gamma}'(x), \widetilde{\gamma}(x) \rangle = 0$. Now if we shift the parameter of $\widetilde{\gamma}$ by x, it follows that $\widetilde{\gamma}$ is orthogonal to $\widetilde{\gamma}(a)$ and $\widetilde{\gamma}(b)$. Thus to prove Proposition 5.2 it suffices to show:

Lemma 5.3. *Let* $\gamma\colon [a, b] \to \mathbf{R}^2$ *be a constant speed curve which is locally convex with respect to* o. *Suppose that* $\alpha(a) = \pi/2 = \alpha(b)$, *and* $|\gamma| \geq 1$. *Then* γ *admits a spiral decomposition.*

To prove this lemma first recall that, as we had mentioned at the end of chapter 3, the height function $|\gamma|$ of a constant speed curve is Lipschitz. In particular $|\gamma|$ is absolutely continuous

19

and so it satisfies the fundamental theorem of calculus:

$$|\gamma(s)| - |\gamma(t)| = \int_t^s |\gamma|'(t)dt \qquad (5.2)$$

for every pair of points $s < t \in [a, b]$. Furthermore, let us reiterate that if the speed of γ is equal to C then

$$\alpha(t) = \cos^{-1}\left(\frac{|\gamma|'(t)}{C}\right), \qquad (5.3)$$

almost everywhere.

Proof of Lemma 5.3. Let X be the set of points $t \in [a, b]$ such that γ has a local support line at $\gamma(t)$ which is orthogonal to $\gamma(t)$. Then it follows from (Equation 5.3) that $|\gamma|' = 0$ almost everywhere on X. Also note that X is closed, since the limit of any sequence of support lines of a convex body is a support line. Consequently each (connected) component U of $[a, b] \setminus X$ is an open subinterval of $[a, b]$. We claim that $\gamma|_{\overline{U}}$ is a strict spiral, possibly after switching the direction of $\gamma|_{\overline{U}}$, which will complete the proof.

To establish the above claim first note that by (Equation 5.3), $|\gamma|'$ cannot vanish at any differentiable point of $|\gamma|$ on U, for any such point would belong to X. We will show that either $|\gamma|' > 0$ almost everywhere on U or else $|\gamma|' < 0$ almost everywhere on U. To this end we start by orienting each local support line ℓ of γ consistent with the orientation of γ at the point of contact with ℓ, so that the *angle* between any support line ℓ and the position vector of its point of contact will be consistently defined along γ. Now suppose, towards a contradiction, that there are subsets X and Y of U with nonzero measure such that $|\gamma|' > 0$ on X and $|\gamma|' < 0$ on Y. Since $X \cup Y$ is dense in U, there exists a point $r \in U$ which is a limit both of X and Y. More specifically, there are sequences of differentiable points t_i, s_i converging to r such that $|\gamma|'(t_i) > 0$ and $|\gamma|'(s_i) < 0$, which in turn implies that $\alpha(t_i) < \pi/2$ and $\alpha(s_i) > \pi/2$ by (Equation 5.3). Hence, since the limit of support lines to a convex body is a support line, there exists a local support line through r which makes an angle $\leq \pi/2$ with r, and there also exists a support line through r which makes

an angle $\geq \pi/2$ with r. So we conclude that there exists a local support line orthogonal to r, which is not possible by definition of U. Hence $|\gamma|'$ is always positive or always negative at differentiable points of $|\gamma|$ in U as claimed.

Now it follows from (Equation 5.2) that $|\gamma|$ is strictly monotone on $\overline{U}$. Next, let t_0 be the boundary point of $\overline{U}$ which forms the minimum point of $|\gamma|$ on $\overline{U}$. We have to show that $\gamma|_{\overline{U}}$ is orthogonal to $\gamma(t_0)$. If $t_0 = a$, b this already holds by assumption. So suppose that $t_0 \in (a, b)$. Then $t_0 \in X$, and so γ has a local support line ℓ at $\gamma(t_0)$ which is orthogonal to $\gamma(t_0)$. Since γ is locally convex with respect to o, locally γ lies below ℓ. On the other hand, since t_0 is the minimum point of $|\gamma|$ on $\overline{U}$, then, near $\gamma(t_0)$, $\gamma|_{\overline{U}}$ lies above the circle S with radius $|\gamma(t_0)|$ centered at o. Thus $\gamma|_{\overline{U}}$ must be orthogonal to $\gamma(t_0)$ as desired. We conclude then that $\gamma|_{\overline{U}}$ is a spiral, after switching the direction of $\gamma|_{\overline{U}}$ if necessary, so that $\gamma(t_0)$ becomes its initial point. $\qquad \square$

CHAPTER 6

FORMULAS FOR EFFICIENCY OF LINE SEGMENTS

Here we derive a number of formulas for the horizon, and therefore efficiency, of line segments, which may be checked using [8]. Suppose that we have a line segment $p_0 p_1$ (with $p_0 \neq p_1$) such that the line generated by $p_0 p_1$ avoids $\mathrm{int}(B^3)$, see Figure 6.1. For
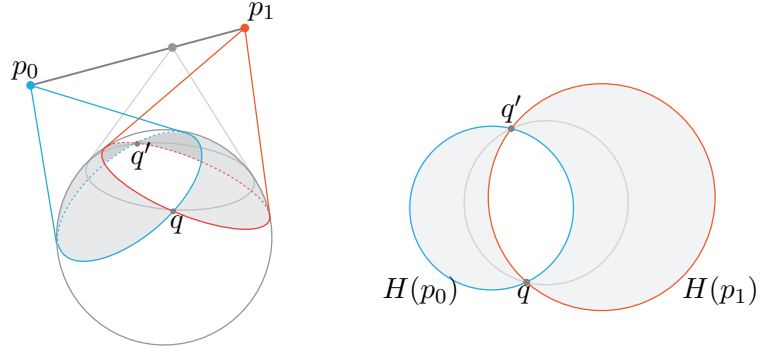

Figure 6.1: The geometric representation of the horizon of line segments

each point p on $p_0 p_1$, let C_p be the (inspection) cone generated by all rays which emanate from p and pass through a point of B^3. Let $H(p)$ be the set of points where ∂C_p touches $\mathbf{S}^2$, i.e., the *horizon circle* from the point of view of p (as had been mentioned earlier in the proof of Proposition 3.1). Then $H(p_0 p_1)$ is the area of the union of all horizon circles $H(p)$. Let $C_i := C_{p_i}$, $H_i := H(p_i)$, and $\{q, q'\} := H_0 \cap H_1$; it is possible that $q = q'$ which happens precisely when the line through p_0 and p_1 is tangent to $\mathbf{S}^2$. Note that all horizon circles $H(p)$ pass through q and q', because the triangles $p_0 p_1 q$ and $p_0 p_1 q'$ lie on planes which are tangent to $\mathbf{S}^2$. Thus $H(p_0 p_1)$ consists of the two lunar regions determined by H_0 and H_1, if $q \neq q'$; otherwise, $H(p_0 p_1)$ is the region lying inside one of the circles and outside the other. More precisely, if D_i denote the (inspection) disks in $\mathbf{S}^2$ bounded by H_i, which lie inside C_i, then

$$H(p_0 p_1) = A(D_0) + A(D_1) - 2A(D_0 \cap D_1), \tag{6.1}$$

where A stands for area. Now we may use basic spherical trigonometry to compute $H(p_0 p_1)$ as follows. To start, note that if we set

$$h_i := |p_i|, \quad \text{and} \quad \ell := |p_0 p_1|$$

then the radii of D_i and the distance in $\mathbf{S}^2$ between the centers of D_i are given respectively by

$$\rho_i := \cos^{-1}\left(\frac{1}{h_i}\right), \quad \text{and} \quad d := \angle(p_0, p_1) = \cos^{-1}\left(\frac{h_0^2 + h_1^2 - \ell^2}{2 h_0 h_1}\right).$$

It is a basic fact that

$$A(D_i) = 4\pi \sin^2\left(\frac{\rho_i}{2}\right).$$

Furthermore it is known that [33], if H_0 and H_1 intersect, then:

$$
\begin{aligned}
A(D_0 \cap D_1) = 2\Bigg(&\cos^{-1}\left(\frac{\cos(\rho_0)\cos(\rho_1) - \cos(d)}{\sin(\rho_0)\sin(\rho_1)}\right) \\
&- \cos^{-1}\left(\frac{\cos(\rho_1) - \cos(d)\cos(\rho_0)}{\sin(d)\sin(\rho_0)}\right)\cos(\rho_0) \\
&\qquad - \cos^{-1}\left(\frac{\cos(\rho_0) - \cos(d)\cos(\rho_1)}{\sin(d)\sin(\rho_1)}\right)\cos(\rho_1)\Bigg).
\end{aligned}
$$

Substituting these formulas in (Equation 6.1) yields the following formula for $H(p_0 p_1)$:

$$
\begin{aligned}
H(h_0, h_1, \ell) = 4\Bigg(\frac{1}{h_0}\sin^{-1}&\left(\frac{h_1^2 - h_0^2 - \ell^2}{\sqrt{(h_0^2 - 1)\left((h_0 + h_1)^2 - \ell^2\right)\left(\ell^2 - (h_1 - h_0)^2\right)}}\right) + \\
\frac{1}{h_1}\sin^{-1}\left(\frac{h_0^2 - h_1^2 - \ell^2}{\sqrt{(h_1^2 - 1)\left((h_0 + h_1)^2 - \ell^2\right)\left(\ell^2 - (h_1 - h_0)^2\right)}}\right) &+ \cos^{-1}\left(\frac{h_0^2 + h_1^2 - \ell^2 - 2}{2\sqrt{(h_0^2 - 1)(h_1^2 - 1)}}\right)\Bigg).
\end{aligned}
$$

$$(6.2)$$

Furthermore note that

$$h_1 = \sqrt{h_0^2 + \ell^2 + 2 h_0 \ell \cos(\alpha)}, \quad \text{where} \quad \alpha := \angle(p_0, p_0 p_1).$$

23

Then we obtain an alternative expression for $H(p_0 p_1)$:

$$H(h_0, \ell, \alpha) := 4\left(\cos^{-1}\left(\frac{h_0^2 + h_0 \ell \cos(\alpha) - 1}{\sqrt{(h_0^2 - 1)(h_1^2 - 1)}} \right) \right.$$
$$\left. + \frac{1}{h_0} \sin^{-1}\left(\frac{\cot(\alpha)}{\sqrt{h_0^2 - 1}} \right) - \frac{1}{h_1} \sin^{-1}\left(\frac{h_0 \cos(\alpha) + \ell}{h_0 \sin(\alpha)\sqrt{h_1^2 - 1}} \right) \right).$$

In particular, if $\alpha = \pi/2$, i.e., $p_0 p_1$ is orthogonal to the position vector of its initial vertex p_0, then we obtain a formula for one-edge spirals:

$$H(h_0, \ell) = 4\left(\cos^{-1}\left(\frac{\sqrt{h_0^2 - 1}}{\sqrt{h_0^2 + \ell^2 - 1}} \right) - \frac{1}{\sqrt{h_0^2 + \ell^2}} \sin^{-1}\left(\frac{\ell}{h_0\sqrt{h_0^2 + \ell^2 - 1}} \right) \right).$$
$$(6.3)$$

Finally, if we set $\ell = \sqrt{h_1^2 - h_0^2}$ in the last expression we obtain another formula for the horizon of one-edge spirals

$$H(h_0, h_1) := 4\left(\cos^{-1}\left(\frac{\sqrt{h_0^2 - 1}}{\sqrt{h_1^2 - 1}} \right) - \frac{1}{h_1} \sin^{-1}\left(\frac{\sqrt{h_1^2 - h_0^2}}{h_0\sqrt{h_1^2 - 1}} \right) \right). \qquad (6.4)$$

The graph of the corresponding efficiency function $E(h_0, h_1) := H(h_0, h_1)/\sqrt{h_1^2 - h_0^2}$, for $h_1 \geq h_0 \geq 1$ is shown in Figure 6.2. Note that $E(h_0, h_1) \leq 2$, and equality holds only
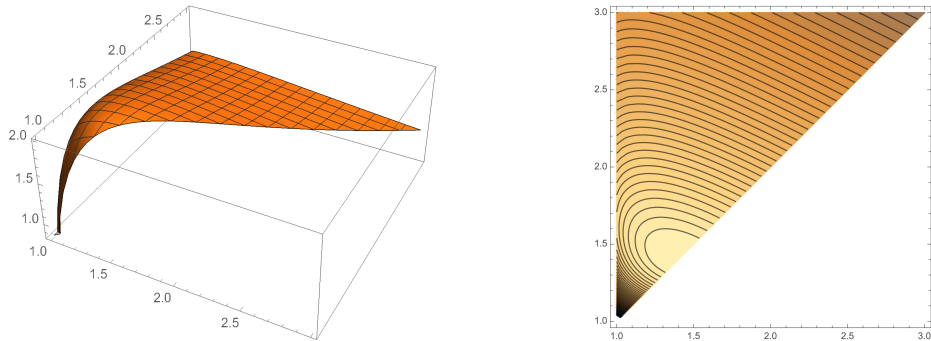

Figure 6.2: Graph of the efficiency function of a spiral segment by initial and final heights

when $h_0 = h_1 = \sqrt{2}$, or the spiral has constant height $\sqrt{2}$. Below we will prove that all spirals satisfy these properties.

CHAPTER 7

UPPER BOUND FOR EFFICIENCY OF SPIRALS

Here we apply the formulas derived in the last to show that the efficiency of spirals is bounded above by 2, via a variational argument applied to polygonal curves. First we need to show that spirals form a locally compact space and the efficiency functional is continuous on that space. To this end we start by extending the definition of a spiral as follows. We say that $\gamma\colon [a, b] \to \mathbf{R}^2$ is a *(generalized) spiral* provided that either γ is a spiral as defined earlier in chapter 5, or else γ is a constant map with $|\gamma| \geq 1$. We also extend the definition of efficiency by setting

$$E(\gamma) := \frac{4\sqrt{|\gamma|^2 - 1}}{|\gamma|^2}, \qquad \text{when} \qquad L(\gamma) = 0. \tag{7.1}$$

So it follows from Proposition 3.1 that, when $L(\gamma) = 0$, $E(\gamma)$ is the efficiency of a curve of constant distance $|\gamma|$ from the origin. Note that then $E(\gamma) \leq 2$, and $E(\gamma) = 2$ only when $|\gamma| = \sqrt{2}$. The space of spirals $\gamma\colon [a, b] \to \mathbf{R}^2$, with the topology induced on it by the uniform metric (Equation 2.2), will be denoted by $\mathcal{S}([a, b])$. To show that E is continuous on $\mathcal{S}([a, b])$ first we observe that:

Lemma 7.1. *Let* $\gamma\colon [a, b] \to \mathbf{R}^2$ *be a constant speed spiral with* $L(\gamma) \neq 0$. *Then*

$$\alpha(t) \geq \sin^{-1}\left(\frac{|\gamma(a)|}{|\gamma(t)|}\right), \tag{7.2}$$

at all differentiable points $t \in [a, b]$ *of* γ.

Proof. We may assume, for convenience, that the speed of γ is one. Then taking the cosine

of both sides of (Equation 7.2) and squaring yields

$$\langle \gamma, \gamma' \rangle^2 \leq |\gamma|^2 - r^2. \tag{7.3}$$

Recall that by (Equation 5.1), $\alpha(t) \leq \pi/2$ which in turn yields that $\langle \gamma, \gamma' \rangle \geq 0$. Thus (Equation 7.3) is equivalent to (Equation 7.2). To establish (Equation 7.3), first assume that γ is $\mathcal{C}^{1,1}$. Then the left hand side of (Equation 7.3) is Lipschitz; therefore, it is differentiable almost everywhere and satisfies the fundamental theorem of calculus. Furthermore, since γ is locally convex with respect to o, $\langle \gamma, \gamma'' \rangle \leq 0$ almost everywhere. So, since $\langle \gamma'(a), \gamma(a) \rangle = 0$, and $\langle \gamma, \gamma' \rangle \geq 0$,

$$\langle \gamma(t), \gamma'(t) \rangle^2 = 2 \int_a^t \langle \gamma(s), \gamma'(s) \rangle \big(1 + \langle \gamma(s), \gamma''(s) \rangle\big) ds \leq 2 \int_a^t \langle \gamma(s), \gamma'(s) \rangle ds = |\gamma(t)|^2 - r^2,$$

as desired. To establish the general case we consider the outer parallel curves γ_ε of γ at distance $\varepsilon > 0$. These curves are given by setting $\gamma_\varepsilon(a) := \gamma(a) + \varepsilon \gamma(a)/|\gamma(a)|$, and requiring that γ_ε maintain constant distance ε from γ. Since γ is locally convex, γ_ε is $\mathcal{C}^{1,1}$ [34, Prop. 2.4.3]. Furthermore, it is not difficult to see that γ_ε is a spiral. So γ_ε satisfies (Equation 7.3). Next note that for each differentiable point $\gamma(t)$ of γ there exists a unique point $\gamma_\varepsilon(t_\varepsilon)$ of γ_ε which is closest to $\gamma(t)$. Then $\alpha(t) = \alpha_\varepsilon(t_\varepsilon)$ where $\alpha_\varepsilon := \angle(\gamma_\varepsilon, \gamma'_\varepsilon)$. Thus

$$\alpha(t) = \alpha_\varepsilon(t_\varepsilon) \geq \sin^{-1}\left(\frac{r + \varepsilon}{|\gamma_\varepsilon(t_\varepsilon)|} \right).$$

Letting $\varepsilon \to 0$ completes the proof. $\qquad\square$

Since for a spiral $\gamma(t)$ with $L(\gamma) \neq 0$, $\alpha(t) \leq \pi/2$, the last lemma shows that $\alpha(t) \to \pi/2$ as $\gamma(t) \to \gamma(a)$. This observation, together with some basic convex analysis, yields:

Lemma 7.2. *The efficiency functional E is continuous on the space of spirals $\mathcal{S}([a, b])$.*

Proof. For convenience we may assume that $[a, b] = [0, 1]$. Let $\gamma_k \colon [0, 1] \to \mathbf{R}^2$ be a sequence of spirals converging to a spiral $\gamma \colon [0, 1] \to \mathbf{R}^2$. We have to show that $E(\gamma_k) \to$

$E(\gamma)$. To this end, we may assume that all spirals have constant speed. First suppose that $L(\gamma) = 0$. If $L(\gamma_k) = 0$ as well, then we obtain the desired result by (Equation 7.1). So we may assume that $L(\gamma_k) > 0$, by passing to a subsequence. Then, by Proposition 3.1

$$E(\gamma_k) = \int_0^1 \int_0^{2\pi} \frac{1}{|\gamma_k(t)|^2} \left| \sqrt{|\gamma_k(t)|^2 - 1} \sin\big(\alpha_k(t)\big) \cos(\theta) + \cos\big(\alpha_k(t)\big) \right| d\theta dt. \quad (7.4)$$

Note that $\gamma_k(a) \to \gamma(a)$ and $\gamma_k(t) \to \gamma(t) = \gamma(a)$. So $\gamma_k(t) \to \gamma_k(a)$. Consequently, by Lemma 7.1, $\alpha_k(t) \to \pi/2$. So, since the integrand in (Equation 7.4) is bounded, the dominated convergence theorem yields that

$$E(\gamma_k) \to \int_0^1 \int_0^{2\pi} \frac{\sqrt{|\gamma(a)|^2 - 1}}{|\gamma(a)|^2} |\cos(\theta)| \, d\theta dt = 4 \frac{\sqrt{|\gamma(a)|^2 - 1}}{|\gamma(a)|^2} = E(\gamma),$$

as desired. Next suppose that $L(\gamma) > 0$, then we may assume that $L(\gamma_k) > 0$ as well. So, again (Equation 7.4) holds. By assumption $\gamma_k \to \gamma$ uniformly. Furthermore, since γ and γ_k are locally convex, it follows that $\gamma_k' \to \gamma'$ almost everywhere on $[0,1]$. This can be shown by representing γ_k, γ locally as graphs of convex functions and applying well-known results on convergence of derivatives from classical convexity theory; e.g., see [35, C(9), p. 20], [36, Lem. 2], or [37]. Alternatively, one could give a more direct geometric argument as follows. Let ℓ_k be the tangent line of γ_k at $\gamma_k(t)$. Then ℓ_k is a local support line of γ_k and so it converges to a local support line ℓ of γ at $\gamma(t)$. But γ is differentiable at $\gamma(t)$. Thus ℓ must be the tangent line of γ at $\gamma(t)$. So $\alpha_k \to \alpha$ almost everywhere on $[0,1]$. Thus by the dominated convergence theorem

$$E(\gamma_k) \to \int_0^1 \int_0^{2\pi} \frac{1}{|\gamma(t)|^2} \left| \sqrt{|\gamma(t)|^2 - 1} \sin\big(\alpha(t)\big) \cos(\theta) + \cos\big(\alpha(t)\big) \right| d\theta dt = E(\gamma),$$

which completes the proof. □

Next we use the above results to bound the efficiency of polygonal spirals. A *polygonal curve P* is a collection of line segments determined by a sequence of points $p_0, \ldots, p_m \in$

27

$\mathbf{R}^2$ which are successively distinct, i.e. $p_{i+1} \neq p_i$ for $i = 0, \ldots, m - 1$. We also allow for the possibility that P may be degenerate, i.e., consist of a single point. We use the formal notation $P = (p_0, \ldots, p_m)$ to specify a polygonal curve. The points p_i are called *vertices* of P, and the segments $p_i p_{i+1}$, form the *edges* of P. Each polygonal curve P admits a unique constant speed parameterization $\gamma_P \colon [0, 1] \to P$, with $\gamma_p(0) = p_0$ which traces the edges of P. The *distance* between a pair of polygonal curves P^1, P^2 is defined as $\mathrm{dist}\left(\gamma_{P^1}, \gamma_{P^2}\right)$, the uniform metric defined by (Equation 2.2). Let $\mathcal{P}^m$ denote the space of polygonal curves with at most m edges in $\mathbf{R}^2$. We endow $\mathcal{P}^m$ with the topology induced by dist. Then any sequence of polygonal curves $P_i \in \mathcal{P}^m$ which is confined to a bounded region of $\mathbf{R}^2$ will have a converging subsequence. So $\mathcal{P}^m$ is locally compact. We say that $P \in \mathcal{P}^m$ is a *polygonal spiral* provided that γ_P is a spiral. Note that a polygonal spiral is always strict. Let $\mathcal{S}^m$ be the collection of polygonal spirals with at most m edges.

Lemma 7.3. *The space of polygonal spirals $\mathcal{S}^m$ is locally compact, for every $m \geq 0$.*

Proof. Since the space of polygonal curves $\mathcal{P}^m$ is locally compact, it is enough to check that $\mathcal{S}^m$ is closed in $\mathcal{P}^m$. Let $P_k \in \mathcal{S}^m$ be a sequence of polygonal spirals converging to a polygonal curve $P \in \mathcal{P}^m$. If $L(P) = 0$, then P already belongs to $\mathcal{S}^m$ and there is nothing to prove. We may suppose then that $P = (p_0, \ldots, p_\ell)$ with $\ell > 0$. It is clear that the distance of points of P from the origin must be increasing. Furthermore, P will be locally convex by Blaschke's selection principle [27, Thm. 7.3.8] on convergence of convex bodies. Finally note that there exists a sequence of edges E_k of P_k such that the initial point of E_k converges to p_0 while the final point of E_k converges to another point of $p_0 p_1$. So E_k becomes parallel to $p_0 p_1$. Furthermore note that the initial point p_0^k of P_k converges to p_0. Thus the initial point of E_k must converge to p_0^k. So, by Lemma 7.1, E_k becomes orthogonal to p_0^k and therefore to p_0. Thus $p_0 p_1$ must be orthogonal to p_0, which completes the proof. $\qquad \square$

Lemma 7.4. *The efficiency of any polygonal spiral is at most 2.*

Proof. Fix an integer $m \geq 0$, and number $R > 1$. By Lemma 7.3 there exists a polygonal spiral $P = (p_0, \ldots, p_k)$ which maximizes E among elements of $\mathcal{S}^m$ which lie in the ball of radius R centered at o. We need to show that $E(P) \leq 2$. If P is a singleton, this is guaranteed by (Equation 7.1). So we may assume that $k \geq 1$. Then $r := |p_0| < R$. Note that for $-\varepsilon < t < \varepsilon$ there exists a point p_0^t such that p_0^t is orthogonal to $p_0^t p_1$, and $|p_0^t| = (1 + t)r$, assuming that ε sufficiently small. Indeed p_0^t lies on an arc of the circle of radius $|p_1|/2$ which is centered at the midpoint of op_1; see Figure 7.1. Furthermore,
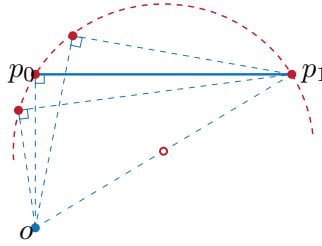
Figure 7.1: Perturbation of initial point of polygonal spiral

choosing ε sufficiently small, we can ensure that $P^t := (p_0^t, p_1, \ldots, p_k)$ is locally convex. Thus P^t will be a spiral provided that $|p_0^t| \geq 1$, which will be the case for small ε provided that $r > 1$ or else $t \geq 0$. Let us assume first that $r > 1$. Then P^t will be a spiral in $B_R(o)$ for $-\varepsilon < t < \varepsilon$. Let $L(t)$, $H(t)$, and $E(t)$ denote respectively the length, horizon, and efficiency of P^t. Then

$$0 = E'(0) = \frac{H'(0)L(0) - H(0)L'(0)}{L(0)^2},$$

which in turn yields

$$\frac{H'(0)}{L'(0)} = \frac{H(0)}{L(0)} = E(P).$$

To compute $L'(0)$ note that

$$L(t) = L(p_0^t p_1) + L\big((p_1, \ldots, p_k)\big) = \sqrt{|p_1|^2 - (r + t)^2} + L\big((p_1, \ldots, p_k)\big).$$

So it follows that

$$L'(0) = -\frac{r}{\sqrt{|p_1|^2 - r^2}}.$$

Next, to compute $H'(0)$, note that

$$H(t) = H(p_0^t p_1) + H\big((p_1, \ldots, p_k)\big).$$

Furthermore, by (Equation 6.3) we have

$$H(p_0^t p_1) = H\big(r + t, L(p_0^t p_1)\big).$$

Now a computation [8] yields that

$$H'(0) = \frac{d}{dt} H\big(r + t, L(p_0^t p_1)\big)\Big|_{t=0} = -\frac{4}{r}\frac{\sqrt{r^2 - 1}}{\sqrt{|p_1|^2 - r^2}}. \tag{7.5}$$

So we conclude that

$$E(P) = \frac{H'(0)}{L'(0)} = 4\frac{\sqrt{r^2 - 1}}{r^2} \leq 2,$$

as desired. It remains to show that our earlier assumption that $r > 1$ was justified. Suppose then, towards a contradiction, that $r = 1$. Then $E(t)$ and $H(t)$ will still be well defined for $t \geq 0$, and so will their right-hand derivatives at 0. By (Equation 7.5), $H'_+(0) = 0$. Thus

$$E'_+(0) = \frac{-H(0)L'(0)}{L(0)^2} = \frac{H(0)}{L(0)^2}\frac{1}{\sqrt{|p_1|^2 - 1^2}} > 0.$$

So $E(t) > E(0)$, or $E(P^t) > E(P)$, for small $t > 0$ which is the desired contradiction. $\square$

The above lemma together with Lemma 7.2 now yields the main result of this section:

Proposition 7.5. *The efficiency of any spiral is at most* 2*.*

Proof. Let γ be a spiral. If $L(\gamma) = 0$, then it is a polygonal spiral and so $E(\gamma) \leq 2$ by Lemma 7.4. Thus we may assume that $L(\gamma) > 0$ and $\gamma\colon [a, b] \to \mathbf{R}^2$ has unit speed. For

$k = 3, 4, \ldots$, we construct a sequence of polygonal curves $P^k = (p_0^k, \ldots, p_k^k)$ converging to γ as follows. Set $p_i^k := \gamma(iL/k)$, for $i = 1, \ldots, k-1$, and let p_0^k, p_k^k be the points on the rays $o\gamma(a)$ and $o\gamma(b)$ respectively such that p_0^k is orthogonal to $p_0^k p_1^k$ and p_k^k is orthogonal to $p_{k-1}^k p_k^k$; see Figure 7.2. Note that $E(P^k) \to E(\gamma)$ by Lemma 7.2. So, to complete
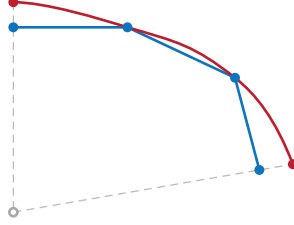
Figure 7.2: A polygonal approximation to a spiral

the proof, it suffices to show that $E(P^k) \le 2$. To see this note that since γ is a spiral, P^k is locally convex with respect to o, assuming that k is sufficiently large. Thus P^k may be partitioned into a collection of spirals by Lemma 5.3. The efficiency of each of these spirals is at most 2, by Lemma 7.4. Hence, by (Equation 1.3), $E(P^k) \le 2$ as desired. $\quad\square$

As we had mentioned in the introduction, Proposition 7.5 together with Proposition 5.2 now establishes inequality (Equation 1.1). The rest of this work will be concerned with characterizing the case where equality holds in (Equation 1.1).

CHAPTER 8

INSTANTANEOUS EFFICIENCY

Here we show that a spiral with initial height $\geq \sqrt{2}$ assumes its maximum efficiency 2 only when it has constant height $\sqrt{2}$. This argument is based on the notion of instantaneous efficiency, which was used in the proof of Proposition 3.1, and will be developed further below. Let $\gamma\colon [a,b] \to \mathbf{R}^3$ be a constant speed curve with $|\gamma| \geq 1$, and $t \in [a,b]$ be a differentiable point of γ with $|\gamma'(t)| \neq 0$. Then $\alpha(t)$ is well defined. Now we define the *instantaneous efficiency* of γ at t as

$$E_\gamma(t) := \int_0^{2\pi} \big| F(|\gamma(t)|, \alpha(t), \theta) \big| d\theta, \tag{8.1}$$

where

$$F(h, \alpha, \theta) := \frac{1}{h^2} \left(\sqrt{h^2 - 1} \sin(\alpha) \cos(\theta) + \cos(\alpha) \right).$$

Note that if t is a differentiable point of $t \mapsto H(\gamma|_{[a,t]})$, then by (Equation 3.2) $E_\gamma(t) = \frac{d}{dt} H\left(\gamma|_{[a,t]} \right)$. So $E_\gamma(t)$ is the rate of change of horizon along γ. Furthermore, by Proposition 3.1,

$$E(\gamma) = \frac{1}{b-a} \int_a^b E_\gamma(t) dt \leq \sup_{[a,b]} E_\gamma(t). \tag{8.2}$$

Thus to find an upper bound for $E(\gamma)$ it suffices to bound E_γ. To this end we derive a more explicit formula for E_γ by computing the integral in (Equation 8.1) as follows. Let

$$\Omega := \left\{ (h, \alpha) \,\Big|\, h \geq 1, \; \sin^{-1}\left(\frac{1}{h} \right) \leq \alpha \leq \frac{\pi}{2} \right\}$$

be the phase space of all possible values for pairs of heights and angles $(|\gamma(t)|, \alpha(t))$, at differentiable points of curves γ which lie outside $\mathbf{S}^2$ and whose tangent lines avoid

$\mathrm{int}(B^3)$. For every $(h, \alpha) \in \Omega$ we set

$$\theta_0 = \theta_0(h, \alpha) := \cos^{-1}\left(-\frac{\cot(\alpha)}{\sqrt{h^2 - 1}}\right).$$

Since $\sin(\alpha) \geq 1/h$, $|\cot(\alpha)/\sqrt{h^2 - 1}| \leq 1$. So θ_0 is well defined. Also note that $F(h, \alpha, \pm\theta_0) = 0$. Now we may compute that [8]

$$\begin{aligned}
\int_0^{2\pi} |F(h, \alpha, \theta)|\, d\theta &= \int_{-\theta_0}^{\theta_0} F(h, \alpha, \theta)\, d\theta - \int_{\theta_0}^{2\pi - \theta_0} F(h, \alpha, \theta)\, d\theta \\
&= \frac{4}{h^2}\left(\sqrt{h^2 - 1}\sin(\alpha)\sin(\theta_0) + \left(\theta_0 - \frac{\pi}{2}\right)\cos(\alpha)\right) \\
&= \frac{4}{h^2}\left(\sqrt{h^2 \sin^2(\alpha) - 1} + \sin^{-1}\left(\frac{\cot(\alpha)}{\sqrt{h^2 - 1}}\right)\cos(\alpha)\right).
\end{aligned}$$

Thus if we set

$$\mathcal{E}(h, \alpha) := \frac{4}{h^2}\left(\sqrt{h^2 \sin^2(\alpha) - 1} + \cos(\alpha)\sin^{-1}\left(\frac{\cot(\alpha)}{\sqrt{h^2 - 1}}\right)\right), \qquad (8.3)$$

then we obtain

$$E_\gamma(t) = \mathcal{E}\big(|\gamma(t)|, \alpha(t)\big),$$

for curves whose tangent lines avoid $\mathrm{int}(B^3)$. We also set

$$\mathcal{E}(h) := \mathcal{E}\left(h, \frac{\pi}{2}\right) = 4\frac{\sqrt{h^2 - 1}}{h^2} \leq 2.$$

Note that $\mathcal{E}(h) = 2$ only if $h = \sqrt{2}$. For any set $X \subset [a, b]$, with nonzero measure $\mu(X)$ we define

$$E(\gamma|_X) := \frac{1}{\mu(X)}\int_X E_\gamma(t)\,dt.$$

The above discussion has established that

Proposition 8.1. *Let $\gamma\colon [a, b] \to \mathbf{R}^3$ be a constant speed curve with $|\gamma| \geq 1$. Suppose that*

33

the tangent lines of γ avoid $\mathrm{int}(B^3)$. *Then, for any set $X \subset [a, b]$ with nonzero measure*

$$E(\gamma|_X) = \frac{1}{\mu(X)} \int_X \mathcal{E}\big(|\gamma(t)|, \alpha(t)\big) dt,$$

where $\mathcal{E}$ is given by (Equation 8.3).

The values of the instant efficiency function $\mathcal{E}(h, \alpha)$ on the phase space Ω are shown in Figure 8.1. These values range from 0 to about 2.6, and the red contour line in the right
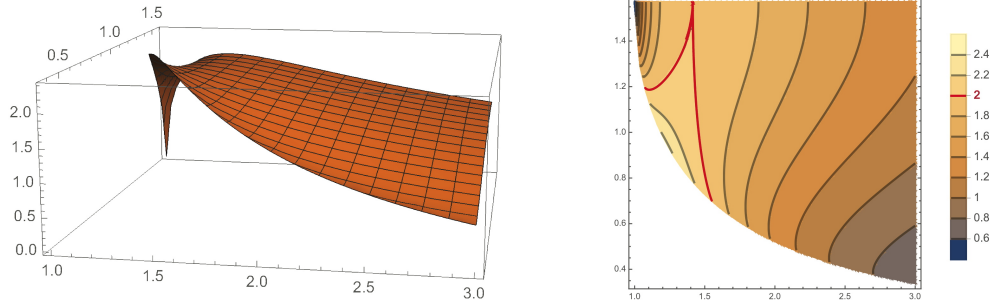
Figure 8.1: The instantaneous efficiency by height and angle

diagram corresponds to the value $\mathcal{E}(h, \alpha) = 2$. The cusp of this contour line lies at the point $(\sqrt{2}, \pi/2)$, which corresponds to the conjectured minimal curve. The region bounded by this contour line has $\mathcal{E}(h, \alpha) \geq 2$, and separates the phase space into two components where $\mathcal{E}(h, \alpha) \leq 2$. In one of these components $h \geq \sqrt{2}$ and in the other $h \leq \sqrt{2}$. This may be regarded as the reason why we need to treat the spirals with initial height above $\sqrt{2}$ separately from those with initial height below $\sqrt{2}$, which will be treated in the next section.

Now let $\gamma \colon [a, b] \to \mathbf{R}^3$ be a spiral with initial height $r := |\gamma(a)| \geq \sqrt{2}$. We will show that then the instantaneous efficiency $E_\gamma(t) \leq 2$, for almost all $t \in [a, b]$, which in turn implies that $E(\gamma) \leq 2$ by (Equation 8.2). The main idea behind the proof is as follows. Any curve $\gamma \colon [a, b] \to \mathbf{R}^3$ which lies outside $\mathbf{S}^2$ and whose tangent lines avoid $\mathrm{int}(B^3)$, generates an associated mapping

$$t \longmapsto (|\gamma(t)|, \alpha(t)) \in \Omega$$

34

where Ω is the phase space described above, and $t \in [a, b]$ are differentiable points of γ with $|\gamma'(t)| \neq 0$. So $E_\gamma(t) \leq 2$ provided that the associated mapping of γ avoids the region in Ω where $\mathcal{E}(h, \alpha) > 2$. An illustrative example is provided in the case where γ traces a straight line segment (which is orthogonal to the position vector of its initial point). Then the associated curves are given by $\alpha(t) = \sin^{-1}(r/|\gamma(t)|)$. These curves for different values of their initial height r are depicted in yellow and blue in Figure 8.2. The blue curve corresponds to $r = \sqrt{2}$. Recall that the cusp of the red contour curve, which bounds the
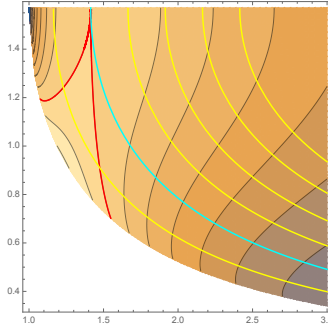
Figure 8.2: The instantaneous efficiency along with lines of equality for $\alpha = \sin^{-1}(r/|\gamma(t)|)$ for different values of r

region with $\mathcal{E}(h, \alpha) \geq 2$ has coordinates $(\sqrt{2}, \pi/2)$. Thus Figure 8.2 shows that if $r \geq \sqrt{2}$, then the instantaneous efficiency of γ is always below 2. We prove this in the next lemma. Recall that $\mathcal{E}(h) := \mathcal{E}(h, \pi/2)$.

Lemma 8.2. *For $h \geq \sqrt{2}$ and $\sin^{-1}(\sqrt{2}/h) \leq \alpha \leq \pi/2$,*

$$\mathcal{E}(h, \alpha) \leq \mathcal{E}(h) \leq 2.$$

Proof. Since $\mathcal{E}(h, \pi/2) = \mathcal{E}(h) \leq 2$, it suffices to check that $\alpha \mapsto \mathcal{E}(h, \alpha)$ is nondecreas-

ing. So we compute [8]

$$\frac{\partial \mathcal{E}}{\partial \alpha}(h, \alpha) = \frac{4}{h^2}\left(\cot(\alpha)\sqrt{h^2\sin^2(\alpha) - 1} - \sin(\alpha)\sin^{-1}\left(\frac{\cot(\alpha)}{\sqrt{h^2 - 1}}\right)\right)$$

$$\geq \frac{4}{h^2}\left(\cot(\alpha) - \sin^{-1}\left(\frac{\cot(\alpha)}{\sqrt{(\sqrt{2}/\sin(\alpha))^2 - 1}}\right)\right)$$

$$= \frac{4}{h^2}\left(\frac{\cos(\alpha)}{\sqrt{1 - \cos^2(\alpha)}} - \sin^{-1}\left(\frac{\cos(\alpha)}{\sqrt{1 + \cos^2(\alpha)}}\right)\right).$$

Now if we let $x := \cos(\alpha)$, then it remains to check that the following expression is non-negative

$$\frac{x}{\sqrt{1 - x^2}} - \sin^{-1}\left(\frac{x}{\sqrt{1 + x^2}}\right)$$

for $0 \leq x \leq 1$. To see this note that the above expression vanishes for $x = 0$. Furthermore, its derivative is given by $1/(1 - x^2)^{3/2} - 1/(x^2 + 1)$ which is indeed nonnegative for $0 \leq x \leq 1$. $\qquad\square$

The last lemma, together with Lemma 7.1 and Proposition 8.1, yields:

Proposition 8.3. *Let $\gamma\colon [a, b] \to \mathbf{R}^2$ be a constant speed spiral with initial height $r \geq \sqrt{2}$. Then*

$$E(\gamma) \leq \frac{1}{b - a}\int_a^b \mathcal{E}(|\gamma(t)|)dt \leq \mathcal{E}(r) \leq 2.$$

Equality in the second inequality holds only when $|\gamma| \equiv r$, and equality in the third inequality holds only when $r = \sqrt{2}$.

CHAPTER 9

SPIRALS WITH MAXIMUM EFFICIENCY

Here we refine the variational method employed in chapter 7 to show that the efficiency of any spiral assumes its maximum value only when it has constant height $\sqrt{2}$. We start by considering one edge spirals $P = (p_0, p_1)$. By a *lifting* of P we mean any polygonal curve $\widetilde{P} = (\widetilde{p}_0, p_1)$ where $\widetilde{p}_0 = \lambda p_0$ for $\lambda > 1$.

Lemma 9.1. *Let $P = (p_0, p_1)$ be a spiral. For any lifting $\widetilde{P}$ of P, $H(P) < H(\widetilde{P})$.*

Proof. Suppose $\widetilde{P} := (\widetilde{p}_0, p_1)$, see Figure 9.1. Since $p_0 p_1$ is orthogonal to p_0, the horizon
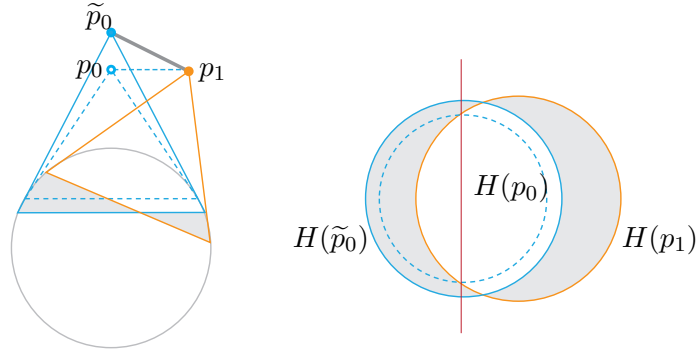
Figure 9.1: The effects of the perturbation of the initial point of a spiral segment radially on the horizon of that segment

circle $H(p_1)$ (depicted in orange) bisects the horizon circle $H(p_0)$ (depicted in dotted blue line), because the two planes which contain $p_0 p_1$ and are tangent to $\mathbf{S}^2$ intersect $H(p_0)$ at a pair of its antipodal points. Thus the area that is gained by the horizon, as p_0 rises to $\widetilde{p}_0$ exceeds the area which is lost. $\qquad\square$

Let $P = (p_0, p_1)$ be a spiral and set $r := |p_0|$, $R := |p_1|$. For every $h \in [r, R]$, let $\widetilde{P}^h := (\widetilde{p}_0^h, p_1)$ be the lifting of P such that the distance of $\widetilde{p}_0^h p_1$ to o is equal to h. Let q^h be the closest point of $\widetilde{p}_0^h p_1$ to o; see Figure 9.2. Set $\widetilde{P}_+^h := (q^h, p_1)$ and $\widetilde{P}_-^h := (\widetilde{p}_0^h, q^h)$.

Figure 9.2: The splitting of a lifted segment

Lemma 9.2. *Let $P = (p_0, p_1)$ be a spiral with initial height r, and final height R. Then for every $r \leq \rho \leq R$,*

$$H(P) \leq \int_r^\rho w(h)\mathcal{E}(h)dh + H(\widetilde{P}_+^\rho),$$

where $\int_r^\rho w(h)dh = L(P) - L(\widetilde{P}_+^\rho)$, and $w \geq r/\sqrt{R^2 - r^2}$.

Proof. If the desired inequality holds for all $r > 1$, then it also holds for $r = 1$ by continuity. So we may assume that $r > 1$. We claim that

$$w(h) := -\frac{d}{ds}L(\widetilde{P}_+^s)\Big|_{s=h} = -\frac{d}{ds}\sqrt{R^2 - s^2}\Big|_{s=h} = \frac{h}{\sqrt{R^2 - h^2}}$$

is the desired weight function. Clearly $\int_r^\rho w(h)dh = L(P) - L(\widetilde{P}_+^\rho)$ and $w \geq r/\sqrt{R^2 - r^2}$. Set $\Delta h := (\rho - r)/n$, $h_i := r + i\Delta h$, and $q^i := q^{h_i}$. We define a sequence of liftings as follows. Set $P^0 := P$. Once P^i is defined, let $\widetilde{p}_0^i$ be its initial point, q^i be its closest point to o, and set $P_-^i := (\widetilde{p}_0^i, q^i)$, $P_+^i := (q^i, p_1)$. Then we define $P^{i+1} := \left(\widetilde{P_+^i}\right)^{h_{i+1}}$; see Figure 9.3. By Lemma 9.1 $H(P_+^i) < H(P^{i+1}) = H(P_-^{i+1}) + H(P_+^{i+1})$. Applying this

Figure 9.3: The repeated lifting and splitting of a spiral segment

inequality iteratively yields

$$H(P) \leq \sum_{i=1}^{n} H(P_-^i) + H(\widetilde{P}_+^\rho).$$

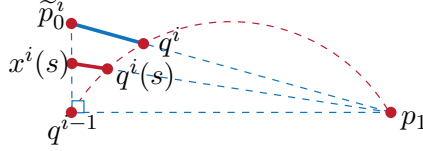Now, for $0 \leq s \leq \Delta h$, let $q^i(s) := q^{h_{i-1}+s}$; see Figure 9.4. Furthermore let $x^i(s)$ be the

Figure 9.4: Parameterizing the lifting and splitting of a segment

point where the line passing through $q^i(s)$ and p_1 intersects $q^{i-1}\widetilde{p}_0^i$. Set $\sigma_s^i := x^i(s)q^i(s)$.

Let $f_i(s) := H(\sigma_s^i)$. By (Equation 6.3), $f_i(s) = H\big(h_{i-1} + s, L(\sigma_i(s))\big)$. So f_i is $\mathcal{C}^\infty$ on $[0, \Delta h]$ provided that $h_{i-1} + s > 1$ or $h_{i-1} > 1$, which is the case since $r > 1$. We have $H(P_-^i) = f_i(\Delta h) - f_i(0) \leq f_i'(0)\Delta h + C_i(\Delta h)^2$, where $C_i := \sup_{[0,\Delta h]} f_i''(s)/2 < \infty$. Note that C_i depends continuously on q^i. So C_i are bounded above by some constant C, independent of i, which yields

$$H(P_-^i) \leq f_i'(0)\Delta h + C(\Delta h)^2 = f_i'(0)\Delta h + C\frac{(\rho - r)^2}{n^2}.$$

Next we compute that

$$f_i'(0) = \frac{d}{ds}L(\sigma_s^i)\Big|_{s=0} E(\sigma_0^i) + L(\sigma_0^i)\frac{d}{ds}E(\sigma_s^i)\Big|_{s=0} = \frac{d}{ds}L(\sigma_s^i)\Big|_{s=0}\mathcal{E}(h_{i-1}).$$

If we let $\tau_s^i := q^i(s)p_1$ then at $s = 0$, $\frac{d}{ds}L(\sigma_s^i) + \frac{d}{ds}L(\tau_s^i) = \frac{d}{ds}|x^i(s)p_1| = 0$, because $x^i(0)p_1 = q^{i-1}p_1$ is orthogonal to $q^{i-1}\widetilde{p}_0^i$. Now recall that $q^h p_1 = \widetilde{P}_+^h$. Thus

$$\frac{d}{ds}L(\sigma_s^i)\Big|_{s=0} = -\frac{d}{ds}L(\tau_s^i)\Big|_{s=0} = -\frac{d}{ds}L(\widetilde{P}_+^{h_{i-1}+s})\Big|_{s=0} = w(h_{i-1}).$$

The last four displayed expressions yield

$$\sum_{i=1}^{n} H(P_-^i) \le \sum_{i=0}^{n-1} \left(w(h_i)\mathcal{E}(h_i)\Delta h + C\frac{(\rho - r)^2}{n^2} \right) = \sum_{i=0}^{n-1} w(h_i)\mathcal{E}(h_i)\Delta h + C\frac{(\rho - r)^2}{n}.$$

Letting $n \to \infty$ completes the proof. $\qquad\square$

The last lemma via an induction yields:

Lemma 9.3. *Let P be a polygonal spiral with initial height r, and final height R. Then*

$$H(P) \le \int_r^R w(h)\mathcal{E}(h)dh, \qquad (9.1)$$

where $\int_r^R w(h)dh = L(P)$, and $w \ge r/\sqrt{R^2 - r^2}$.

Proof. If P has only one edge (Equation 9.1) holds by Lemma 9.2. Suppose that (Equation 9.1) holds for spirals with n edges and let $P = (p_0, \ldots, p_{n+1})$. Let ρ be the distance of the line spanned by p_1p_2 from o and q be the closest point of that line to the origin. Then $P' := (q, p_2, \ldots, p_m)$ is a spiral with n edges. Note that $H(P) = H\big((p_0, p_1)\big) + H\big((p_1, \ldots, p_{n+1})\big)$, and by Lemma 9.2, $H\big((p_0, p_1)\big) \le \int_r^\rho w_0(h)\mathcal{E}(h)dh + H\big((q, p_1)\big)$. Thus

$$H(P) \le \int_r^\rho w_0(h)\mathcal{E}(h)dh + H(P'),$$

where $\int_r^\rho w_0(h)dh = L(p_0p_1) - L(qp_1) = L(P) - L(P')$, and $w_0 \ge r/\sqrt{|p_1|^2 - r^2} \ge r/\sqrt{R^2 - r^2}$. By the inductive hypothesis

$$H(P') \le \int_\rho^R w_1(h)\mathcal{E}(h)dh,$$

where $\int_r^\rho w_1(h)dh = L(P')$, and $w_1 \ge \rho/\sqrt{R^2 - \rho^2} \ge r/\sqrt{R^2 - r^2}$. Set $w := w_0$ for $h < \rho$ and $w := w_1$ for $h \ge \rho$. Then the last two displayed inequalities yield (Equation 9.1)
. $\qquad\square$

Now we prove the main result of this section, which extends Proposition 7.5:

Proposition 9.4. *For any spiral γ, $E(\gamma) \leq 2$ with equality only if $|\gamma| \equiv \sqrt{2}$.*

Proof. Lemma 9.3 together with Lemma 7.2 yields $E(\gamma) \leq 2$ via a polygonal approximation. Next suppose that $E(\gamma) = 2$. Let r, R be the initial and final heights of γ. If $r = R$, then $2 = E(\gamma) = \mathcal{E}(|\gamma|)$ which yields $|\gamma| \equiv \sqrt{2}$. Suppose towards a contradiction that $r < R$. Let P_i, $i = 1, 2, \ldots$ be a sequence of polygonal spirals converging to γ, with initial and final heights r_i, R_i. We may assume for convenience that $r \leq r_i < R_i \leq R$. Let w_i be the weight functions for P_i given by Lemma 9.3. Set $\overline{w}_i := w_i/L(P_i)$ on $[r_i, R_i]$ and $\overline{w}_i := 0$ elsewhere. Then $\int_r^R \overline{w}_i(h)dh = 1$. By Lemma 7.2, for any given $\varepsilon > 0$, we may choose i so large that $E(P_i) \geq 2 - \varepsilon$. Then by Lemma 9.3,

$$2 - \varepsilon \leq E(P_i) \leq \int_r^R \overline{w}_i(h)\mathcal{E}(h)dh \leq \sup_{[r,R]} \mathcal{E} \leq 2.$$

So $\sup_{[r,R]} \mathcal{E} = 2$. Since $\mathcal{E} = 2$ only at $\sqrt{2}$, $[r, R] \ni \sqrt{2}$. So the set of heights $h \in [r, R]$ with $\mathcal{E}(h) \geq 2 - \sqrt{\varepsilon}$ forms a subinterval $[x_\varepsilon^-, x_\varepsilon^+]$. It follows that

$$2 - \varepsilon \leq \int_r^R \overline{w}_i(h)\mathcal{E}(h)dh \leq -\sqrt{\varepsilon}\left(\int_r^{x_\varepsilon^-} \overline{w}_i(h)dh + \int_{x_\varepsilon^+}^R \overline{w}_i(h)dh \right) + 2.$$

So $\int_r^{x_\varepsilon^-} \overline{w}_i(h)dh + \int_{x_\varepsilon^+}^R \overline{w}_i(h)dh \leq \sqrt{\varepsilon}$. But $\overline{w}_i \geq 1/\left(L(P_i)\sqrt{(R/r)^2 - 1}\right)$. Thus

$$R - r \leq \sqrt{\varepsilon}L(P_i)\sqrt{(R/r)^2 - 1} + x_\varepsilon^+ - x_\varepsilon^-.$$

Letting $\varepsilon \to 0$, we obtain $r = R$, since $x_\varepsilon^\pm \to \sqrt{2}$, and $L(P_i)$ is bounded above. Hence we arrive at the desired contradiction. $\qquad\square$

CHAPTER 10

PROOF OF THEOREM 1.1

By Proposition 2.3 there exists a minimal inspection curve $\gamma \colon [a, b] \to \mathbf{R}^3$, which we may assume to have constant speed by Lemma 2.1. As we described in chapter 3, to establish (Equation 1.1) it suffices to show that $E(\gamma) \leq 2$. By Proposition 4.1, $E(\gamma) = E(\widetilde{\gamma})$ where $\widetilde{\gamma}$ is the unfolding of γ. Furthermore, by Proposition 5.2, $\widetilde{\gamma}$ admits a spiral decomposition, generated by a collection of mutually disjoint open sets $U_i \subset [a, b]$, $i \in I$. Set $U_0 := [a, b] \setminus \cup_i \overline{U}_i$, and let $\widetilde{\gamma}_i := \widetilde{\gamma}|_{\overline{U}_i}$, $\widetilde{\gamma}_0 := \widetilde{\gamma}|_{U_0}$. Then

$$E(\widetilde{\gamma}) = \frac{H(\widetilde{\gamma})}{L(\widetilde{\gamma})} = \frac{1}{L(\widetilde{\gamma})} \sum_i H(\widetilde{\gamma}_i) = \frac{1}{L(\widetilde{\gamma})} \left(L(\widetilde{\gamma}_0) E(\widetilde{\gamma}_0) + \sum_i L(\widetilde{\gamma}_i) E(\widetilde{\gamma}_i) \right), \quad (10.1)$$

where we define $L(\widetilde{\gamma}_0) := \int_{U_0} |\widetilde{\gamma}_0'(t)| dt$. So $L(\widetilde{\gamma}_0) + \sum_i L(\widetilde{\gamma}_i) = L(\widetilde{\gamma})$. If $L(\widetilde{\gamma}_0) = 0$, then we may disregard the first term in the summation above. Otherwise, by definition of spiral decomposition, $\widetilde{\alpha}(t) = \pi/2$ for almost all $t \in U_0$. Thus, by Proposition 8.1,

$$E(\widetilde{\gamma}_0) = \frac{1}{\mu(U_0)} \int_{U_0} \mathcal{E}\left(|\widetilde{\gamma}(t)|, \frac{\pi}{2} \right) dt = \frac{1}{\mu(U_0)} \int_{U_0} \mathcal{E}(|\widetilde{\gamma}(t)|) dt \leq 2. \quad (10.2)$$

Furthermore, by Proposition 7.5,

$$E(\widetilde{\gamma}_i) \leq 2, \quad (10.3)$$

assuming $U_i \neq \emptyset$. So it follows that $E(\widetilde{\gamma}) \leq 2$, as desired.

It remains to characterize the case of equality in (Equation 1.1), which corresponds to $E(\gamma) = 2$. Then $E(\widetilde{\gamma}) = 2$, which yields that the terms $E(\widetilde{\gamma}_i)$ and $E(\widetilde{\gamma}_0)$ in (Equation 10.1) must all be equal to 2. But the inequality in (Equation 10.3) must be strict by Propositions 8.3 and 9.4, since $\widetilde{\gamma}_i$ are strict spirals by definition of spiral decomposition. So $\widetilde{\gamma}$ cannot contain any strict spirals or $U_i = \emptyset$, which means that $U_0 = [a, b]$ or $\widetilde{\gamma}$ has constant height.

Furthermore, equality in (Equation 10.2) implies that $\mathcal{E}(|\widetilde{\gamma}(t)|) \equiv 2$ which can happen only when $|\widetilde{\gamma}(t)| \equiv \sqrt{2}$. So we conclude that γ has constant height $\sqrt{2}$, since unfoldings preserve height.

Now let $\overline{\gamma} := \gamma/\sqrt{2}$ be the projection of γ into $\mathbf{S}^2$. Then $L(\overline{\gamma}) = L(\gamma)/\sqrt{2} = 4\pi/\sqrt{2}$. Recall that, since γ is an inspection curve, the horizon circles generated by points of γ cover $\mathbf{S}^2$. Since $|\gamma| \equiv \sqrt{2}$, these circles have (spherical) radius $\pi/4$ and are centered at points of $\overline{\gamma}$. Thus $\overline{\gamma}$ satisfies the hypothesis of the following proposition, which will complete the proof of Theorem 1.1.

Proposition 10.1. *Let $\gamma\colon [a,b] \to \mathbf{S}^2$ be a closed constant speed curve with $L(\gamma) = 4\pi/\sqrt{2}$. Suppose that the distance between any point of $\mathbf{S}^2$ and γ is at most $\pi/4$. Then γ is a simple $C^{1,1}$ curve which traces consecutively 4 semicircles of length $\pi/\sqrt{2}$.*

It remains then to establish the above proposition. To this end we need:

Lemma 10.2 (Crofton-Blaschke-Santalo [38])**.** *Let $\gamma\colon [a,b] \to \mathbf{S}^2$ be a rectifiable curve, and for every point $p \in \mathbf{S}^2$, and $0 \le \rho \le \pi/2$, let $C_\rho(p) \subset \mathbf{S}^2$ denote the circle of radius ρ centered at p. Then*

$$L(\gamma) = \frac{1}{4\sin(\rho)} \int_{p\in\mathbf{S}^2} \#\gamma^{-1}\big(C_\rho(p)\big)\, dp.$$

For the rest of this we assume that γ satisfies the hypothesis of the last proposition. Then, since $L(\gamma) = 4\pi/\sqrt{2}$, applying the last lemma with $\rho = \pi/4$ to γ yields

$$\operatorname*{Ave}_{p\in\mathbf{S}^2} \#\gamma^{-1}\big(C_{\frac{\pi}{4}}(p)\big) = \frac{1}{4\pi} \int_{p\in\mathbf{S}^2} \#\gamma^{-1}\big(C_{\frac{\pi}{4}}(p)\big) dp = \frac{1}{4\pi} L(\gamma)\, 4\sin\left(\frac{\pi}{4}\right) = 2.$$

Furthermore, note that $C_{\frac{\pi}{4}}(p)$ must intersect γ for all $p \in \mathbf{S}^2$, since the distance of p from γ cannot be bigger than $\pi/4$ by assumption. So, since γ is closed, $\#\gamma^{-1}(C_{\frac{\pi}{4}}(p)) \ge 2$ for almost all $p \in \mathbf{S}^2$. Now since the average of $\#\gamma^{-1}(C_{\frac{\pi}{4}}(p))$ is 2, it follows that

Lemma 10.3. *For almost every $p \in \mathbf{S}^2$,*

$$\#\gamma^{-1}\big(C_{\frac{\pi}{4}}(p)\big) = 2.$$

43

By a *side* of a circle C in $\mathbf{S}^2$ we mean either of the two closed disks in $\mathbf{S}^2$ bounded by C. If the radius of C is less than $\pi/2$, then the disk with radius less than $\pi/2$ will be called the *inside* of C and the other disk will be called the *outside* of C. Furthermore, by *strictly inside* or *strictly outside* we mean the interior of inside and interior of outside respectively.

Lemma 10.4. *For any point $p \in \mathbf{S}^2$, the portion of γ which lies outside $C_{\frac{\pi}{4}}(p)$ has length at least π.*

Proof. By assumption, γ intersects $C_{\frac{\pi}{4}}(-p)$, which has distance $\pi/2$ from $C_{\frac{\pi}{4}}(p)$. Furthermore, since γ is closed, there must exist at least two segments of γ which connect $C_{\frac{\pi}{4}}(-p)$ and $C_{\frac{\pi}{4}}(p)$. $\square$

For the rest of this section, we will assume that γ is reparameterized so that $[a, b] = [0, 2\pi]$, and identify $[0, 2\pi]$ with the unit circle $\mathbf{S}^1 \simeq \mathbf{R}/(2\pi\mathbf{Z})$. Furthermore we fix an orientation on $\mathbf{S}^1$. Then for every pair of distinct points t, $s \in \mathbf{S}^1$, we let $[t, s]$ denote the segment in $\mathbf{S}^1$ with end points t and s whose orientation from t to s agrees with the orientation of $\mathbf{S}^1$.

Lemma 10.5. *For every $t \in \mathbf{S}^1$, the tangent cone $T_t\gamma$ is a line.*

Proof. Let $s_i \in \mathbf{S}^1$ be a sequence of points converging to t from the left hand side (with respect to the orientation of $\mathbf{S}^1$). Since γ has non-vanishing speed, it cannot be locally constant. Thus we may assume, after passing to a subsequence, that $\gamma(s_i) \neq \gamma(t)$. Then the secant rays ℓ_i in $\mathbf{R}^3$ which emanate from $\gamma(t)$ and pass through $\gamma(s_i)$ are well-defined. Let ℓ be a limit of ℓ_i. Similarly, we can consider the secant rays ℓ_i' generated by points $s_i' \in \mathbf{S}^1$ converging to t from the right hand side, and let ℓ' be a limit of ℓ_i'. We claim that the angle between ℓ and ℓ' is π. Suppose not. Then there exist points s, $s' \in \mathbf{S}^1$ arbitrary close to t and with $(s, s') \ni t$ such that the angle between the geodesic segments $\gamma(t)\gamma(s)$ and $\gamma(t)\gamma(s')$ in $\mathbf{S}^2$ is less than π. Consequently, there exists an open set S of circles of radius $\pi/4$ in $\mathbf{S}^2$ such that for every $C \in S$ we have $\gamma(s)$, $\gamma(s')$ lie strictly inside C while $\gamma(t)$ lies strictly outside C. Thus, by Lemma 10.3, there exists a circle $C \in S$ which

intersects γ in only two points. So the portion of γ which lies outside C is a subset of $\gamma([s, s'])$. But since s and s' may be chosen arbitrarily close to t, the length of $\gamma([s, s'])$ may be arbitrarily small. Hence we obtain the desired contradiction via Lemma 10.4. So the angle between ℓ and ℓ' is π as claimed. Now since ℓ and ℓ' where arbitrary limits of the right and left secant rays of γ at t, and all these limits are tangent to $\mathbf{S}^2$, it follows that ℓ and ℓ' are unique. Hence $T_t\gamma = \ell \cup \ell'$ which completes the proof. $\square$

Now that γ has a well-defined tangent line at each point, we may talk about whether γ is tangent or transversal to any circle in $\mathbf{S}^2$. Furthermore, by the above lemma, for any point $t \in \mathbf{S}^1$, the left and right unit tangent vectors of γ given by

$$u_\gamma^\pm(t) := \lim_{s \to t^\pm} \frac{\gamma(s) - \gamma(t)}{|\gamma(s) - \gamma(t)|}$$

are well-defined with $u_\gamma^+(t) = -u_\gamma^-(t)$, where $s \to t^+$ (resp. $s \to t^-$) means that s approaches t from the right (resp. left) hand side, with respect to the orientation of $\mathbf{S}^1$.

Lemma 10.6. *Let $C \subset \mathbf{S}^2$ be a circle of radius $\pi/4$. Suppose that there exists an interval $[t, s] \subset \mathbf{S}^1$ such that $\gamma(t)$ lies on C while $\gamma((t, s])$ lies strictly inside C. Then γ is transversal to C at $\gamma(t)$.*

Proof. Suppose towards a contradiction that γ is tangent to C at $\gamma(t)$. Let C' be a circle of radius $\pi/4$ in $\mathbf{S}^2$ which passes through $\gamma(t)$ and is transversal to C at $\gamma(t)$ with $u_\gamma^+(t)$ pointing outside C'. Then there exist $r \in (t, s)$ such that $\gamma(r)$ lies strictly outside C'. Furthermore, choosing C' sufficiently close to C, we can ensure that $\gamma(s)$ lies strictly inside C'. Next, by perturbing the center of C', we may find another circle C'' of radius $\pi/4$ such that $\gamma(t)$ and $\gamma(s)$ lie strictly inside C'' while $\gamma(r)$ lies strictly outside C''. Since C'' may be chosen freely from an open set of circles in $\mathbf{S}^2$, we may assume by Lemma 10.3 that C'' intersects γ at only two points. Thus the portion of γ lying outside C'' is a subset of $\gamma([t, s])$. But $\gamma([t, s])$ can have arbitrarily small length, since we may choose s as close to t as desired. Thus we obtain a contradiction by Lemma 10.4. $\square$

We say that a circle $C \subset \mathbf{S}^2$ *supports* γ at a point p of γ provided that C passes through p and γ lies on one side of C. Furthermore, if the radius of C is less than $\pi/2$, then we assume that γ lies outside C.

Lemma 10.7. *Through each point of γ there pass a pair of support circles of radius $\pi/4$ which lie outside each other.*

Proof. Let C be one of the two circles of radius $\pi/4$ which are tangent to γ at $\gamma(t)$. Suppose towards a contradiction that there exists a point $t' \in \mathbf{S}^1$ such that $\gamma(t')$ lies strictly inside C. Let D be the disk of radius $\pi/4$ bounded by C, and I be the closure of the component of $\gamma^{-1}(\mathrm{int}(D))$ which contains t'. By Lemma 10.3 γ cannot lie entirely in C. Thus I is a proper interval in $\mathbf{S}^1$. By Lemma 10.6, γ is transversal to C at the end points of I. In particular there are points s_1, $s_2 \in \mathbf{S}^1$ close to each of the end points of I such that $\gamma(s_i)$ lie strictly outside C, and t, s_1, t', s_2 are arranged cyclically in $\mathbf{S}^1$. Now by perturbing the center of C we may find a circle C' of radius $\pi/4$ such that $\gamma(t)$ and $\gamma(t')$ lie strictly inside C', while $\gamma(s_i)$ lie strictly outside C'. It follows that C' intersects γ at least 4 times. Thus we obtain the desired contradiction via Lemma 10.3, since C' may be chosen freely from an open set of circles in $\mathbf{S}^2$. $\square$

In the terminology of [11], the conclusion of Lemma 10.7 means that γ has *double positive support*. A set X in a Riemannian manifold has *positive support* provided that for some constant $r > 0$ there passes a ball of radius r through each point of X whose interior is disjoint from X. If there are two such balls at each point of X, whose interiors are disjoint from each other, we say that X has *double* positive support. The last two lemmas now yield:

Lemma 10.8. *γ is simple.*

Proof. Suppose towards a contradiction that there are distinct points t, $s \in \mathbf{S}^1$ such that $\gamma(t) = \gamma(s)$. Let r_1, r_2 be points of $\mathbf{S}^1$ which lie in the interior of different segments of $\mathbf{S}^1$ determined by s and t, so that s, r_1, t, r_2 are cyclically arranged in $\mathbf{S}^1$. By Lemma 10.7,

there exists a circle C of radius $\pi/4$ in $\mathbf{S}^2$ which supports γ at $\gamma(t) = \gamma(s)$. Furthermore, by Lemma 10.3, γ cannot lie completely on C. So we may choose r_i so that at least one of the points $\gamma(r_1)$, $\gamma(r_2)$ lies strictly outside C. Then we may translate C to obtain a circle C' of the same radius such that $\gamma(t) = \gamma(s)$ lies strictly inside C' while $\gamma(r_i)$ lie strictly outside C'. Hence $C' \cap \gamma$ must consist of at least 4 points. Furthermore C' may be chosen from an open set of circles of radius $\pi/4$ in $\mathbf{S}^2$. Thus we obtain a violation of Lemma 10.3, which completes the proof. $\qquad\square$

For a topological hypersurface X in a Riemannian manifold, double positive support is equivalent to the notion of *positive reach* introduced by Federer [39]. Thus the last two lemmas imply that γ has positive reach, which in turn yields:

Lemma 10.9. γ *is* $\mathcal{C}^{1,1}$.

Proof. Since γ has finite length, there exists a point in $\mathbf{S}^2 \setminus \gamma$, which we may assume to be $(0, 0, 1)$ after a rotation. Let $\pi \colon \mathbf{S}^2 \setminus \{(0, 0, 1)\} \to \mathbf{R}^2$ be the stereographic projection, and set $\widetilde{\gamma} := \pi \circ \gamma$. It suffices to show that $\widetilde{\gamma}$ is $\mathcal{C}^{1,1}$. To this end note that, since stereographic projections preserve circles, and by Lemma 10.7 γ has double positive support, then $\widetilde{\gamma}$ has double positive support as well. Furthermore, by Lemma 10.8, γ is simple. So, by Jordan's curve theorem, γ has two sides in $\mathbf{S}^2$. The support circles of γ must lie in opposite sides of γ at each point; otherwise the tangent cone would be a ray (or γ would have a cusp) which is not possible by Lemma 10.5. Thus the support circles of $\widetilde{\gamma}$ must lie on the opposite sides of $\widetilde{\gamma}$ as well. Consequently $\widetilde{\gamma}$ is $\mathcal{C}^{1,1}$ by [11, Thm. 1.2]; see also [40, prop. 1.4]. $\qquad\square$

Next we observe that:

Lemma 10.10. *Let C be a support circle of γ of radius $\pi/4$. Then $C \cap \gamma$ is either a pair of antipodal points of C or else is a semicircle of C.*

Proof. We claim that (i) every closed semicircle of C intersects γ, and (ii) every open semicircle of C intersects γ in a connected set. These two properties easily imply that

$\gamma \cap C$ is either a pair of antipodal points of C, a closed semicircle of C, or the entire C. The last possibility is not allowed, because by Lemma 10.8 γ is simple; therefore, if γ covers C, it must coincide with C, which would violate Lemma 10.3. So establishing the above claims will complete the proof.

To establish the first claim suppose, towards a contradiction, that there exists a closed semicircle of C which does not intersect γ. Then moving the center of C by a small distance towards the center of that semicircle yields a circle C' of radius $\pi/4$ which is disjoint from γ. Obviously all circles of radius $\pi/4$ which are close to C' will be disjoint from γ as well, which would violate Lemma 10.3.

To establish the second claim suppose, towards another contradiction, that there exists an open semicircle S of C which intersects γ in a disconnected set. Then there exist points t_1, t_2, $s \in \mathbf{S}^1$, with $s \in (t_1, t_2)$ such that $\gamma(t_i) \in S$ while $\gamma(s)$ lies strictly outside C. Furthermore note that either $\gamma((t_2, t_1))$ lies entirely on C or not. In the former case there exist a point $s' \in (t_2, t_1)$ such that $\gamma(s')$ lies in the open semicircle of C which is disjoint from S; in the latter case there exists a point $s' \in (t_2, t_1)$ such that $\gamma(s')$ lies strictly outside C. In either case, moving the center of C by a small distance towards the midpoint of S will yield a circle C' of radius $\pi/4$ such that $\gamma(t_i)$ lie strictly inside C' while $\gamma(s)$, $\gamma(s')$ lie strictly outside C'. But t_1, s, t_2, s' are cyclically arranged in $\mathbf{S}^1$. So perturbing the center of C' yields an open set of circles of radius $\pi/4$ each of which intersects γ at least 4 times. Thus again we obtain the desired contradiction via Lemma 10.3. □

The last lemma leads to the proof of Proposition 10.1 via the notion of *nested partitions of a circle* employed in [12] as a weaker version of a device developed by Umehara and Thorbergsson [15, 16] to prove 4-vertex type theorems for closed curves. To describe this approach note that since, by Lemma 10.8, γ is simple, it bounds a topological disc $D \subset \mathbf{S}^2$. By Lemmas 10.7 and 10.9 through each point $p \in \gamma$ there passes a circle C_p of radius $\pi/4$ which lies in D. Furthermore, it follows from Lemma 10.9 that C_p is unique. Thus if we

set

$$[p] := \gamma \cap C_p, \qquad \text{and} \qquad P := \{[p] \mid p \in \gamma\},$$

then P will be a partition of γ, which is a topological circle. A partition P of a circle is said to be *nested* provided that no element of P separates the components of any other element, i.e., for every $[p] \in P$ and $q \in \gamma \setminus [p]$, $[q]$ lies in a connected component of $\gamma \setminus [p]$.

Lemma 10.11. *The partition P of γ is nested.*

Proof. Suppose that P is not nested. Then there are distinct support circles C, C' of γ of radius $\pi/4$ contained in D such that C has points in different components of $\gamma \setminus C'$. In particular neither $C \cap \gamma$ nor $C' \cap \gamma$ is connected. So by Lemma 10.10, C and C' must intersect γ in precisely two points each, say $C \cap \gamma = \{p, q\}$ and $C' \cap \gamma = \{p', q'\}$. Since D is simply connected, each of the segments pq and qp of C separate D into two components. Thus each of the segments $p'q'$ and $q'p'$ of C' must intersect each of the segments pq and qp. Furthermore, each of these intersections must occur in the interior of the segments, because the interior of each segment is disjoint from γ. Thus C and C' intersect at least 4 times. So $C = C'$, which is the desired contradiction. $\qquad\square$

Finally we need to invoke the following fact which has already been established in [12, Lem. 2.2]. A partition is called *nontrivial* provided that it contains more than one element.

Lemma 10.12. *Any nontrivial nested partition of a topological circle contains at least two elements which are connected subsets of the circle.*

So there are two distinct elements $[p_1]$, $[p_2] \in P$ such that $C_{p_i} \cap \gamma$ is a connected set. Consequently, by Lemma 10.10, $C_{p_i} \cap \gamma$ are semicircles. Thus γ contains a pair of disjoint semicircles which curve toward D. Similarly, by repeating the above argument for the other domain D' in $\mathbf{S}^2$ bounded by γ, we obtain two disjoint semicircles in γ which curve toward D'. The interior of the semicircles which curve toward D must be disjoint from the interior of the semicircles which curve toward D'. Thus all 4 semicircles have mutually

49

disjoint interiors. Finally, since each semicircle has length $\pi/\sqrt{2} = L(\gamma)/4$, it follows that the semicircles cover γ, which completes the proof of Proposition 10.1.

REFERENCES

[1] V. A. Zalgaller, "Extremal problems on the convex hull of a space curve," *Algebra i Analiz*, vol. 8, no. 3, 1–13 (Translation in St. Petersburg Math. J. **8** (1997), no. 3, 369–379), 1996.

[2] J. O'Rourke, "Shortest closed curve to inspect a sphere (question posted on mathoverflow)," *mathoverflow.net/questions/69099/shortest-closed-curve-to-inspect-a-sphere*, Jun. 2011.

[3] J.-B. Hiriart-Urruty, "Du calcul différentiel au calcul variationnel: Un aperçu de l'évolution de p. fermata nos jours," *Quadrature*, no. 70, pp. 8–18, 2008.

[4] M. Ghomi, "The length, width, and inradius of space curves," *Geom. Dedicata*, vol. 196, pp. 123–143, 2018.

[5] H. T. Croft, K. J. Falconer, and R. K. Guy, *Unsolved problems in geometry*, ser. Problem Books in Mathematics. New York: Springer-Verlag, 1991, pp. xvi+198, Unsolved Problems in Intuitive Mathematics, II, ISBN: 0-387-97506-3.

[6] S. R. Finch and J. E. Wetzel, "Lost in a forest," *Amer. Math. Monthly*, vol. 111, no. 8, pp. 645–654, 2004.

[7] M. Ghomi and J. Wenk, *Shortest closed curve to inspect a sphere*, 2020.

[8] ——, "Spherical-Inspection.nb," *Mathematica Package, available at http://www.math.gatech.edu/~ghomi/MathematicaNBs/Spherical-Inspection.nb*, 2020.

[9] J. Cantarella, R. B. Kusner, and J. M. Sullivan, "On the minimum ropelength of knots and links," *Invent. Math.*, vol. 150, no. 2, pp. 257–286, 2002.

[10] M. Ghomi, "Torsion of locally convex curves," *Proc. Amer. Math. Soc.*, vol. 147, no. 4, pp. 1699–1707, 2019.

[11] M. Ghomi and R. Howard, "Tangent cones and regularity of real hypersurfaces," *J. Reine Angew. Math.*, vol. 697, pp. 221–247, 2014.

[12] M. Ghomi, "Boundary torsion and convex caps of locally convex surfaces," *J. Differential Geom.*, vol. 105, no. 3, pp. 427–486, 2017.

[13] H. Kneser, "Neuer beweis des vierscheitelsatzes," *Christian Huygens*, vol. 2, pp. 315–318, 1922.

[14] R. C. Bose, "On the number of circles of curvature perfectly enclosing or perfectly enclosed by a closed convex oval," *Math. Z.*, vol. 35, no. 1, pp. 16–24, 1932.

[15] M. Umehara, "A unified approach to the four vertex theorems. I," in *Differential and symplectic topology of knots and curves*, ser. Amer. Math. Soc. Transl. Ser. 2, vol. 190, Providence, RI: Amer. Math. Soc., 1999, pp. 185–228.

[16] G. Thorbergsson and M. Umehara, "A unified approach to the four vertex theorems. II," in *Differential and symplectic topology of knots and curves*, ser. Amer. Math. Soc. Transl. Ser. 2, vol. 190, Providence, RI: Amer. Math. Soc., 1999, pp. 229–252.

[17] V. A. Zalgaller, "The problem of the shortest space curve of unit width," *Mat. Fiz. Anal. Geom.*, vol. 1, no. 3-4, pp. 454–461, 1994.

[18] ——, "Shortest inspection curves for a sphere," *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, vol. 299, no. Geom. i Topol. 8, pp. 87–108, 328, 2003.

[19] S. Finch, "A translation of Zalgaller's "The shortest space curve of unit width" (1994)," *arXiv:1910.02729*, 2019.

[20] I. J. Schoenberg, "An isoperimetric inequality for closed curves convex in even-dimensional Euclidean spaces," *Acta Math.*, vol. 91, pp. 143–164, 1954.

[21] P. Tilli, "Isoperimetric inequalities for convex hulls and related questions," *Trans. Amer. Math. Soc.*, vol. 362, no. 9, pp. 4497–4509, 2010.

[22] R. Bellman, "A minimization problem," *Bulletin of the AMS*, vol. 62, p. 270, 1956.

[23] P. Brass, W. Moser, and J. Pach, *Research problems in discrete geometry*. Springer, New York, 2005, pp. xii+499, ISBN: 978-0387-23815-8; 0-387-23815-8.

[24] K. Fässler and T. Orponen, "Curve packing and modulus estimates," *Trans. Amer. Math. Soc.*, vol. 370, no. 7, pp. 4909–4926, 2018.

[25] R. Norwood and G. Poole, "An improved upper bound for Leo Moser's worm problem," *Discrete Comput. Geom.*, vol. 29, no. 3, pp. 409–417, 2003.

[26] L. Cesari, "Rectifiable curves and the Weierstrass integral," *Amer. Math. Monthly*, vol. 65, pp. 485–500, 1958.

[27] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, ser. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2001, vol. 33, pp. xiv+415, ISBN: 0-8218-2129-6.

[28] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces*, ser. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 2004, vol. 25, pp. viii+133, ISBN: 0-19-852938-4.

[29] H. Federer, *Geometric measure theory*, ser. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969, pp. xiv+676.

[30] J. Choe and R. Gulliver, "The sharp isoperimetric inequality for minimal surfaces with radially connected boundary in hyperbolic space," *Invent. Math.*, vol. 109, no. 3, pp. 495–503, 1992.

[31] ——, "Density of a minimal submanifold and total curvature of its boundary," *Comm. Anal. Geom.*, vol. 25, no. 3, pp. 545–556, 2017.

[32] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, expanded, ser. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2014, vol. 151, pp. xxii+736, ISBN: 978-1-107-60101-7.

[33] A. Tovchigrechko and I. A. Vakser, "How common is the funnel-like energy landscape in protein-protein interactions?" *Protein science*, vol. 10, no. 8, pp. 1572–1583, 2001.

[34] L. Hörmander, *Notions of convexity*, ser. Progress in Mathematics. Boston, MA: Birkhäuser Boston Inc., 1994, vol. 127, pp. viii+414, ISBN: 0-8176-3799-0.

[35] A. W. Roberts and D. E. Varberg, *Convex functions*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1973, pp. xx+300, Pure and Applied Mathematics, Vol. 57.

[36] M. Tsuji, "On F. Riesz' fundamental theorem on subharmonic functions," *Tohoku Math. J. (2)*, vol. 4, pp. 131–140, 1952.

[37] I. B. Lacković, "On the behaviour of sequences of left and right derivatives of a convergent sequence of convex functions," *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz.*, no. 735-762, 19–27 (1984), 1982.

[38] L. A. Santaló, "Integral formulas in Crofton's style on the sphere and some inequalities referring to spherical curves," *Duke Math. J.*, vol. 9, pp. 707–722, 1942.

[39] H. Federer, "Curvature measures," *Trans. Amer. Math. Soc.*, vol. 93, pp. 418–491, 1959.

[40] A. Lytchak, "Almost convex subsets," *Geom. Dedicata*, vol. 115, pp. 201–218, 2005.